Design and Analysis of Mixture-of-Mixture Experiments

Lulu Kang[1], V. Roshan Joseph[1] and William A. Brenneman[2]

[1]H. Milton Stewart School of Industrial and Systems Engineering

Georgia Institute of Technology, Atlanta, GA 30332

lkang@isye.gatech.edu, roshan@isye.gatech.edu

[2]Mason Business Center, The Procter & Gamble Company

8700 Mason-Montgomery Road, Mason, OH 45040

brenneman.wa@pg.com

Abstract

In mixture-of-mixture (MoM) experiments, major components are defined as the components which themselves are mixtures of some other components, called minor components. In other MoM experiments, components are divided into different categories. Each category is called major component, and the components within it are minor components. The special structure of the MoM experiment makes the design and modeling approaches different from a typical mixture experiment. In this paper, we propose a new modeling approach called the major-minor model. It overcomes some of the limitations of the multiple-Scheffé model, which has been studied and used for MoM experiments. We also introduce a major-minor D-optimal design strategy for designing the experiment. Some examples are given to illustrate the methodology and we apply the proposed design and modeling approach to a MoM experiment conducted to formulate a new Pringles potato crisp.

KEY WORDS: Mixture Experiment; Categorized-Components; Mixture-of-Mixture; Scheffé model; D-optimal

1. INTRODUCTION

In some mixture experiments the mixture components themselves are made up of other sub-components, or more generally, the mixture components can be divided up into different categories or groups. These types of mixture experiments have been called *mixture-of-mixture* (MoM) experiments or *categorized-components* mixture experiments. To inherit the terms from previous works, each mixture component or category is called a *major component* and the mixture components contained in the major components or categories are called *minor components*.

Assume there are M major components in a mixture experiment, and let c_i be the proportion of the ith major component. They meet the constraint:

$$\sum_{i=1}^{M} c_i = 1, \quad 0 \leq c_i \leq 1, \quad i = 1, 2, \ldots, M. \tag{1}$$

Each major component is composed of m_i minor components, whose proportions with respect to c_i are x_{ij}, such that,

$$\sum_{j=1}^{m_i} x_{ij} = 1, \quad \text{and} \quad 0 \leq x_{ij} \leq 1 \tag{2}$$
$$i = 1, 2, \ldots, M; \quad j = 1, 2, \ldots, m_i.$$

The constraints (1) and (2) result from the nature of the MoM experiments. There could be additional bounds and linear constraints (3) and (4) on the major and minor components.

$$L_i \leq c_i \leq U_i, \quad i = 1, 2, \ldots, M, \tag{3}$$
$$C_k \leq \textstyle\sum_{i=1}^{N} a_{i,k} c_i \leq D_k, \quad \text{for the } k\text{th constraint.}$$

$$l_{ij} \leq x_{ij} \leq u_{ij}, \quad i = 1, 2, \ldots, M; j = 1, 2, \ldots, m_i, \tag{4}$$
$$c_{ik} \leq \textstyle\sum_{j=1}^{m_i} b_{j,ik} x_{ij} \leq d_{ik}, \quad \text{for the } ik\text{th constraint.}$$

Our work is motivated by a practical mixture-of-mixture experiment. The goal of the experiment is to formulate a new Pringles potato crisp that will be packaged in a bag. In

Table 1: Boundaries for major components.

Name	Low	Center	High
A (c_1)	0.601	0.622	0.643
B (c_2)	0.34	0.36	0.38
C (c_3)	0.017	0.018	0.019
Total		1	

Table 2: Boundaries for minor components for A.

Name	Low	Center	High
A_1 (x_{11})	0.835	0.87	0.905
A_2 (x_{12})	0.095	0.13	0.165
Total		1	

this experiment, there are three major components: A, B, and C, whose lower and upper bounds are shown in Table 1. The major component A consists of two minor components: A_1 and A_2; the major component B also consists of two minors: B_1 and B_2; while C can be viewed as a pure material. Table 2 and 3 show the lower and upper bounds for the four minor components. All of the constraints can be written as follows:

$$\begin{aligned} c_1 + c_2 + c_3 &= 1, & 0.601 &\le c_1 \le 0.643, \\ 0.34 \le c_2 &\le 0.38, & 0.017 &\le c_2 \le 0.019, \\ x_{11} + x_{12} &= 1, & 0.835 &\le x_{11} \le 0.905, \\ x_{21} + x_{22} &= 1, & 0.9 &\le x_{21} \le 0.98. \end{aligned} \tag{5}$$

In general, MoM experiments can be transformed into the typical mixture experiments by constructing the appropriate linear constraints. To see how this is done, let $X_{ij} = x_{ij}c_i$, for $j = 1, \dots, m_i$, $i = 1, \dots, M$. The constraint (1) and (3) on the major components become constraints on X_{ij}, where c_i is replaced by $\sum_{j=1}^{m_i} X_{ij}$ for $i = 1, \dots, M$ in (1) and (3). In addition, the constraints (4) on the minor components can be converted into the

Table 3: Boundaries for minor components for B.

Name	Low	Center	High
B_1 (x_{21})	0.9	0.94	0.98
B_2 (x_{22})	0.02	0.06	0.1
Total		1	

linear constraints on $X_{i,j}$ as follows:

$$
\begin{aligned}
& l_{ij}\sum_{j=1}^{m_i} X_{ij} \le X_{ij} \le u_{ij}\sum_{j=1}^{m_i} X_{ij} \quad i=1,2,\ldots,M; j=1,2,\ldots,m_i,\\
& c_{ik}\sum_{j=1}^{m_i} X_{ij} \le \sum_{j=1}^{m_i} b_{j,ik}X_{ij} \le d_{ik}\sum_{j=1}^{n_i} X_{ij} \quad \text{for the } ik\text{th constraint.}
\end{aligned}
$$

For example, in the Pringles formulation experiment, we essentially have five components which have to meet the following constraints:

$$
\begin{aligned}
X_{11}+X_{12}+X_{21}+X_{22}+c_3=1, &\qquad 0.017 \le c_3 \le 0.019 \qquad (6)\\
0.601\times 0.835 \le X_{11} \le 0.643\times 0.905, &\qquad 0.34\times 0.9 \le X_{21} \le 0.38\times 0.98,\\
0.601\times 0.095 \le X_{12} \le 0.643\times 0.165, &\qquad 0.34\times 0.02 \le X_{22} \le 0.38\times 0.1,\\
0.601 \le X_{11}+X_{12} \le 0.643, &\qquad 0.34 \le X_{21}+X_{22} \le 0.38,\\
0.905X_{12}-0.095X_{11} \ge 0, &\qquad 0.98X_{22}-0.02X_{21} \ge 0,\\
0.835X_{12}-0.165X_{11} \le 0, &\qquad 0.9X_{22}-0.1X_{21} \le 0.
\end{aligned}
$$

This naturally raises the question, “why do we have to treat the MoM experiments differently from the usual mixture experiments?” To answer this we have to think through the goals and objectives of MoM experiments. First, the information about how the major components interact with each other and how they interact with their minor components are usually of interest to practitioners. Hence we want to model the mixture-of-mixture structure so as to study the blending properties between and within major components. Moreover, transforming a MoM experiment into a typical mixture experiment will result in an increase in the number of components and complicated constraints. Again, consider the Pringles experiment. As a MoM experiment, apart from the basic "sum-to-1" constraints, it only has upper and lower bounds on each major/minor component as shown in (5). However, after the transformation it has six linear constraints as well as bounds on the components in (6). The design space becomes an irregularly shaped convex polyhedron in high dimension, which makes it harder to construct an efficient design. Moreover, the modeling aspect also faces challenges: the well known collinearity or ill-conditioning problem are likely to appear because the data are collected from a highly constrained region (see Cornell and Gorman, 2003 and Prescott et. al., 2002), introducing stability problems with parameter estimation.

The early works on MoM experiments have assumed that the proportions of the major components are fixed. Lambrakis (1968, 1969) first introduced the multiple-Scheffé model and multiple-lattice design strategy. Other methods were also used such as in Cornell and Good (1970) and Cornell (1971). Cornell and Ramsey (1998) extended the multiple-Scheffé model to the case where both the major and minor components can be varied. Since then the multiple-Scheffé model became very popular. See Piepel (1999), Dingstad et. al. (2003), Borges et. al. (2007), and Didier et. al. (2007) for several case studies. Despite of its popularity, the Multiple-Scheffé model has some limitations which will be discussed in the next section.

This paper is organized as follows. In Section 2, we first review the development of the multiple-Scheffé model, and then point out some problems in its extension to deal with more general MoM experiments. In the next section, we propose a new modeling approach, which we call *major-minor model*, and compare it with the multiple-Scheffé model. In Section 4, we propose a new design called *major-minor D-optimal design* for MoM experiments. In Section 5, we apply our proposed design and modeling methods to the Pringles MoM experiment. In Section 6, we compare the prediction performance of major-minor model and multiple-Scheffé model using simulations and the paper concludes with a summary in Section 7.

2. MULTIPLE-SCHEFFÉ MODEL

The multiple-Scheffé model was first introduced to study the MoM experiment in which the proportions of the major components are fixed. Essentially, the multiple-Scheffé model is a *product* model. Let $f_i(x_{i1}, \ldots, x_{im_i})$ be a Scheffé type of model for the minors of the ith major component. For those major components that only have single component ($m_i = 1$), we take $f_i \equiv 1$. When all c_i's are fixed, the multiple-Scheffé model is a product of $f_i(x_{i1}, \ldots, x_{im_i})$:

$$f(\boldsymbol{x}, \boldsymbol{\gamma}) = \prod_{i=1}^{M} f_i(x_{i1}, \ldots, x_{im_i}). \tag{7}$$

Note that the orders of f_i can be different from each other. The multiple-Scheffé model is still a linear model and $\boldsymbol{\gamma}$ are the coefficient parameters of the polynomial after the product is expanded. Consider the Pringles experiment, in which $M = 3$, $m_1 = m_2 = 2$ and $m_3 = 1$. For the minor components, we use the linear Scheffé model:

$$f_i(x_{i1}, x_{i2}) = \beta_{i1}x_{i1} + \beta_{i2}x_{i2}, \quad \text{for} \quad i = 1, 2.$$

If the major components A, B, and C have fixed proportion in the experiment, the multiple-Scheffé model is given by:

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{\gamma}) &= (\beta_{11}x_{11} + \beta_{12}x_{12})(\beta_{21}x_{21} + \beta_{22}x_{22}) \\ &= \gamma_1 x_{11}x_{21} + \gamma_2 x_{11}x_{22} + \gamma_3 x_{12}x_{21} + \gamma_4 x_{12}x_{22}. \end{aligned}$$

There are many MoM experiments such as the Pringles experiment where the proportions of major components are also variables. Cornell and Ramsey (1998) extended the multiple-Scheffé model to this more general setting. The extended multiple-Scheffé model is given by

$$G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = h(c_1, \ldots, c_M) \times \prod_{i=1}^{M} f_i(x_{i1}, \ldots, x_{im_i}). \tag{8}$$

For instance, if we use a linear Scheffé model $h(c_1, c_2, c_3) = \alpha_1 c_1 + \alpha_2 c_2 + \alpha_3 c_3$ for the major components in the Pringles example, then the multiple-Scheffé model in (8) can be written as:

$$\begin{aligned} G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) &= h(c_1, c_2, c_3) \times f_1(x_{11}, x_{12}) \times f_2(x_{21}, x_{22}) \qquad (9) \\ &= (\gamma_1 x_{11}x_{21} + \gamma_2 x_{11}x_{22} + \gamma_3 x_{12}x_{21} + \gamma_4 x_{12}x_{22})c_1 \\ &+ (\gamma_5 x_{11}x_{21} + \gamma_6 x_{11}x_{22} + \gamma_7 x_{12}x_{21} + \gamma_8 x_{12}x_{22})c_2 \\ &+ (\gamma_9 x_{11}x_{21} + \gamma_{10} x_{11}x_{22} + \gamma_{11} x_{12}x_{21} + \gamma_{12} x_{12}x_{22})c_3. \end{aligned}$$

The model has 12 unknown coefficients (γ's) and is without an intercept.

Note that when $\boldsymbol{c}$ is fixed, the multiple-Scheffé model (8) reduces to (7). Moreover, the extended multiple-Scheffé model has a clear structure to represent the mixture-of-mixture relations; it is able to include all the possible blending properties among major and minor components. Even though this extension seems natural, the model has some limitations.

One limitation is that the multiple-Scheffé model in (8) does not apply to the case when the proportion of any major component is zero. In other words, every major component must have a positive lower bound L_i. However, in some MoM experiments, it is of interest to explore the region where some major components are absent. For example, if c_1 is 0 in (9), then the model becomes

$$(\gamma_{11}^2 x_{11}x_{21} + \gamma_{12}^2 x_{11}x_{22} + \gamma_{21}^2 x_{12}x_{21} + \gamma_{22}^2 x_{12}x_{22})c_2 +$$
$$(\gamma_{11}^3 x_{11}x_{21} + \gamma_{12}^3 x_{11}x_{22} + \gamma_{21}^3 x_{12}x_{21} + \gamma_{22}^3 x_{12}x_{22})c_3.$$

Since the first major component is absent, all of its minor components should also be absent. But the multiple-Scheffé model still contains the two minors x_{11} and x_{12}, which contradicts intuition.

There is yet another aspect of the multiple-Scheffé model that counters intuition. Take the multiple-Scheffé model in (9) as an example. Since the major component model is linear, there are no nonlinear blending (NLB) terms c_1c_2, c_1c_3, and c_2c_3 in the model. However, there are NLB terms between the minors from different majors, such as $x_{11}x_{21}c_1$. Clearly, if the minors of two majors have NLB terms, so should their majors because the minors are part of the majors. This problem will manifest itself whenever the model h for the major components is linear.

Furthermore, the number of terms or unknown parameters in multiple-Scheffé models are usually quite large and will increase rapidly as the number of minor or major components increases, or as the models for minors or majors become larger. As a result, large experimental designs are required to support multiple-Scheffé models.

3. MAJOR-MINOR MODEL

To model the mixture-of-mixture structure and overcome the limitations of the multiple-Scheffé model, we propose a new model for the general MoM experiments. Similar to multiple-Scheffé models, we independently specify the model h for major components and model f_i for minor components of major component i. In this proposed model, a Scheffé model is used for h. The linear and the quadratic Scheffé models are given by

$$h(c_1, \ldots, c_M) = \sum_{i=1}^{M} \alpha_i c_i,$$
$$h(c_1, \ldots, c_M) = \sum_{i=1}^{M} \alpha_i c_i + \sum_{1 \leq i < j \leq M} \alpha_{ij} c_i c_j.$$

Different from multiple-Scheffé model, we assume that the coefficients of major components depend only on their respective minor components. Let $\alpha_i = f_i(x_{i1}, \ldots, x_{im_i})$ and $\alpha_{ij} = f_i(x_{i1}, \ldots, x_{im_i}) \times f_j(x_{j1}, \ldots, x_{jm_j})$. As in the multiple-Scheffé model, $f_i \equiv 1$ if $m_i = 1$. Thus, the proposed model can be written as follows:

$$G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = \sum_{i=1}^{M} f_i(x_{i1}, \ldots, x_{im_i}) c_i \tag{10}$$
$$G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = \sum_{i=1}^{M} f_i(x_{i1}, \ldots, x_{im_i}) c_i + \sum_{1 \leq i < j \leq M} f_i(x_{i1}, \ldots, x_{im_i}) f_j(x_{j1}, \ldots, x_{jm_j}) c_i c_j \tag{11}$$

Again, (10) and (11) are linear models because $\boldsymbol{\gamma}$ represent the unknown coefficients in the expanded polynomials. By assuming the coefficients of major components' proportions c_i and their nonlinear blending properties $c_i c_j$ are the functions of their own minors, this model clearly incorporates the mixture-of-mixture structure. We call it the *major-minor* model. In the following section, we will explain the major-minor model in detail under the two scenarios: fixed and variable $\boldsymbol{c}$.

3.1 Fixed Major Components

In this part, we study the major-minor model when the major proportions $\boldsymbol{c}$ are fixed. Consider an example: $M = 2$, $m_1 = m_2 = 2$. Assume the model for major components are quadratic Scheffé models, and f_1, f_2 are of first order. If f_1 and f_2 are Scheffé models, then the major-minor model is given by

$$\begin{aligned} G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = & (\gamma_1 x_{11} + \gamma_2 x_{12}) c_1 + (\gamma_3 x_{21} + \gamma_4 x_{22}) c_2 \\ & + (\gamma_5 x_{11} x_{21} + \gamma_6 x_{12} x_{21} + \gamma_7 x_{11} x_{22} + \gamma_8 x_{12} x_{22}) c_1 c_2. \end{aligned}$$

When c_1, c_2, and c_3 are fixed, the major-minor model reduces to:

$$\gamma_1 x_{11} + \gamma_2 x_{12} + \gamma_3 x_{21} + \gamma_4 x_{22} + \gamma_5 x_{11} x_{21} + \gamma_6 x_{12} x_{21} + \gamma_7 x_{11} x_{22} + \gamma_8 x_{12} x_{22} \tag{12}$$

Note that a linear model f_i can be supported by two design points (x_{i1}^j, x_{i2}^j) $j = 1, 2$. Taking all the combinations of (x_{i1}^j, x_{i2}^j) $j = 1, 2$ and $i = 1, 2$, we find that there are only four combinations. Intuitively, these four points should be enough to support the reduced model when $\boldsymbol{c}$ are fixed. However, (12) is not identifiable by the four-point design. This identifiability problem can be avoided by using models with independent variables instead of the Scheffé models. Suppose we use $f_i = \beta_0 + \beta_1 x_{i1}$ $i = 1, 2$, then the major-minor model is given by

$$G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}') = (\gamma_1' + \gamma_2' x_{11})c_1 + (\gamma_3' + \gamma_4' x_{21})c_2 + (\gamma_5' + \gamma_6' x_{11} + \gamma_7' x_{21} + \gamma_8' x_{11} x_{21})c_1 c_2. \quad (13)$$

When $\boldsymbol{c}$ is fixed, (13) is reduces to

$$\gamma_0' + \gamma_1' x_{11} + \gamma_2' x_{21} + \gamma_3' x_{11} x_{21},$$

which can be estimated using the four-point design. As stated below, this result holds more generally. The proof is given in the Appendix. Let $\tilde{f}_i$ be f_i without the intercept and $\tilde{M}$ be the number of major components whose $m_i > 1$. Also, denote the reduced major-minor model by $f(\boldsymbol{x}, \boldsymbol{\gamma})$.

Proposition 1. *If each f_i is a model with independent variables, then the major-minor model reduces to:*

$$f(\boldsymbol{x}, \boldsymbol{\gamma}) = \gamma_0 + \sum_{i=1}^{\tilde{M}} \tilde{f}_i(x_{i1}, \ldots, x_{im_i}),$$

and

$$f(\boldsymbol{x}, \boldsymbol{\gamma}) = \gamma_0 + \sum_{i=1}^{\tilde{M}} \tilde{f}_i(x_{i1}, \ldots, x_{im_i}) + \sum_{1 \le i < j \le \tilde{M}} \tilde{f}_i(x_{i1}, \ldots, x_{im_i}) \times \tilde{f}_j(x_{j1}, \ldots, x_{jm_j}),$$

when h is a linear Scheffé model and a quadratic Scheffé model, respectively. Moreover, $f(\boldsymbol{x}, \boldsymbol{\gamma})$ can always be supported by a crossed design (taking all the combinations) of all the designs supporting the models $f_1, \ldots, f_M$.

3.2 Varied Major Components

The advantages of the major-minor model become more obvious when the proportions of the major components are also varied. Similar to the multiple-Scheffé model, we can

Table 4: Number of terms contained in the multiple-Scheffé model and the MoM model

	multiple-Scheffé				major-minor			
	1	2	3	4	1	2	3	4
1	6	12	18	24	6	9	12	15
2	12	24	36	48	9	13	17	21
3	18	36	54	72	12	17	22	27
4	24	48	72	96	15	21	27	33

choose linear, quadratic, or even cubic models for h and f_i. For the Pringles experiment, if we choose linear models for both major and minor components, the major-minor model is given by:

$$G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = (\gamma_1 + \gamma_2 x_{11})c_1 + (\gamma_3 + \gamma_4 x_{21})c_2 + \gamma_5 c_3. \tag{14}$$

If we use linear f_i and quadratic h, the major-minor model is:

$$\begin{aligned} G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = {} & (\gamma_1 + \gamma_2 x_{11})c_1 + (\gamma_3 + \gamma_4 x_{21})c_2 + \gamma_5 c_3 + (\gamma_{10} + \gamma_{11} x_{11})c_1 c_3 \\ & + (\gamma_{12} + \gamma_{13} x_{21})c_2 c_3 + (\gamma_6 + \gamma_7 x_{11} + \gamma_8 x_{21} + \gamma_9 x_{11} x_{21})c_1 c_2. \end{aligned} \tag{15}$$

Of course, it is not necessary that all minor models f_i have to be the same. Different models can be used for minors of different major components based on physical knowledge. This proposed model can also study the inter- and intra- major component blending properties throughly. For the simplest linear models as in (14), the blending properties of a major component with its own minors can be studied; for linear f_i and quadratic h as in (15), the blending properties of a major component with its own minors, and the blending properties between any two different majors and their minors can be studied.

The major-minor model is a much smaller model compared to the multiple-Scheffé model. In Table 4, we compare the number of terms contained in the multiple-Scheffé model and the major-minor model. Suppose there are three major components with $m_3 = 1$. The number of minor components m_1 and m_2 are varied from 1 to 4 along the row and column directions in Table 4. We use linear models for minor components and quadratic models for major components. Obviously, the major-minor model contains much less terms than the multiple-Scheffé model and thus it is more economical.

Note that if $c_i = 0$ in the major-minor model, then all the minor components of major i will disappear from the model. This makes intuitive sense and is not shared by the multiple-Scheffé model. For example, if $c_1 = 0$ in (14) or (15), all the terms related to this major component and its minors are removed. Also, the NLB terms between minors belonging to different majors will be included if and only if the NLB term between the two majors is included.

Example 1. In this example, we use the photoresist-coating experiment that has been studied in Cornell and Ramsey (1998). This experiment is a MoM experiment involving two major components, each of which has two minor components. This experiment was used to illustrate the "quadratic × quadratic × quadratic" multiple-Scheffé model. It contains 27 terms in all. Correspondingly, the design contains $3^3 = 27$ points and is a crossed design of three 3-point designs supporting each quadratic model. The three points for (x_{i1}, x_{i2}) $i = 1, 2$ are (0,1), (0.5, 0.5), and (1,0). For (c_1, c_2) the design is (0.75,0.25), (0.5,0.5), and (0.25, 0.75). Some of the points in the crossed design are replicated twice, which resulted in a 42-run experiment.

Using quadratic f_1, f_2, and h, the major-minor model is given by

$$G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = (\gamma_1 + \gamma_2 x_{11} + \gamma_3 x_{11}^2)c_1 + (\gamma_4 + \gamma_5 x_{21} + \gamma_6 x_{21}^2)c_2 + (\gamma_7 + \gamma_8 x_{11} + \gamma_9 x_{21} + \gamma_{10} x_{11}^2 + \gamma_{11} x_{11} x_{21} + \gamma_{12} x_{21}^2 + \gamma_{13} x_{11}^2 x_{21} + \gamma_{14} x_{11} x_{21}^2 + \gamma_{15} x_{11}^2 x_{21}^2)c_1 c_2,$$

which contains only 15 parameters. To fit this model, 15 points are selected from the set of 27 points according to the D-optimal criterion. If any of the selected points is replicated in the original design, we also include its replication. This resulted in a new design containing 24 runs. The remaining 18 runs are used to test the prediction of the major-minor model. In Figure 1, the predicted values are plotted against the observed values. The predictions are reasonably close to the observations indicating good prediction accuracy. The multiple-Scheffé model used in Cornell and Ramsey (1998) has an R^2 of 0.9996, which is larger than that of the major-minor model ($R^2 = 0.9989$). This is because the multiple-Scheffé model contains 27 terms, which is equal to the number of design points. Therefore, it will produce an interpolating model if there were no replications. The major-minor model contains 12 less terms, but achieves comparable performance in fitting and prediction.

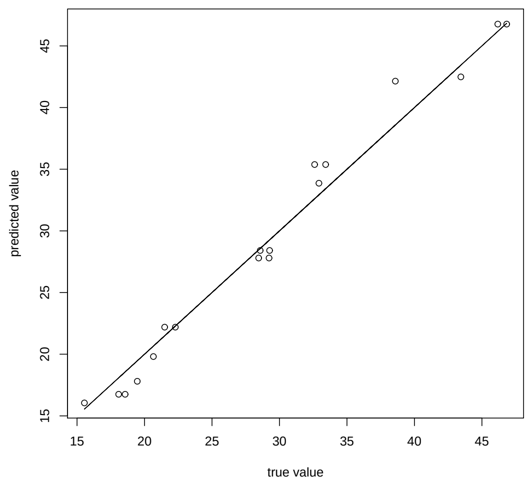

Figure 1: *Actual v.s. prediction from the major-minor model for the 18 points.*

4. MAJOR-MINOR D-OPTIMAL DESIGN

In design of mixture experiments, when there are more complicated constraints and the design space is a non-simplex polyhedron, D or I-optimal designs are used. For the major-minor model, we can proceed in the same way. Recall that in specifying the major-minor model, we separately specify a model for major components h and models for minor components f_i, and then "combine" them into the major-minor model. This suggests the following two-stage design strategy: (i) specify the design D_0 of the major components to support h, and the designs D_i of the minor components to support f_i, $i = 1, 2, \ldots, M$; (ii) "combine" the design D_0, D_1, $\ldots$, D_M to support the major-minor model.

To construct designs $D_0, D_1, \ldots, D_M$, we can use the methods for constructing the classic mixture experiments. If the design space of $(c_1, \ldots, c_M)$ or $(x_{i1}, \ldots, x_{iM})$ is simplex, we simply use a simplex-lattice design to construct D_i, $i = 0, 1, \ldots, M$. For non-simplex design space, optimal designs with D or I criteria are usually used. Another method is to choose the design points following the recommendations from Snee and Marguardt (1974) and Snee (1975), all of which are summarized in Chapter 4 of Cornell (2002). The rule of thumb is that an efficient design consists of a subset of the extreme vertices, edge centroids, constraint plane centroids, and the overall centroid. If M, m_1, $\ldots$, m_M are not large and the constraints are not complex, we can compute the vertices and centroids by hand; otherwise, Cornell (2002) has summarized many algorithms such as CONVRT (Piepel, 1988) that can be used to compute them.

The next stage is to "combine" $D_0, D_1, \ldots, D_M$ into one design. First of all, it is always desirable to include all the combinations of the design points, i.e., $D_0 \otimes D_1 \otimes \cdots \otimes D_M$, where $\otimes$ stands for Kronecker product. Therefore, if the budget allows, we can use the crossed design $D_0 \otimes D_1 \otimes \cdots \otimes D_M$. Otherwise, a subset of the crossed design that can support the major-minor model can be used. The optimal subset of runs can be found by using criteria such as D, A, or E. Here we choose the D-optimal design criterion. It can be easily done using an existing optimal design algorithm since the candidate set is already provided. We name our proposed design the *major-minor D-optimal* design. An example is given below.

Example 2. Didier et. al. (2007) carried out a MoM experiment to develop complex culture media used in recombinant protein production. Two categories of components (major components) are involved: hexose (H) and energy provider (E). Hexose is a mixture of three hexoses H_1, H_2, and H_3. Energy provider is a mixture of three energy providers E_1, E_2, and E_3. In their paper, the designs for the minor components have been provided. The design S_1 for the hexoses is a 10-point $\{3, 3\}$ simplex-lattice design as shown in Figure 2(a). The design S_2 for the energy providers is a constrained simplex design with a centroid as shown in Figure 2(b). However, the amounts of the two major components are fixed at $c_H = 83.1\%$ and $c_E = 16.9\%$ in their experiments. Here we assume that c_H and c_E can be varied at three values $(c_H, c_E) = (0.748, 0.252)$, $(0.831, 0.169)$, and $(0.914, 0.086)$ as in Figure 2(c), denoted as design S_3.

Take the crossed design as the candidate set, which contains $10 \times 5 \times 3 = 150$ points. Next we choose the models so that we can decide the run size of the experiment accordingly. We assume a quadratic model f_1 for H_1, H_2, H_3, and a linear model f_2 for E_1, E_2, E_3. For the major components we assume a quadratic model. In all, the major-minor model

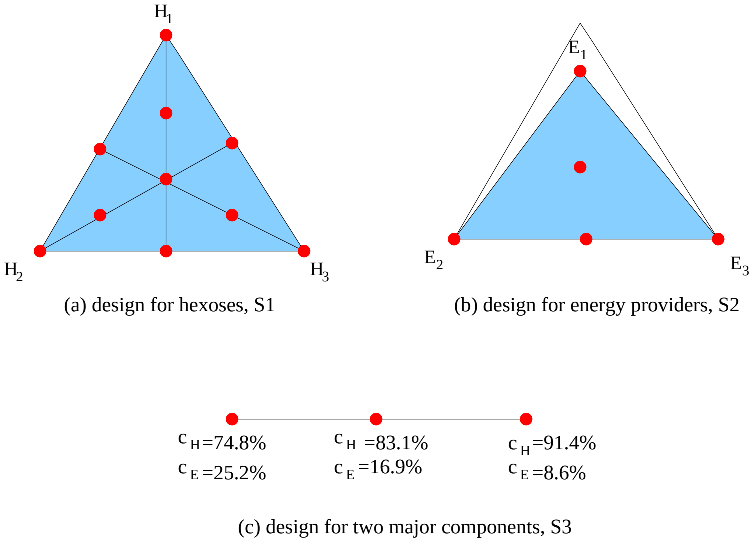

Figure 2: *Designs for major and minor components.*

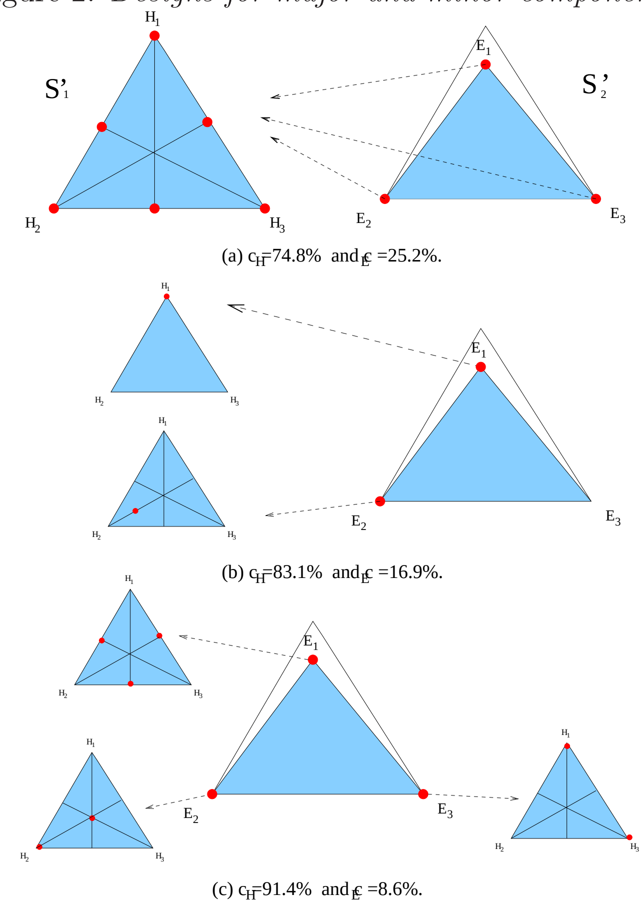

Figure 3: *D-optimal design containing 27 points.*

to be fitted contains 27 terms and is given by

$$
\begin{aligned}
G(\boldsymbol{x},\boldsymbol{c},\boldsymbol{\gamma}) = {} & (\gamma_1 + \gamma_2 x_{11} + \gamma_3 x_{12} + \gamma_4 x_{11}^2 + \gamma_5 x_{12}^2 + \gamma_6 x_{11}x_{12})c_1 + (\gamma_7 + \gamma_8 x_{21} + \gamma_9 x_{22})c_2 \\
& + (\gamma_{10} + \gamma_{11}x_{11} + \gamma_{12}x_{12} + \gamma_{13}x_{11}^2 + \gamma_{14}x_{12}^2 + \gamma_{15}x_{11}x_{12} + \\
& + \gamma_{16}x_{21} + \gamma_{17}x_{11}x_{21} + \gamma_{18}x_{12}x_{21} + \gamma_{19}x_{11}^2 x_{20} + \gamma_{20}x_{12}^2 x_{21} + \gamma_{21}x_{11}x_{12}x_{21} + \\
& + \gamma_{22}x_{21} + \gamma_{23}x_{11}x_{21} + \gamma_{24}x_{12}x_{21} + \gamma_{25}x_{11}^2 x_{20} + \gamma_{26}x_{12}^2 x_{21} + \gamma_{27}x_{11}x_{12}x_{21})c_1 c_2.
\end{aligned}
$$

Since the model has 27 coefficients, the D-optimal design can contain as few as 27 points. In Figure 3, the 27-point D-optimal design is plotted. Figure 3(a) shows that when $c_1 = 74.8\%$ and $c_2 = 25.2\%$, the minors E_1, E_2, E_3 are chosen to be the three vertices of S_2, denoted as design S_2'. At each of the vertices, the three vertices and three mid-edge points from S_1 are selected for H_1, H_2, H_3, denoted as design S_1'. Therefore, when $(c_H, c_E) = (0.748, 0.252)$, the design is a crossed design $S_1' \otimes S_2'$ of 18 points in all. The remaining 9 design points are shown in Figure 3 (b) and (c). Increasing the run size to 54, the augmented D-optimal design is just the crossed design of $S_1' \otimes S_2' \otimes S_3$.

5. PRINGLES MOM EXPERIMENTS

Consider the Prinlges MoM experiment described in the Introduction. The purpose of this experiment is to reformulate the Pringles potato crisps for the new product to be packaged in a bag. Many quality characteristics are measured for each experimental run, and we focus on percentage of fat (% Fat) and hardness of the potato crisps (Hardness). Among them, Hardness is of great importance because the package of the new crisp is changed from a canister to a bag, which makes the potato crisps more susceptible to breaking. To avoid this, we want to find the optimal formulation so that the Hardness is maximized, while at the same time reducing the percentage of fat.

Due to the time and budget constraints, the total number of runs can not exceed 16. Here we assume the same model f_i for minors of both A and B. Therefore, only the major-minor model (14) and (15) can be considered. Here we prefer the more parsimonious model (14). The design of this experiment follows the strategy in Section 3. Because f_1 and f_2 are both first order models, two design points are sufficient. Thus, according to the

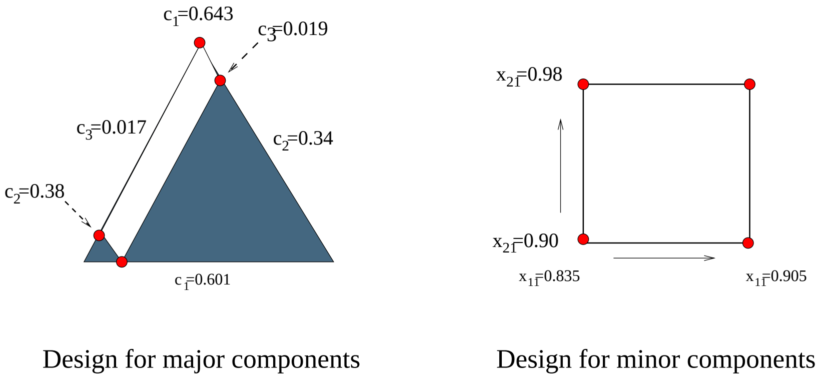

Figure 4: *Design D_0 (left) and $D_1 \otimes D_2$ (right).*

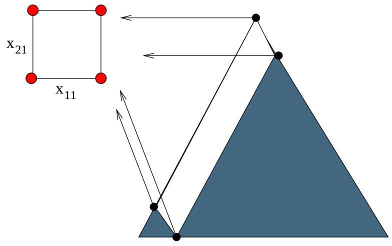

Figure 5: *Design for Pringles Experiment D.*

boundary conditions on (x_{11}, x_{12}) and (x_{21}, x_{22}), we choose $D_1(x_{11}) = (0.835, 0.905)$ and $D_2(x_{21}) = (0.9, 0.98)$. The design space of the major components is a two-dimensional parallelogram, which has four vertices. To fit a linear model h, it is sufficient for D_0 to consists of the four of vertices. Figure 4 contains the design D_0 and the crossed design $D_1 \otimes D_2$ for the four minor components. For this experiment, we can simply take the crossed design $D = D_0 \otimes D_1 \otimes D_2$ as the combined design, since its total number of runs is 16. Design D is graphed in Figure 5. The experimental design and the measurements are shown in Table 5.

Next, we fit the major-minor model (14). The coefficient estimates and test statistics are shown in Table 6. The corresponding multiple-Scheffé model given in (9) has 12 terms. In Table 7, we compare the fitting of the two models. The R^2 for the multiple-Scheffé model is larger than that of the major-minor model. However, this could be due to over fitting because the multiple-Scheffé model has seven more terms than the major-minor model. To check the prediction performance, we computed the leave-one-out cross validation errors $cv_i = y_i - \hat{y}_{-i}$, where $\hat{y}_{-i}$ is the prediction at the ith point by the model

Table 5: Pringles MoM Experiment Data

Run	x_{11}	x_{12}	x_{21}	x_{22}	c_1	c_2	c_3	% Fat	Hardness
1	0.835	0.165	0.90	0.10	0.603	0.38	0.017	35.040	4.835
2	0.835	0.165	0.90	0.10	0.643	0.34	0.017	32.100	6.375
3	0.835	0.165	0.98	0.02	0.603	0.38	0.017	37.800	3.625
4	0.835	0.165	0.98	0.02	0.643	0.34	0.017	33.300	5.500
5	0.905	0.095	0.90	0.10	0.643	0.34	0.017	31.320	6.875
6	0.905	0.095	0.90	0.10	0.603	0.38	0.017	34.026	5.250
7	0.835	0.165	0.90	0.10	0.601	0.38	0.019	34.140	5.000
8	0.835	0.165	0.90	0.10	0.641	0.34	0.019	31.968	6.250
9	0.835	0.165	0.98	0.02	0.601	0.38	0.019	36.990	3.625
10	0.905	0.095	0.98	0.02	0.603	0.38	0.017	35.970	4.250
11	0.905	0.095	0.90	0.10	0.601	0.38	0.019	33.870	5.250
12	0.835	0.165	0.98	0.02	0.641	0.34	0.019	33.438	4.875
13	0.905	0.095	0.98	0.02	0.643	0.34	0.017	33.144	4.940
14	0.905	0.095	0.90	0.10	0.641	0.34	0.019	32.106	6.165
15	0.905	0.095	0.98	0.02	0.641	0.34	0.019	33.660	5.565
16	0.905	0.095	0.98	0.02	0.601	0.38	0.019	35.520	4.875

fitted without the ith point. The mean squared cross validation $MSCV = \frac{1}{n}\sum_{i=1}^{n} cv_i^2$ is also shown in Table 7. It is much smaller for the major-minor model than that of the multiple-Scheffé model. This clearly shows that the major-minor model is a much better model than the multiple-Scheffé model.

At last, we find the best formulation that can maximize the hardness of the crisp while reducing the percentage of the fat. The fitted major-minor models for % Fat and Hardness are:

$$\% \text{ Fat} = (22.611 - 14.440x_{11})c_1 + (17.051 + 66.753x_{21})c_2 - 52.951c_3,$$

$$\text{Hardness} = (8.786 + 8.658x_{11})c_1 + (20.966 - 37.641x_{21})c_2 + 13.506c_3.$$

To maximize Hardness, the optimal setting is $(x_{11}, x_{12}, x_{21}, x_{22}, c_1, c_2, c_3)$=(0.905, 0.095, 0.9, 0.1, 0.643, 0.34, 0.017), which is one of the vertices of the polyhedron (run # 5 in Table 5). At this setting, the predicted Hardness and % Fat are 6.528 and 31.460. To minimize the % Fat, the optimal setting is $(x_{11}, x_{12}, x_{21}, x_{22}, c_1, c_2, c_3)$=(0.905, 0.095, 0.9, 0.1, 0.641, 0.34, 0.019), which is also a vertex and very close to the previous optimal solution of Hardness (run # 14 in Table 5). At this vertex, the predicted Hardness and % Fat are 6.521 and 31.335. Because the observed values of Hardness and % Fat are slightly

Table 6: Coefficients Estimates of the Major-minor Model

	% Fat			Hardness		
Terms	Coeff. Est.	t-value	Pr(>|t|)	Coeff. Est.	t-value	Pr(>|t|)
c_1	22.611	3.522	0.005 *	8.786	2.054	0.065
c_2	17.051	1.682	0.121	20.966	3.105	0.010 *
c_3	-52.951	-0.398	0.698	13.506	0.152	0.882
c_1x_{11}	-14.440	-2.321	0.040 *	8.658	2.089	0.061
c_2x_{21}	66.753	7.105	0.000 *	-37.641	-6.014	0.000 *

Table 7: Comparison of the major-minor and multiple-Scheffé model

	% Fat		Hardness	
	Major-Minor	Multiple-Scheffé	Major-Minor	Multiple-Scheffé
MSE	0.2938	0.2068	0.1303	0.0612
R^2	0.9364	0.9837	0.8877	0.9808
$MSCV$	0.4277	0.8273	0.1894	0.2447

better for run # 5, we may choose the first optimal setting as the new formulation of the Pringles crisp.

6. SIMULATION STUDY

In this section, we study the prediction accuracy of the major-minor model and the multiple-Scheffé model using simulations. Consider a MoM experiment that has two major components A and B, each of which has two minor components A_1, A_2 and B_1, B_2, whose corresponding proportions are X_{A_1}, X_{A_2}, X_{B_1}, and X_{B_2}. They satisfy the constraint $X_{A_1} + X_{A_2} + X_{B_1} + X_{B_2} = 1$. We assume two models:

I. $f_I(\boldsymbol{X}) = 3X_{A_1} + X_{A_2} + 5X_{B_1} + X_{B_2} + 12X_{A_1}X_{B_1}$, and

II. $f_{II}(\boldsymbol{X})y = 5 + (X_{A_1} - 0.5)^2$,

and simulate data from the two models by adding the random error $y = f_k(\boldsymbol{X}) + \epsilon$, $k = I, II$, where $\epsilon \sim N(0, 1)$. In addition, we assume two boundary conditions for the major components: first with $0 \leq c_i \leq 1$ and second with $0.25 \leq c_i \leq 0.75$, for $i = 1, 2$, and then fit the resulting data using multiple-Scheffé model and major-minor model. Since both of the true underlying models are different from the multiple-Scheffé and major-minor models, we can make a reasonable comparison between their performances.

We use linear f_1, f_2, and h for both multiple-Scheffé and major minor models:

$$\text{Multiple-Scheffé:} \quad G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = (\gamma_1 x_{11} x_{21} + \gamma_2 x_{11} x_{22} + \gamma_3 x_{12} x_{21} + \gamma_4 x_{12} x_{22}) c_1 + (\gamma_5 x_{11} x_{21} + \gamma_6 x_{11} x_{22} + \gamma_7 x_{12} x_{21} + \gamma_8 x_{12} x_{22}) c_2, \tag{16}$$

$$\text{Major-Minor:} \quad G(\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\gamma}) = (\gamma_1 + \gamma_2 x_{11}) c_1 + (\gamma_3 + \gamma_4 x_{21}) c_2. \tag{17}$$

The design given in Table 8 contains 16 runs which can support both (16) and (17). The columns c_1 and c_2 are the major components for the first boundary condition; columns c_1' and c_2' are for second boundary conditions. In fact, c_1' and c_2' are values of c_1 and c_2 scaled into interval $[0.25, 0.75]$. Figure 6 shows the comparisons of the mean squared prediction error, $1/n \sum_{i=1}^{n} (f_k(\boldsymbol{X}_i) - \hat{y}_i)^2$, between the two models (16) and (17) using 100 data sets simulated from Model I and II. Even though the major-minor model has only half of the terms compared to multiple-Scheffé model, it gives a superior fit to the data for both the models I and II and both the boundary conditions. The improvement obtained by using major-minor model is quite substantial for model II. It can also be seen that the performance of multiple-Scheffé model deteriorates when the major components take values closer to 0 (as in the first boundary condition).

7. CONCLUSION

Multiple-Scheffé models have certain limitations for modeling in a general mixture-of-mixture problem. In this article, we have introduced a new type of model called major-minor model to overcome these limitations. The new model is smaller and thus, can be estimated with fewer data. This saves a lot of time and money for the experimenter. Moreover, the new model captures the mixture-of-mixture structure and can work even when some of the major components are absent in the mixture. We have also proposed a D-optimal-based strategy to efficiently design mixture-of-mixture experiments. The modeling and design approach was then successfully applied in a real problem of optimizing the formulation of a potato crisp.

Table 8: Design for simulation Model I.

Run	x_{11}	x_{12}	x_{21}	x_{22}	c_1	c_2	c'_1	c'_2
1	0	0	1	0	0.00	1.00	0.25	0.75
2	0	0	0	1	0.00	1.00	0.25	0.75
3	1	0	0	0	1.00	0.00	0.75	0.25
4	0	1	0	0	1.00	0.00	0.75	0.25
5	0	1	1	0	0.25	0.75	0.375	0.625
6	0	1	0	1	0.25	0.75	0.375	0.625
7	1	0	1	0	0.25	0.75	0.375	0.625
8	1	0	0	1	0.25	0.75	0.375	0.625
9	0	1	1	0	0.50	0.50	0.5	0.5
10	0	1	0	1	0.50	0.50	0.5	0.5
11	1	0	1	0	0.50	0.50	0.5	0.5
12	1	0	0	1	0.50	0.50	0.5	0.5
13	0	1	1	0	0.75	0.25	0.625	0.375
14	0	1	0	1	0.75	0.25	0.625	0.375
15	1	0	1	0	0.75	0.25	0.625	0.375
16	1	0	0	1	0.75	0.25	0.625	0.375

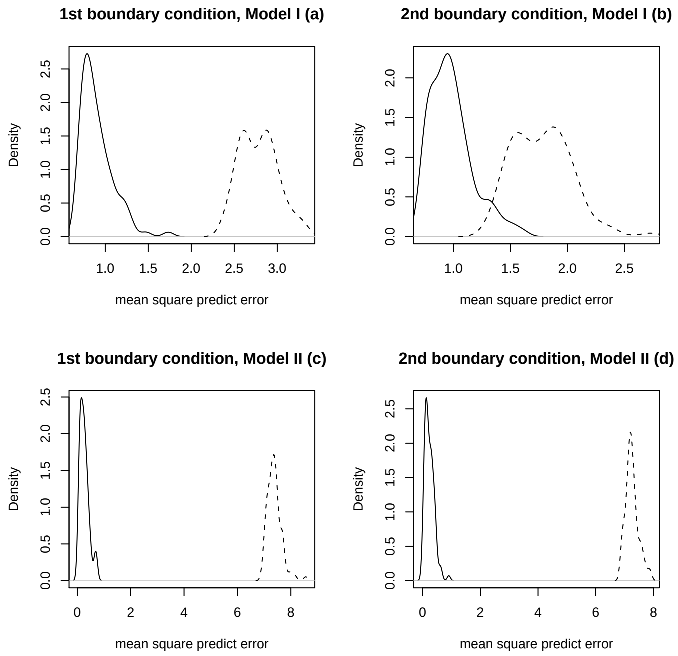

Figure 6: *Comparison of multiple-Scheffé model (dashed) and major-minor model (solid) for the 100 simulated data sets from Models I and II with two boundary conditions.*

ACKNOWLEDGEMENTS

This research was supported by U.S. National Science Foundation grant CMMI-0448774.

APPENDIX: PROOF OF PROPOSITION 1

Write the model f_i in row vector form–$[1, \boldsymbol{f}_i]$, in which 1 is the intercept, $\boldsymbol{f}_i$ is a row vector contains all the other terms. For example, if f_i is a 2nd order polynomial for a three minors x_{i1}, x_{i2}, x_{i3}, whose sum is 1, then f_i can be written as $[1, x_{i1}, x_{i2}, x_{i1}x_{i2}, x_{i1}^2, x_{i2}^2] = [1, \boldsymbol{f}_i]$.

First, we show that $f(\boldsymbol{x}, \boldsymbol{\gamma})$ is a sub-model contained in the product model $f_1 \times f_2 \times \cdots \times f_M$. The product model can also be written as

$$\begin{aligned} f_1 \times \cdots \times f_M &= [1, \boldsymbol{f}_i] \otimes \cdots \otimes [1, \boldsymbol{f}_M] \\ &= [1, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_M, \boldsymbol{f}_1 \otimes \boldsymbol{f}_2, \ldots, \boldsymbol{f}_{M-1} \otimes \boldsymbol{f}_M, \ldots]. \end{aligned}$$

If model h for major components is linear, $f(\boldsymbol{x}, \boldsymbol{\gamma})$ contains $[1, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_M]$; if h is quadratic, $f(\boldsymbol{x}, \boldsymbol{\gamma})$ contains $[1, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_M, \boldsymbol{f}_1 \otimes \boldsymbol{f}_2, \ldots, \boldsymbol{f}_{M-1} \otimes \boldsymbol{f}_M]$; both of which are sub-models of the product model. Similar proof can be applied when h is cubic or higher order model.

Next, we show that the product model is always supported by a crossed design. Let D_i be a n_i-run design supporting f_i and U_i be the model matrix for f_i. The crossed design, i.e., the design taking all the combinations for all the D_i, can be written as $D_1 \otimes D_2 \otimes \cdots \otimes D_M$. The model matrix for the product model is $U = U_1 \otimes U_2 \otimes \cdots \otimes U_M$. Note that

$$(U^T U)^{-1} = (U_1^T U_1)^{-1} \otimes \cdots \otimes (U_M^T U_M)^{-1}.$$

Because $U_i^T U_i$ is invertible for $i = 1, \ldots, M$, $U^T U$ is also invertible. Thus, the crossed design always supports the product model. Because $f(\boldsymbol{x}, \boldsymbol{\gamma})$ is a sub-model of the product model, it is also supported by the crossed design.

References

Borges, C., Bruns, E. R., Almeida, A. A., and Scarminio, I. S. (2007), "Mixture-mixture design for fingerprint optimization of chromatographic mobile phases and extraction solutions for Camellia sinensis," *Analytica Chimica Acta*, 595, 28-37.

Cornell, J. A., and Good, I. J. (1970), "The Mixture Problem for Categorized Components," *Journal of the American Statistical Association*, 65, 339-355.

Cornell, J. A. (1971), "Process Variables in the Mixture Problem for Categorized Components," *Journal of the American Statistical Association*, 66, 42-48.

Cornell, J. A., and Ramsey, P. J. (1998), "A Generalized Mixture Model for Categorized-Components Problems With an Application to a Photoresist-Coating Experiment," *Technometrics*, 40, 48-61.

Cornell, J. A. (2002), *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data* (3rd ed.), New York: Wiley.

Cornell, J. A., and Gorman, J. W. (2003), "Two New Mixture Models: Living Collinearity but Removing Its Influence," *Journal of Quality Technology*, 35, 78-88.

Didier, C., Etcheverrigaray, M., Kratje, R., and Goicoechea, H. C. (2007), "Crossed mixture design and multiple response analysis for developing complex culture media used in recombinant protein production," *Chemometrics and Intelligent Laboratory Systems*, 86, 1-9.

Dingstad, G., Egelandsdal, B., and NæS, T. (2003), "Modeling methods for corssed mixture experiments—a case study from sausage production," *Chemometrics and intelligent laboratory systems*, 66, 175-190.

Lambrakis, D. P. (1968), "Experiments with Mixtures: A Generalization of the Simplex-Lattice Design," *Journal of the Royal Statistical Society*, Ser. B, 30, 123-136.

Lambrakis, D. P. (1969), "Experiments with Mixtures: Estimated Regression Function of the Multiple-Lattice Design," *Journal of the Royal Statistical Society*, Ser. B, 31, 276-284.

Piepel, G. F. (1988), "Programs for Generating Extreme Vertices and Centroids of Linearity Constrained Experimental Regions," *Journal of Quality Technology*, 20, 125-139.

Piepel, G. F. (1999), "Modeling Methods for Mixture-of-Mixtures Experiments Applied to a Tablet Formulation Problem," *Pharmaceutical Development and Technology*, 4, 593-606.

Prescott, P., Dean, A. M., Draper, N. R., and Lewis, S. M. (2002), "Mixture Experiments: ILL-Condioning and Quadratic Model Specification," *Technometrics*, 44, 260-268.

Snee R. D., and D. W. Marquardt (1974), "Extreme Vertices Designs for Linear Mixture Models," *Technometrics*, 16, 399-408.

Snee R. D. (1975) "Experimental Designs for Quadratic Models in Constrained Mixture Spaces," *Technometrics*, 17, 149-159.