

Optimal designs for estimating the parameters in weighted power-mean-mixture models

R.L.J. Coetzer

*Sasol Technology Research and Development, 1 Klasie Havenga Road,
Sasolburg, 1947, South Africa.*

W.W. Focke

*Institute of Applied Materials, Department of Chemical Engineering,
University of Pretoria, South Africa.*

In the mixing of fluids, a mixture may be viewed conceptually as a hypothetical collection of fluid clusters. In this context a mixture model is defined by prescriptions for (a) estimating fluid cluster properties and (b) combining them to yield an overall mixture property. A particular flexible form is obtained from using generalized weighted-power-means with the weighting based on global mole fractions $x_i, 0 \leq x_i \leq 1, \sum_i x_i = 1, i = 1, 2, \dots, q$. Optimal designs for estimating the parameters in Scheffé S- and K-polynomials are well known. In this paper we present optimal designs for estimating the parameters in the generalized weighted-power-mean mixture models, which may be nonlinear in the pure and binary interaction parameters. We illustrate the practical value of applying optimal designs for mixture variables through design efficiencies. The designs are derived for modeling viscosity from three-component mixtures.

Key Words: Design efficiency. Mixture variables. Optimal designs. Power-mean-mixture models.

1 Introduction

Research and development in the chemical and chemical engineering disciplines rely heavily on the development of accurate empirical, semi-empirical, and theoretical equations that express mixture properties in terms of compositions and pure components attributes. Mixture models or mixing rules are applied in chemical technology to assist with the design of process plants and the optimization of product formulations. Consider a macroscopically homogeneous fluid composed of q different molecular species, $x_q \in S_q$, where S_q is the q -dimensional simplex. The mixture composition is quantified by the vector

$\underline{x} \in S_q$ of normalized weights, e.g. mole fractions. The mixture components are subject to the constraints:

$$\begin{aligned} 0 \leq x_l \leq 1, \quad l = 1, 2, \dots, q \\ \sum_l x_l = 1 \end{aligned} \tag{1}$$

Recently, Focke et al. (2007) proposed a generalized weighted double power mean of orders r and s (Focke et al., 2007) as a mixture model, i.e. for $r, s \neq 0$, it takes the form:

$$\eta = \left[\sum_{k=1}^q x_k \left(\sum_{l=1}^q x_l a_{kl}^s \right)^{r/s} \right]^{1/r} \tag{2}$$

Equation (2) is homogeneous of order 1 in the coefficients a_{kl} , that is, for all λ ,

$$m^{[r,s]}(\underline{x}, [\lambda a_{kl}]) = \lambda m^{[r,s]}(\underline{x}, [a_{kl}])$$

Special forms of model (2) are derived for specific values of r and s . For example, for $r = s = 1$, (2) reduces to the second-order Scheffé K-polynomial (Draper and Pukelsheim, 1998). Focke et al. (2007) discussed the estimation of the power-mean-mixture model for predicting viscosity from knowledge of binary behaviour. They recommended that $r = -5/6$ and $s = 1/2$ are the optimal parameters for the prediction of viscosity. However, the collection of appropriate data for the accurate estimation of the parameters in the power-mean-mixture models have not been discussed in great detail previously in the chemical engineering literature.

Specifically, optimal designs may be employed for specifying the design points that are optimal according to some criterion for estimating the parameters in the model. The design of physical experiments in mixture variables has received text book length discussions for the estimation of models linear in the parameters, such as the Scheffé polynomials (see Cornell (2002) and Atkinson et al. (2007), Chapter 11). Snee (1975) recommended that the support points of an optimal design for fitting a second-order model over a constrained region is a subset of the extreme vertices, edge centroids, constraint plane centroids and the overall centroid. Snee (1979) extended this methodology to multicomponent constraints. Draper and Pukelsheim (1999) presented optimal designs for estimating the parameters in the homogeneous K-polynomials in mixture variables.

The special difficulty of model (2) is that, depending on the values of the parameters r, s , the model may be nonlinear in the parameters. Therefore, for models nonlinear in the parameters, specifying an optimal design is dependent on initial guesses for the parameters. Atkinson et al. (2007) discussed the derivation of optimal designs for nonlinear models. Atkinson and Bogacka (2002); Atkinson et al. (1998) discussed the derivation of optimal designs for estimating

the order and the rate of reactions from kinetic models. This paper presents methodology of optimal experimental designs, and the derivation thereof, for a number of special forms for the weighted power-mean-mixture model for three component mixtures. With the design methodology presented, experiments in mixture variables can be specified that are optimal according to a design criterion for estimating the parameters in the weighted power-mean-mixture model most precisely.

2 Experimental Design Theory

2.1 Notation and theory

In this section we establish notation and provide an introduction to the theory of optimal experimental design. A nomenclature of the symbols used on the paper is listed in Appendix B. Let y be a random response variable with constant variance σ^2 , then at the setting of mixture variables $\underline{x} = (x_1, x_2, \dots, x_q)^T$, subject to constraints (1), y has expected value:

$$E(y) = \eta(\underline{x}, \underline{\phi}) = \left[\sum_{k=1}^q x_k \left(\sum_{l=1}^q x_l a_{kl}^s \right)^{r/s} \right]^{1/r} \quad (3)$$

where $\underline{\phi}$ is the $p \times 1$ vector of parameters containing all the pure and binary interactions a_{kl} , $k, l = 1, 2, \dots, q$, and the orders r, s (Focke et. al., 2007). We consider a continuous design on the simplex $S_q = \left\{ x_k, k = 1, 2, \dots, q \mid \sum_k x_k = 1, 0 \leq x_k \leq 1 \right\}$ which allocates weights to n distinct sets of mixture variables and which can be specified as:

$$\xi = \left\{ \begin{array}{cccc} \underline{x}_1 & \underline{x}_2 & \cdots & \underline{x}_n \\ \omega_1 & \omega_2 & \cdots & \omega_n \end{array} \right\} \quad (4)$$

where the $\underline{x}_u \in S_q$, $u = 1, 2, \dots, n$, are the values of the q mixture variables, ω_u , $u = 1, 2, \dots, n$, $0 \leq \omega_u \leq 1$, $\sum_{u=1}^n \omega_u = 1$, are the design weights. Note that if N total observations are to be taken, then the number allocated to the set \underline{x}_u is given by $N\omega_u$ which is not necessarily integer. The optimum exact design may be approximated by a design with the number of points at \underline{x}_u the integer closest to $N\omega_u$. Alternatively, and more pragmatically, an exact design can be constructed by applying an exchange algorithm of the type discussed by Atkinson et al. (2007), Chapter 12.

Criteria in optimal design are functions of the information matrix for the set of parameters of interest. The information matrix has as entries the squared and cross-products of the sensitivities of the model to the set of parameters of interest at each experimental design point. For models linear in the parameters, the inverse of this matrix is the variance-covariance matrix of the parameter

estimates. In a model nonlinear in the parameters, this property holds asymptotically. Specifically, the information matrix of a design ξ for the p parameters $\underline{\phi}$ is given by:

$$M(\xi, \underline{\phi}) = F^T W F \quad (5)$$

where

$$F = \begin{pmatrix} \underline{f}^T(\underline{x}_1, \underline{\phi}) \\ \underline{f}^T(\underline{x}_2, \underline{\phi}) \\ \vdots \\ \underline{f}^T(\underline{x}_n, \underline{\phi}) \end{pmatrix}$$

is a $n \times p$ matrix and the u -th row has j -th element

$$f_j(\underline{x}_u, \underline{\phi}) = \frac{\partial \eta(\underline{x}_u, \underline{\phi})}{\partial \phi_j}, j = 1, 2, \dots, p.$$

called the parameter sensitivity, and

$$W = \text{diag}\{\omega_1, \dots, \omega_n\}$$

The information matrix can be written as:

$$M(\xi, \underline{\phi}) = \sum_{u=1}^n \omega_u M(\underline{x}_u, \underline{\phi}) \quad (6)$$

where $M(\underline{x}_u, \underline{\phi}) = \underline{f}(\underline{x}_u, \underline{\phi}) \underline{f}^T(\underline{x}_u, \underline{\phi})$ the information matrix at the u -th experiment. Therefore, for the general form of the model (3), the information matrix depends on the unknown parameters $\underline{\phi}$. In this paper, we adopt a best guess $\underline{\phi}_0$ for the parameters and consider designs which maximize an appropriate function of $M(\xi, \underline{\phi})$ evaluated at $\underline{\phi} = \underline{\phi}_0$. Such designs are termed locally optimum. Alternative strategies are to derive Bayesian designs which require a prior distribution of the parameters (Atkinson and Bogacka, 2002).

2.2 D -optimal designs

In this paper our interest is in specifying the mixture variable settings for the precise estimation of the parameters in the model (3). Therefore, D -optimality is the appropriate criterion. D -optimal designs maximize the determinant of the information matrix, $|M(\xi, \underline{\phi})|$, or equivalently, minimize the asymptotic generalized variance of the parameter estimates. Often, the natural logarithm of the information matrix, $\log |M(\xi, \underline{\phi})|$, is maximized to ensure convexity. Due to the potentially large number of parameters in model (3), obtaining analytical solutions for the maximization of $|M(\xi, \underline{\phi})|$ are cumbersome or mostly impossible. Therefore, we apply a suitable nonlinear constrained optimization algorithm to perform numerical maximization of $|M(\xi, \underline{\phi})|$, evaluated at $\underline{\phi} = \underline{\phi}_0$.

For numerical optimization, the celebrated Equivalence Theorem from Kiefer and Wolfowitz (1960) can be used to evaluate the global optimality or otherwise of a candidate design (Atkinson and Haines, 1996). This theorem relates the maximization of the determinant of the information matrix to the minimization of the maximum variance of prediction over the simplex S_q . Let the standardized variance of prediction at the point \underline{x} be defined as:

$$d(\underline{x}, \xi, \underline{\phi}) = \underline{f}^T(\underline{x}, \underline{\phi}) M^{-1}(\xi, \underline{\phi}) \underline{f}(\underline{x}, \underline{\phi}) \quad (7)$$

The equivalence theorem states that a design, ξ^* , is globally optimal if and only if $d(\underline{x}, \xi^*, \underline{\phi}) \leq p$, the number of parameters in the model, and further that the maximum value is attained at the support points \underline{x}^* of ξ^* . It follows from Caratheodory's Theorem that a continuous D -optimal design is based on, at most, $p(p+1)/2$ design points (Federov (1972), Chapter 2). Often locally D -optimal designs are based on exactly p number of design points, when the weights associated with the support points are equal to $1/p$ (Federov (1972), Chapter 2).

The D -efficiency of a design ξ_1 relative to a design ξ_2 is defined as:

$$Eff_D(\xi_1, \xi_2, \underline{\phi}) = 100 \left(\frac{|M(\xi_1, \underline{\phi})|}{|M(\xi_2, \underline{\phi})|} \right)^{(1/p)} \quad (8)$$

where p is the number of parameters in the model. D -efficiencies are calculated to specify the efficiency of experimental designs for estimating the parameters in the model most precisely. The construction of D -optimal designs have been widely communicated in the design literature for models linear in the parameters in mixture variables. Draper and Pukelsheim (1999) presented D -optimal designs for estimating the parameters in the homogeneous K-polynomials in mixture variables. Cornell (2002) and Atkinson et al. (2007) (Chapter 11) also discussed the derivation of D -optimal designs for linear models in mixture variables. However, the derivation of optimal designs for weighted power-mean-mixture models have not been communicated previously in the design literature.

2.3 D_s -optimal designs

If only a subset of v of the parameters, $\underline{\phi}_1$, is of interest, let the parameters be partitioned as $\underline{\phi} = (\underline{\phi}_1, \underline{\phi}_2)$ with $M_{22}(\xi, \underline{\phi})$ the information matrix for the $p-v$ parameters not of interest, and $\underline{f}_{22}^T(\underline{x}, \underline{\phi})$ the associated vector of sensitivities. Then the D_s -optimal design for $\underline{\phi}_1$ maximizes $|M(\xi, \underline{\phi})|/|M_{22}(\xi, \underline{\phi})|$. The equivalence theorem states that the design ξ^* is D_s -optimal if and only if $d_s(\underline{x}, \xi^*, \underline{\phi}) \leq v$ where

$$d_s(\underline{x}, \xi, \underline{\phi}) = \underline{f}^T(\underline{x}, \underline{\phi}) M^{-1}(\xi, \underline{\phi}) \underline{f}(\underline{x}, \underline{\phi}) - \underline{f}_{22}^T(\underline{x}, \underline{\phi}) M_{22}^{-1}(\xi, \underline{\phi}) \underline{f}_{22}(\underline{x}, \underline{\phi})$$

is the variance of prediction at a point $\underline{x} \in S_q$. The D_s -efficiency of a design ξ_1 relative to a design ξ_2 is defined as:

$$Eff_{D_s}(\xi_1, \xi_2, \underline{\phi}) = 100 \left(\frac{|M(\xi_1, \underline{\phi})| / |M_{22}(\xi_1, \underline{\phi})|}{|M(\xi_2, \underline{\phi})| / |M_{22}(\xi_2, \underline{\phi})|} \right)^{(1/v)} \quad (9)$$

Atkinson and Bogacka (2002) derived D_s -optimal designs for estimating the orders of reaction from a kinetic model.

3 Special Forms and their Optimal Designs

3.1 Continuous D -optimal designs

In this section we use the preceding theory to derive continuous D -optimum designs for some special forms of the power-mean-mixture models (3). We consider the data from Focke et al. (2007) for specifying the initial guesses of the parameter estimates for the derivation of the optimal designs. Furthermore, we will consider designs for $q = 3$ mixture variables only. As a subset of the data in Focke et al. (2007), $n = 68$ data points were used to study the prediction of viscosity as a function of the three components acetone, methanol and water. The data are depicted in Appendix A in the Appendix. Extending the results to more than 3 components is only of mathematical interest and does not contribute to a greater understanding of the importance of applying the methodology of the optimal design of experiments.

For $q = 3$, model (3) becomes

$$\begin{aligned} \eta(\underline{x}, \underline{\phi}) &= \left[\sum_{k=1}^3 x_k \left(\sum_{l=1}^3 x_l a_{kl}^s \right)^{r/s} \right]^{1/r} \\ &= [x_1 (a_{11}^s x_1 + a_{12}^s x_2 + a_{13}^s x_3)^{r/s} + \\ &\quad x_2 (a_{21}^s x_1 + a_{22}^s x_2 + a_{23}^s x_3)^{r/s} + \\ &\quad x_3 (a_{31}^s x_1 + a_{32}^s x_2 + a_{33}^s x_3)^{r/s}]^{1/r} \end{aligned} \quad (10)$$

If all the parameters in model (10) are of interest, with $a_{kl} \neq a_{lk}$, $k, l = 1, 2, 3$, then there are $q = 11$ parameters to be estimated, including r and s . According to the theory, the optimal design will have at least $n = p = 11$ distinct points in the $q = 3$ mixture variables. Therefore, in addition to the $n = 11$ optimal weights required, the D -optimality criterion must be optimized for $p(q+1) = 44$ values with a nonlinear constrained optimization algorithm. This can become very tedious and time consuming. Therefore, to simplify the optimization and interpretation of the results, we will consider special forms of model (10) for given values of r and s .

The special forms considered in this paper are listed in Table 1, together with the number of unknown parameters. Therefore, for deriving D -optimal designs, the parameters a_{kl} , $k, l = 1, 2, 3$ are of interest alone and we assume r and s are fixed and without error. Furthermore, as a result of considering r

Table 1: Special forms of model (10) for given values of r and s .

r, s	Model	Number of parameters	Equation number
$r = s$	$\eta(\underline{x}, \underline{\phi}) = \left(\sum_{k=1}^3 \sum_{l=1}^3 a_{kl}^r x_k x_l \right)^{\frac{1}{r}}$	6	(11)
$r = 1, s = \frac{1}{2}$	$\eta(\underline{x}, \underline{\phi}) = \sum_{k=1}^3 x_k \left(\sum_{l=1}^3 a_{kl}^{(1/2)} x_l \right)^2$	9	(12)
$r = \frac{-5}{6}, s = \frac{1}{2}$	$\eta(\underline{x}, \underline{\phi}) = \left[\sum_{k=1}^3 x_k \left(\sum_{l=1}^3 x_l a_{kl}^{1/2} \right)^{-5/3} \right]^{-6/5}$	9	(14)
$r = 0, s = 0$	$\eta(\underline{x}, \underline{\phi}) = \exp \left(\sum_{k=1}^3 \sum_{l=1}^3 \ln(a_{kl}) x_k x_l \right)$	6	(15)
$r = 0, s = 1$	$\eta(\underline{x}, \underline{\phi}) = \prod_{k=1}^3 \left(\sum_{l=1}^3 a_{kl} x_l \right)^{x_k}$	9	(16)

and s as known, they are not included in the construction of the information matrix (6), i.e. model sensitivities are not calculated with respect to r and s . The power-mean-mixture models listed in Table 1 have been communicated previously in the chemical engineering literature, and are discussed in detail in the following paragraphs for deriving optimal designs.

Model (10) reduces to a quadratic Scheffé K-polynomial in η^r , that is for $r = s$:

$$\eta(\underline{x}, \underline{\phi})^r = \sum_{k=1}^3 \sum_{l=1}^3 a_{kl}^r x_k x_l = \sum_{k=1}^3 a_{kk}^r x_k^2 + \sum_{k=1}^3 \sum_{l>k}^3 (a_{kl}^r + a_{lk}^r) x_k x_l \quad (11)$$

(Draper and Pukelsheim, 1998). The over-parameterization is circumvented by setting $a_{kl} = a_{lk}$. D -optimal designs are well known for Scheffé S-polynomials (Scheffé, 1958) and is discussed by Cornell (2002). Draper and Pukelsheim (1999) utilized the Kiefer ordering for simplex designs and showed that the D -optimal designs for S- and K-polynomials are the same since the models can be interchanged due to the constraint $\sum_k x_k = 1$. However, the K-polynomials are homogeneous of order 1, which is a great advantage over the S-polynomials which are not homogeneous (Draper and Pukelsheim, 1998).

For $r = 1$, the D -optimal design is the simplex-lattice with $n = 6$ design points, i.e. three pure component blends and three 50:50 binary blends, with equal weights $\omega = 1/6$ at each design point. For $r = 2$, model (11) becomes nonlinear in the parameters and the information matrix is a function of the unknown parameters. From Focke et al. (2007), the initial values were specified as: $a_{11} = 0.301, a_{22} = 0.542, a_{33} = 0.892, a_{12} = 0.0774, a_{13} = 0.7685, a_{23} = 1.5716$. The D -optimal design criterion, $|M(\xi, \underline{\phi})|$, was maximized by a suitable nonlinear constrained optimizer. The D -optimal design consists of $n = 6$ design

points, i.e. three pure component blends and the three binary blends (0.6, 0, 0.4), (0.6, 0.4, 0), (0, 0.5, 0.5), with equal weights $w = 1/6$ at each design point. Note that the three components are not equally represented in all three binary experiments. This is due to nonlinearity of the model, and obviously, the result is a function of the initial guesses.

For $r = 3$, the nonlinearity of model (11) increases. For the following initial guesses for $a_{12} = 0.00089$, $a_{13} = 0.6524$, $a_{23} = 1.4067$, with the pure component parameters remaining the same, the optimal design points, in addition to the pure component blends, are (0.722, 0, 0.278), (0.67, 0.33, 0), (0, 0.557, 0.443) for the binary mixtures, with equal weights $w = 1/6$ at each design point. Note the binary mixtures involving x_1 assigns a higher proportion to x_1 . The specific design structure cannot be known without deriving the D -optimal design.

For $r = 1$, $s = 1/2$, model (10) becomes:

$$\eta(\underline{x}, \underline{\phi}) = \sum_{k=1}^3 x_k \left(\sum_{l=1}^3 a_{kl}^{(1/2)} x_l \right)^2 \quad (12)$$

The D -optimal design is depicted in Table 2 for (12), with initial guesses for $a_{12} = 0.7767$, $a_{13} = 0.0001$, $a_{21} = 0.0001$, $a_{23} = 6.0754$, $a_{31} = 2.3898$, $a_{32} = 0.0368$, with the pure component parameters the same as above. The design has $n = 9$ points with equal weight $\omega = 1/9$ assigned to each point.

However, Focke et al. (2007) showed that (12) simplifies to the cubic Scheffé K -polynomial:

$$\begin{aligned} \eta(\underline{x}, \underline{\phi}) = & c_{111}x_1^3 + c_{222}x_2^3 + c_{333}x_3^3 + 3c_{112}x_1^2x_2 + 3c_{122}x_1x_2^2 \\ & + 3c_{113}x_1^2x_3 + 3c_{133}x_1x_3^2 + 3c_{223}x_2^2x_3 + 3c_{233}x_2x_3^2 + 6c_{123}x_1x_2x_3 \end{aligned} \quad (13)$$

where $c_{klm} = (\sqrt{a_{kl}a_{km}} + \sqrt{a_{lk}a_{lm}} + \sqrt{a_{mk}a_{ml}}) / 3$, $k, l, m = 1, 2, 3$. The c_{klm} parameters in (13) as a function of the binary interaction parameters a_{kl} in (12) is a very important result. It illustrates that ternary and higher order Scheffé K -polynomials can be estimated from binary data. Therefore, the D -optimal design in Table 2 can be used for estimating the cubic model (13).

Alternatively, the D -optimal design for estimating the parameters in the cubic model (13) is the simplex-centroid design with $n = 10$, which consists of, in addition to the three pure component blends, binary blends (0, 0.276, 0.724), (0.276, 0, 0.724), (0.276, 0.724, 0), (0.724, 0, 0.276), (0.724, 0.276, 0), (0, 0.724, 0.276), and the ternary blend (0.333, 0.333, 0.333). Equal weight $\omega = 1/10$ is assigned to each design point. See Cornell (2002) (Chapter 2) for a discussion of simplex-centroid designs. Figure 2 depicts the design points for the D -optimal design in Table 2 for model (12), together with the D -optimal design (simplex-centroid) for the cubic Scheffé K -polynomial (13). Clearly the two designs are very similar on the simplex S_q . This is because model (12) simplifies to the cubic K -polynomial (13) with one additional point in the optimal design as depicted in Figure 2.

The D -efficiency of the simplex-centroid design compared to the optimal design in Table 2 is 96% for estimating the parameters in model (12). Therefore,

Table 2: D -optimal design for model (12) with $r = 1, s = 1/2$

	x_1	x_2	x_3	w
1	0.0000	0.4008	0.5992	0.111
2	1.0000	0.0000	0.0000	0.111
3	0.0000	0.0000	1.0000	0.111
4	0.3378	0.3177	0.3444	0.111
5	0.2761	0.7239	0.0000	0.111
6	0.0000	1.0000	0.0000	0.111
7	0.2764	0.0000	0.7236	0.111
8	0.7236	0.2764	0.0000	0.111
9	0.7235	0.0000	0.2765	0.111

the simplex-centroid design is very efficient in estimating the parameters in model (12), and optimal for estimating the parameters in the cubic Scheffé K-polynomial (13). An additional advantage of the simplex-centroid design is that it is not dependent on initial guesses for the parameter estimates.

From a practical perspective, we calculate the D -efficiency of the $n = 68$ data points used in Focke et al. (2007) for estimating the power-mean-mixture models. The D -efficiency of the data in Focke et al. (2007) is equal to 71% for estimating model (12), and equal to 76% for estimating the cubic Scheffé polynomial (13). This indicates that almost 30% more replications are required for the data in Focke et al. (2007) to be as efficient in estimating the above two models. Therefore, although many more mixture data have been collected and fitted, the data is not as efficient as those of the optimal designs. This has great practical impact in terms of the design and analysis of mixture experiments in the chemical and chemical engineering disciplines. It illustrates that the optimal mixture experiments are more important than collecting huge amounts of data.

Focke et al. (2007) showed that $r = -5/6$ and $s = 1/2$ yield the optimal mixing rule for predicting liquid viscosity. Model (10) becomes:

$$\eta(\underline{x}, \underline{\phi}) = \left[\sum_{k=1}^3 x_k \left(\sum_{l=1}^3 x_l a_{kl}^{1/2} \right)^{-5/3} \right]^{-6/5} \quad (14)$$

The D -optimal design for model (14), with initial parameter guesses $a_{12} = 0.66804, a_{13} = 0.7222, a_{21} = 0.84593, a_{23} = 1.2223, a_{31} = 3.88214, a_{32} = 2.6656$, and the pure component parameters the same as above, is depicted in Table 3. The design has $n = 10$ design points with unequal weights assigned to the design points. Note that the ternary mixture, i.e. (0.3632, 0.2931, 0.3436), which is close to the midpoint, has an optimal weight which is almost half of the other design points. The optimal weights are very similar for the other data points. For practical application, this result indicates that the pure and binary mixtures are assigned twice as many replications than the ternary mixture. The D -optimal design is plotted in Figure 2, together with the $n = 68$ data points from Focke et al. (2007). The figure shows that the historical data from Focke

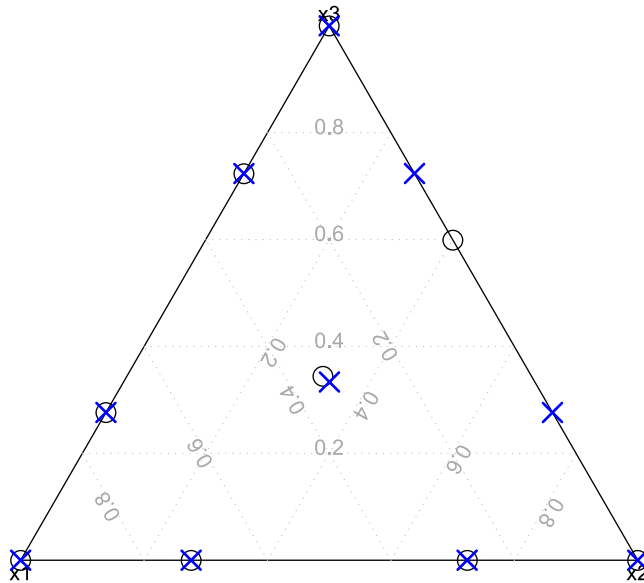


Figure 1: D -optimal design points for $r = 1, s = 1/2$, model (12) (\circ), and for the cubic Scheffé polynomial (\times).

et al. (2007) are populated with mixtures which are scattered over the simplex S_q , although there are fewer mixtures with high proportions of component x_1 present.

The D -efficiency of the $n = 68$ historical data compared to the optimal design with $n = 10$ distinct points is equal to 74%. Therefore, although a significant amount of data have been used to derive the optimal mixing rule for viscosity, it is only 74% as efficient as the optimal mixture design with $n = 10$ distinct points for estimating the proposed model. However, for practical application at least a total of $N = 20$ mixtures would be needed, with one measurement taken at the ternary mixture and two replications taken at all the other mixtures. The D -efficiency of the optimal design for the cubic polynomial (13), i.e. simplex-centroid design, compared to the optimal design in Table 3 for estimating model (14) is 87%. Since the simplex-centroid design is independent of the parameters' estimates, with good efficiency, it may be used in sequential mixture design studies for obtaining initial data for parameter estimation.

D -optimal designs can also be derived for other special forms of the power-mean-mixture model. For $r = 0, s = 0$, we obtain the mixing law for viscosity by Grunberg and Nissan (1949):

$$\eta(\underline{x}, \underline{\phi}) = \exp \left(\sum_{k=1}^3 \sum_{l=1}^3 \ln(a_{kl}) x_k x_l \right) \quad (15)$$

with $a_{kl} = a_{lk}, k, l = 1, 2, 3$. Model (15) has six parameters to estimate. Let

Table 3: D -optimal design for model (14) with $r = -5/6$ and $s = 1/2$.

	x_1	x_2	x_3	w
1	0.0000	0.2516	0.7484	0.1111
2	1.0000	0.0000	0.0000	0.1012
3	0.0000	0.0000	1.0000	0.1111
4	0.6638	0.3362	0.0000	0.0875
5	0.0000	0.5975	0.4025	0.1085
6	0.0000	1.0000	0.0000	0.1093
7	0.1891	0.0000	0.8109	0.1111
8	0.2620	0.7380	0.0000	0.1100
9	0.3632	0.2931	0.3436	0.0462
10	0.5036	0.0000	0.4964	0.1039

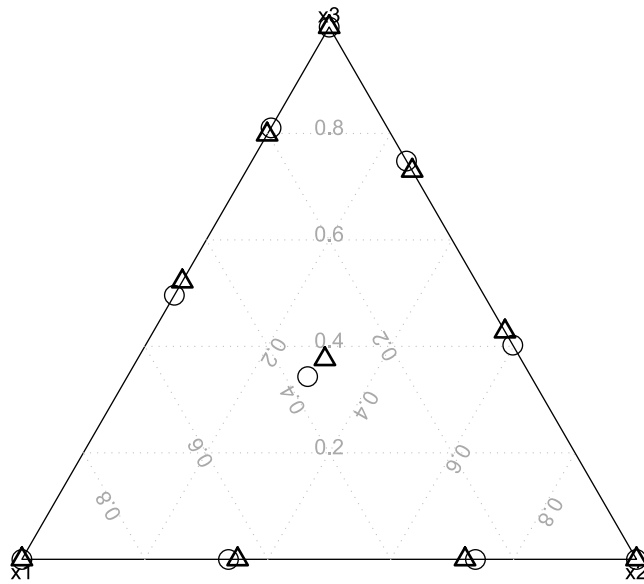


Figure 2: D -optimal design points for $r = -5/6, s = 1/2$, model (14) (○), and for the $n = 68$ data points from Focke et al. (2007) (△).

Table 4: D -optimal design for model (15) with $r = 0, s = 0$.

	x_1	x_2	x_3	w
1	0.3055	0.4516	0.2430	0.1667
2	0.0000	0.4498	0.5502	0.1667
3	0.0000	0.0000	1.0000	0.1667
4	0.0000	0.9463	0.0537	0.1667
5	0.3996	0.0000	0.6004	0.1667
6	0.9524	0.0000	0.0476	0.1667

$a_{11} = 0.301, a_{12} = 0.66804, a_{13} = 0.7222, a_{22} = 0.542, a_{23} = 1.2223, a_{33} = 0.892$, the D -Optimal design is depicted in Table 4. The design has six optimal distinct points with equal weight $\omega = 1/6$ assigned to each point. Note that the optimal design includes only one pure component blend for x_3 , i.e. $(0, 0, 1)$. The simplex-lattice design in three components has also 6 design points (Cornell, 2002). The D -efficiency of the simplex-lattice design compared to the D -optimal design in Table 4 for estimating model (15) is 88%. Therefore, due to the high efficiency, and the fact that the simplex-lattice is independent of the parameter estimates, it may be used to obtain initial parameter estimates.

For $r = 0, s = 1$, we obtain the model from Wilson (1964):

$$\eta(\underline{x}, \underline{\phi}) = \prod_{k=1}^3 \left(\sum_{l=1}^3 a_{kl} x_l \right)^{x_k} \quad (16)$$

Model (16) has nine parameters to estimate. Let $a_{11} = 0.204, a_{12} = 0.0197, a_{13} = 0.0181, a_{21} = 1.295, a_{22} = 0.5482, a_{23} = 0.3149, a_{31} = 9.5214, a_{32} = 7.1334, a_{33} = 0.9309$, the D -Optimal design is depicted in Table 5. The design has nine optimal distinct points with equal weight $\omega = 1/9$ assigned to each point. The designs in Tables 4 and 5 both assign equal weight to each design point confirming the theory discussed in Section 2.1, i.e. $\omega = 1/p$ for p parameters. However, the designs are non-symmetrical due to the non-linearity of the models. In comparison, the simplex-centroid design in three components, which is optimal for the cubic Scheffé K-polynomial, has 10 distinct design points. The D -efficiency of the simplex-centroid design compared to the D -optimal design in Table 5 for estimating model (16) is only 47%. Therefore, care should be taken in using the general simplex-centroid design for estimating model (16), even for obtaining initial parameter estimates. This result is not known without applying the D -optimality criterion.

3.2 Continuous D_s -optimal designs

Model (14), with $r = -5/6$ and $s = 1/2$, was illustrated by Focke et al. (2007) to be the optimal mixing rule for predicting liquid viscosity. Therefore, for illustrating D_s -optimal designs, we consider model (14) and specifically we assume that the parameters $a_{kk}, k = 1, 2, 3$ are not of interest for estimation. This assumption is sensible in analytical chemistry because the behavior of multicomponent

Table 5: D -optimal design for model (16) with $r = 0, s = 1$.

	x_1	x_2	x_3	w
1	0.0000	0.0000	1.0000	0.1111
2	0.0806	0.2353	0.6841	0.1111
3	0.1688	0.0000	0.8312	0.1111
4	0.0832	0.8390	0.0778	0.1111
5	0.0000	0.1443	0.8557	0.1111
6	0.3632	0.4446	0.1923	0.1111
7	0.0000	1.0000	0.0000	0.1111
8	0.0000	0.4732	0.5268	0.1111
9	0.6201	0.0000	0.3799	0.1111

mixtures is naturally affected by the interactions of unlike molecules (Hamad, 1998; Prausnitz et al., 1999; Walas, 1985). Focke et al. (2007) also assumed the pure component properties to be known and only estimated the binary interaction parameters. However, although the pure component properties can be estimated or specified very precisely they remain subject to uncertainty. The objective is to derive the continuous D_s -optimal design for estimating the parameters $a_{kl}, k \neq l = 1, 2, 3$ in model (14), with $a_{kk}, k = 1, 2, 3$ subject to uncertainty.

The D_s -optimal design is depicted in Table 6, with $n = 10$ distinct design points, and unequal weights assigned to each point. The D_s -optimal design is plotted in Figure 3 together with the D -optimal design for estimating model (14). Clearly, the two optimal design are very similar. Specifically, although the optimal weights are very similar for the different design points that lie close to each other, the D_s -optimal design has a smaller weight assigned to the approximate center point compared to the D -optimal design. The D_s -efficiency for the D -optimal design in Table 3 compared to the D_s -optimal design in Table 6 for estimating model (14), with $a_{kk}, k = 1, 2, 3$ subject to uncertainty, is 96%. Therefore, the very high efficiency confirms the similarity of the two designs. The D_s -efficiency for the historical $n = 68$ data points from Focke et al. (2007) (Appendix A) compared to the D_s -optimal design in Table 6 for estimating model (14), with $a_{kk}, k = 1, 2, 3$ subject to uncertainty, is only 71%. Again, this illustrates the practical value of determining the optimal mixture experiments rather than engaging in ad-hoc experimentation.

3.3 Exact designs

For a design ξ_n , with n_u observations at each x_u , and $N = \sum_u n_u, u = 1, 2, \dots, t$, the information matrix is:

$$M(\xi_n, \underline{\phi}) = F^T W_n F$$

where $W_n = \text{diag}\{1/n_1, 1/n_2, \dots, 1/n_t\}$.

Table 6: D_s -optimal design for model (14) with $r = -5/6$ and $s = 1/2$.

	x_1	x_2	x_3	w
1	0.649	0.351	0.000	0.093
2	0.279	0.721	0.000	0.129
3	0.000	0.572	0.428	0.110
4	0.000	1.000	0.000	0.080
5	0.000	0.270	0.730	0.124
6	0.000	0.000	1.000	0.080
7	0.478	0.000	0.522	0.100
8	0.202	0.000	0.798	0.126
9	1.000	0.000	0.000	0.076
10	0.319	0.305	0.376	0.082

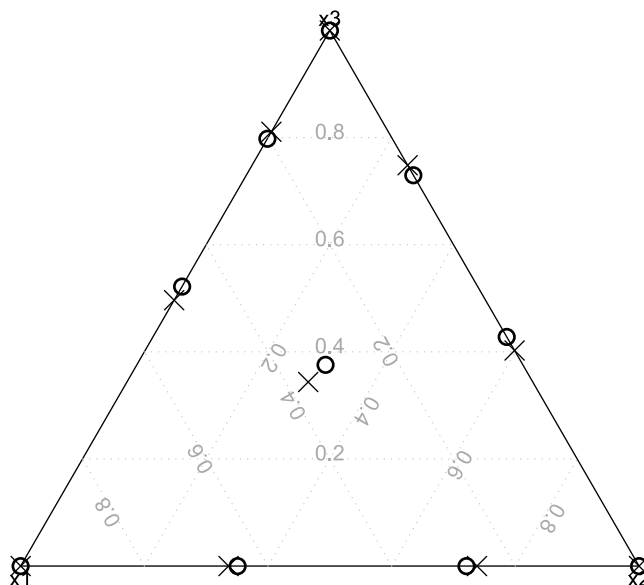


Figure 3: D_s -optimal design points (\circ) and the D -optimal design points (\times) for $r = -5/6, s = 1/2$, model (14).

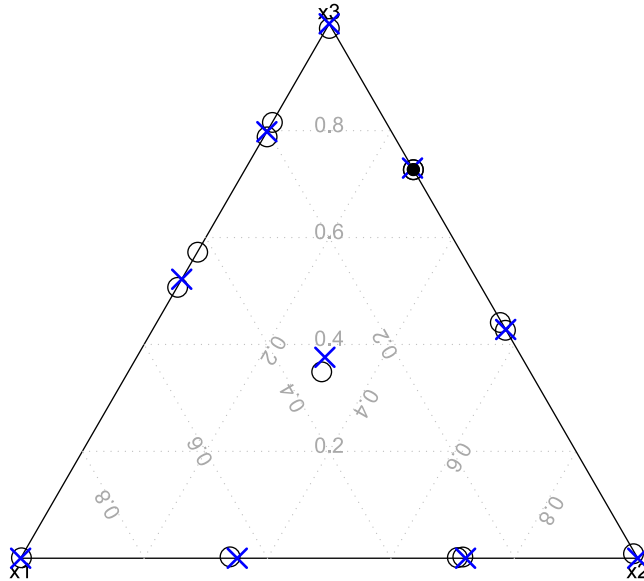


Figure 4: D_s -optimal exact design points (o) and the D_s -optimal continuous design points (x) for $r = -5/6, s = 1/2$, model (14).

Cornell (2002), Chapter 4, and Atkinson et al. (2007), Chapter 16, discuss the construction of exact designs for mixture experiments. In this section we illustrate the exact design for the D_s -optimality criterion, i.e. maximize $|M(\xi_n, \phi)| / |M_{22}(\xi_n, \phi)|$, for estimating the parameters $a_{kl}, k \neq l = 1, 2, 3$ in model (14) under the assumption that $a_{kk}, k = 1, 2, 3$ are subject to uncertainty. For constructing exact designs, an exchange algorithm is deployed which require a candidate set of mixture design points. Therefore, a grid of 10 000 random mixture design points were created for the candidate set. We applied the Federov coordinate-exchange algorithm (Miller and Nguyen, 1994) to select 15 design points from the 10 000 candidate set. 15 Points were specified to evaluate whether some points are replicated because the optimal continuous design yielded 10 points. The exchange algorithm was iterated 3 000 times, and the computer time was 250 minutes in total. There are other exchange algorithms available for generating exact optimal designs, such as the DETMAX algorithm (Galil and Kiefer, 1980; Mitchell, 1974a,b; Welch, 1984).

The D_s -optimal exact design points are plotted in Figure 4. The solid dot indicates the replicated replicated point. Therefore, there are 14 distinct design points with two replications. Also plotted in Figure 4 is the D_s -optimal continuous design in Table 6. Clearly, the two optimal designs are very similar. The D_s -efficiency of the exact design compared to the optimal design in Table 6 is equal to 95%. This example illustrates that exact optimal designs can be generated and may be as efficient as the continuous optimal designs.

4 Discussion and Conclusions

In this paper we presented optimal designs for estimating the parameters in the generalized weighted-power-mean mixture models, which may be nonlinear in the pure and binary interaction parameters. We illustrated the construction of D - and D_s -optimal designs and discussed the designs both from a theoretical and practical perspective. Specifically, design efficiencies were calculated for all the optimal designs. We highlighted the difference between continuous and exact optimal designs, and the construction thereof.

In this paper we derived locally optimal designs, i.e. those designs which depend on an initial guess for the parameters. Therefore, if the values of one or more of the parameters are changed, the optimal design might change due to a change in the parameter sensitivities and consequently the information matrix (6). Uncertainties in the parameter values can be addressed by specifying a prior distribution on the parameters and deriving Bayesian D - and D_s -optimal designs (Atkinson and Bogacka, 2002). Alternatively, mixture experiments may be performed in order to obtain initial parameter estimates.

Specifically, it was illustrated that the common mixture designs, such as the simplex-lattice and simplex-centroid designs, have high efficiency for estimating most of the special forms of the weighted power-mean-mixture models that are nonlinear in the parameters. Therefore, if desired, these designs may be employed for determining initial parameter estimates for employing a sequential design strategy. Atkinson et al. (2007), Chapter 17, illustrates the efficiency of sequential designs for nonlinear models, whereby a preliminary (arbitrary) experimental design is employed for obtaining initial parameter estimates, and then one experiment at a time is added sequentially according to the optimality criterion of interest.

To conclude, it is recommended that the methodology of optimal experimental design should be employed for specifying experiments in mixture variables that are optimal for estimating the parameters in weighted power-mean-mixture models most precisely.

5 Acknowledgments

The authors would like to thank the management of Sasol Technology R&D for their permission to publish this work. Financial support for W.W. Focke, from the Institutional Research Development Programme (IRDP) of the National Research Foundation of South Africa, as well as Xyris Technology CC, is gratefully acknowledged.

Appendix A Mixture data

Exp.	Acetone	Methanol	Water	Viscosity	Exp.	Acetone	Methanol	Water	Viscosity
1	0.146	0.211	0.643	1.250	35	0.702	0.298	0.000	0.329
2	0.494	0.073	0.433	0.635	36	0.692	0.308	0.000	0.331
3	0.280	0.076	0.644	1.049	37	0.645	0.355	0.000	0.340
4	0.243	0.090	0.666	1.129	38	0.537	0.463	0.000	0.347
5	0.073	0.128	0.798	1.448	39	0.359	0.641	0.000	0.377
6	0.250	0.155	0.595	1.037	40	0.297	0.703	0.000	0.394
7	0.433	0.158	0.409	0.658	41	0.210	0.790	0.000	0.423
8	0.244	0.164	0.591	1.037	42	0.122	0.878	0.000	0.465
9	0.098	0.169	0.734	1.409	43	0.000	1.000	0.000	0.542
10	0.300	0.171	0.529	0.899	44	0.000	0.000	1.000	0.892
11	0.270	0.192	0.539	0.948	45	0.000	0.051	0.949	1.107
12	0.354	0.297	0.349	0.663	46	0.000	0.113	0.888	1.339
13	0.057	0.283	0.660	1.422	47	0.000	0.141	0.859	1.421
14	0.416	0.332	0.252	0.543	48	0.000	0.228	0.772	1.559
15	0.288	0.339	0.374	0.745	49	0.000	0.293	0.707	1.558
16	0.184	0.359	0.458	0.957	50	0.000	0.420	0.580	1.431
17	0.450	0.353	0.197	0.488	51	0.000	0.486	0.514	1.331
18	0.094	0.412	0.494	1.139	52	0.000	0.554	0.446	1.214
19	0.307	0.484	0.209	0.570	53	0.000	0.713	0.287	0.963
20	0.244	0.448	0.308	0.713	54	0.000	0.804	0.196	0.814
21	0.326	0.497	0.178	0.532	55	0.000	0.835	0.166	0.777
22	0.106	0.556	0.338	0.893	56	0.000	0.914	0.086	0.660
23	0.106	0.593	0.301	0.841	57	0.000	1.000	0.000	0.542
24	0.220	0.648	0.132	0.541	58	0.000	0.000	1.000	0.892
25	0.183	0.674	0.143	0.573	59	0.058	0.000	0.942	1.207
26	0.271	0.098	0.631	1.048	60	0.146	0.000	0.854	1.372
27	0.345	0.172	0.483	0.799	61	0.206	0.000	0.794	1.289
28	0.170	0.162	0.668	1.234	62	0.302	0.000	0.698	1.087
29	1.000	0.000	0.000	0.301	63	0.460	0.000	0.540	0.749
30	0.878	0.122	0.000	0.326	64	0.527	0.000	0.473	0.648
31	0.790	0.210	0.000	0.325	65	0.618	0.000	0.382	0.510
32	0.771	0.229	0.000	0.329	66	0.770	0.000	0.230	0.396
33	0.722	0.278	0.000	0.330	67	0.820	0.000	0.181	0.381
34	0.706	0.294	0.000	0.332	68	1.000	0.000	0.000	0.301

Appendix B Nomenclature

q	number of components
p	number of parameters
x	mixture components
$\eta(\underline{x}, \underline{\phi})$	model
$\underline{\phi}$	vector of parameters
a_{kl}	parameters
r, s	parameters
ξ	experimental design
ω	design weights
n	number of observations
$M(\xi, \underline{\phi})$	information matrix
$f^T(\underline{x}, \underline{\phi})$	vector of sensitivities
F	matrix of parameter sensitivities
W	matrix of design weights
$d(\underline{x}, \xi, \underline{\phi})$	standardized variance of prediction

References

- Atkinson, A. and Bogacka, B. (2002). Compound and other optimum designs for systems of nonlinear differential equations arising in chemical kinetics. *Chemometrics and intelligent laboratory systems*, 61:17–33.
- Atkinson, A., Bogacka, B., and Bogacki, M. (1998). D- and t-optimum designs for the kinetics of a reversible chemical reaction. *Chemometrics and Intelligent Laboratory Systems*, 43:185–198.
- Atkinson, A., Donev, A., and Tobias, R. (2007). *Optimum experimental designs, with SAS*. Oxford University Press Inc., New York.
- Atkinson, A. and Haines, L. (1996). *Handbook of Statistics*, chapter Designs for Nonlinear and Generalised Linear Models. In: Ghosh, S., Rao, C.R. (Eds.), *Handbook of Statistics*, Elsevier Science, 13, 437-475., pages 437–475. Elsevier Science.
- Cornell, J. (2002). *Experiments with mixtures: Designs, Models and the Analysis of Mixture Data*. Wiley, New York, third edition.
- Draper, N. and Pukelsheim, F. (1998). Mixture models based on homogeneous polynomials. *Journal of Statistical Planning 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.
- Federov, V. (1972). *Theory of optimal experiments*. Academic Press N.Y.
- Focke, W., Sandrock, C., and Kok, S. (2007). Weighted-power-mean mixture model: Empirical mixing laws for liquid viscosity. *Industrial Engineering Chemistry Research*, 46:4660–4666.
- Galil, Z. and Kiefer, J. (1980). Time- and space-saving computer methods, related to mitchell’s detmax, for finding d-optimum designs. *Technometrics*, 22, NO. 3:301–313.
- Grunberg, L. and Nissan, A. (1949). Mixing law for viscosity. *Nature*, 164:799.
- Hamad, E. (1998). Exact limits of mixture properties and excess thermodynamic functions. *Fluid phase equilibria*, 142:163.
- Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, 12:363–366.
- Miller, A. J. and Nguyen, N.-K. (1994). Algorithm AS 295: A federov exchange algorithm for d-optimal design. *Applied Statistics*, 43, NO 4:669–677.
- Mitchell, T. (1974a). An algorithm for the construction of “d-optimal” experimental designs. *Technometrics*, 16:203–210.
- Mitchell, T. (1974b). Computer construction of “d-optimal” first-order designs. *Technometrics*, 16, NO. 2:211–220.
- Prausnitz, J., Lichtenthaler, R., and de Azevedo, E. (1999). *Molecular Thermodynamics of Fluid-Phase Equilibria*. Prentice Hall: Upper Saddle River, NJ.
- Scheffé, H. (1958). Experiments with mixtures. *Journal of the Royal Statistical Society Series B*, 20:344–360.
- Snee, R. (1975). Experimental designs for quadratic models in constrained mixture spaces. *Technometrics*, 17:399–408.
- Snee, R. (1979). Experimental designs for mixture systems with multicomponent constraints. *Communications in Statistics - Theory and Methods*, A8(4):303–326.
- Walas, S. (1985). *Phase Equilibrium in Chemical Engineering*. Butterworth: Boston.
- Welch, W. J. (1984). Computer-aided design of experiments for response estimation. *Technometrics*, 26, NO. 3:217–224.
- Wilson, G. (1964). Vapor-liquid equilibrium. xi. a new expression for the excess free energy of mixing. *Journal of the American Chemistry Association.*, 86:127.