Design and Analysis of Computer Experiments with Branching and Nested Factors

Ying Hung and V. Roshan Joseph
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332
yhung@isye.gatech.edu, roshan@gatech.edu

Shreyes N. Melkote
School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332
smelkote@gatech.edu

Abstract

In many experiments, some of the factors exist only within the level of another factor. Such factors are often called nested factors. A factor within which other factors are nested is called a branching factor. For example, suppose we want to experiment with two processing methods. The factors involved in these two methods can be different. Thus, in this experiment the processing method is a branching factor and the other factors are nested within the branching factor. Design and analysis of experiments with branching and nested factors are challenging and have not received much attention in the literature. Motivated by a computer experiment in a machining process, we develop optimal Latin hypercube designs and kriging methods that can accommodate branching and nested factors. Through the application of the proposed methods, optimal machining conditions and tool edge geometry are attained, which resulted in a remarkable improvement in the machining process.

KEY WORDS: Finite element models; Kriging; Latin hypercube designs.

1. INTRODUCTION

Nested factors are those factors which exist only within the level of another factor. A factor within which other factors are nested is called a branching factor. For example, suppose we want to experiment with two surface preparation methods in printed circuit board (PCB) manufacturing: mechanical scrubbing and chemical treatment. Here, the surface preparation method is the branching factor. Mechanical scrubbing can be optimized by changing the pressure of the scrub and chemical treatment can be optimized by changing the micro-etch rate. The pressure and micro-etch rate are the nested factors. When designing an experiment, these two factors (pressure and micro-etch rate) will be collapsed into a single factor (i.e., they will be assigned to the same column in the design matrix). The physical meaning and levels of the nested factor depends on the corresponding level of the branching factor. Because nested factors can differ with respect to the level of branching factor, designing and analyzing experiments with such factors are not trivial.

Taguchi (1987) has proposed an innovative idea to design experiments with branching and nested factors. He called nested factors as pseudo-factors and the resulting designs as pseudo-factor designs. Phadke (1989) called the same as branching designs. The core idea was to carefully assign branching and nested factors to the columns of orthogonal arrays using linear graphs in such a way that their interactions are estimable. The interaction between branching and nested factors are important because the nested factors differ with respect to the levels of branching factors and thus their effects can change depending on the level of the branching factor. For example, suppose we choose two levels for the pressure and micro-etch rate in the PCB experiment. The quality of the PCB may increase with increase in pressure but may decrease with increase in micro-etch rate. This is shown in Figure 1(a). We can see strong interaction between the branching and nested factors. The interaction effect could be reduced if we knew the effects of the nested factors before the experiment. For example, if we interchange the two levels of pressure, then the interaction is not significant (see Figure 1 (b)). In general, the effects of the factors are not known before the experiment and therefore, we should design the experiments that are capable of estimating the branching-by-nested

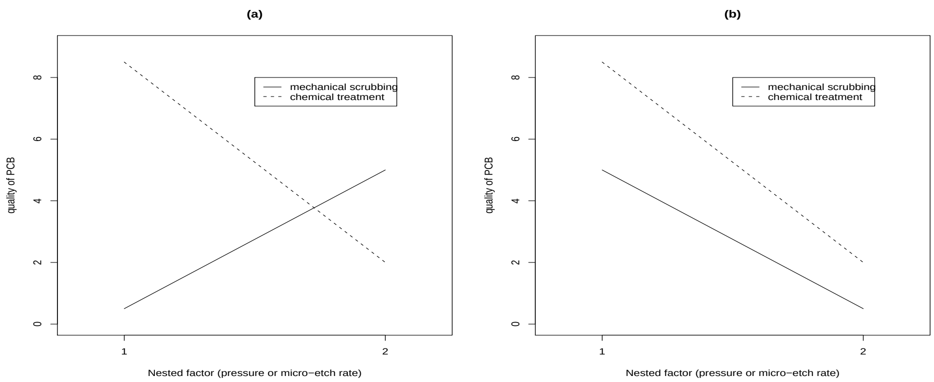

Figure 1: Illustration of branching-by-nested interaction. (a) when the effects are unknown and (b) when the effects are known.

interactions. Although Taguchi's approach using orthogonal arrays and linear graphs are very intuitive, they are not general enough to apply to more complex situations such as the design of a computer experiment.

The designs we consider here are different from the so-called nested designs in the literature (see, e.g., Hicks and Turner, 1999; Montgomery, 2001). In nested designs, the factors are assumed to be similar (for example, different batches of material nested within different suppliers). Therefore, the branching-by-nested interactions can be safely assumed to be negligible, which is not the case in the present problem. Moreover, nested designs are crossed designs and therefore, the complication arising due to the aliasing of effects is not an issue. In contrast, the main focus here is to efficiently design highly fractionated experiments in the space of branching, nested, and other factors. Furthermore, in nested designs the nested factors are usually not the effects of interests and are treated as random effects or block effects, whereas our objective is to simultaneously identify the optimal settings of branching, nested, and other factors.

Our work is motivated by a computer experiment that involves branching and nested factors. The objective of the experiment is to optimize a turning process for hardened bearing steel with a cBN cutting tool (see Figure 2). This process is commonly referred to as hard turning and is of considerable interest to bearing manufacturers as a potential replacement for the grinding process. Since the material being machined is very hard (hardness in excess of 60 Rockwell C), the cutting tool is subjected to large forces, stresses, and temperatures during the operation. In practice, the cutting edge of the tool is shaped such that it can withstand the severe conditions. Two commonly employed cutting edge shapes, hone and chamfer, are shown in Figure 3. Note that Figure 3 represent the idealized view of the instantaneous cutting action in the cross section A-A indicated in Figure 2. These cutting edge shapes are intended to strengthen the cutting edge to bear the large tool stresses generated in cutting. The chamfer tool design can be changed using two factors: chamfer length and chamfer angle, whereas the hone design is fixed. In other words, the two factors length and angle are nested within the chamfer edge and there are no factors nested within hone edge. In our terminology, the tool edge is a branching factor. Thus, when the branching factor takes the level chamfer, there are two additional factors present in the experiment; but when the branching factor takes the level hone, there are no factors. There are a few other factors that are common to both of the tool edges such as the cutting edge radius, tool nose radius, and rake angle. The machining parameters such as cutting speed, feed, and depth of cut are also factors that do not depend on the type of tool edges. To distinguish them from the branching and nested factors, we call them as shared factors. All of the factors involved in this experiment and their allowable ranges are shown in Table 1. The experiments can be performed in computers using a commercially available finite element software AdvantEdge.

Latin hypercube designs (LHDs) are commonly used in computer experiments (McKay, Beckman, and Conover, 1979). A desirable property of a LHD is its one-dimensional balance, i.e., when we project a N-point design onto any factor, there will be N different levels for that factor. Clearly this cannot be satisfied for branching and nested factors. The branching factor is usually a qualitative factor and therefore the number of levels of a branching factor is fixed (it does not depend on the number of runs N). Moreover, the nested factors are different for different levels of the branching factor. Therefore, we need a one-dimensional balance for the nested factors within each level of the branching factor. As an example, consider an experiment with one branching factor z_1, one nested factor v_1, and two shared factors x_1 and x_2. Suppose that the branching factor has two levels and that we want to do the experiment

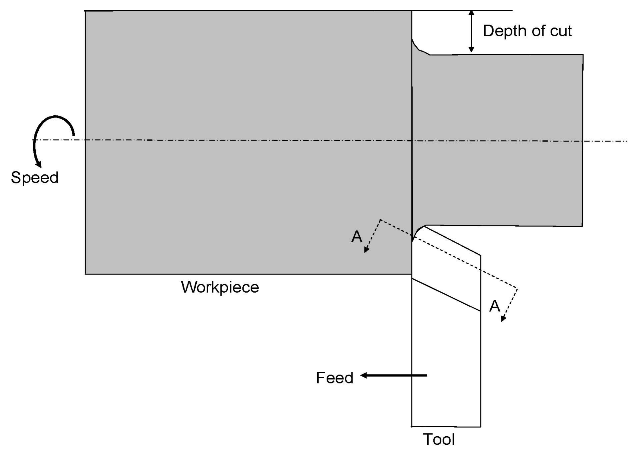

Figure 2: Schematic of turning process; A-A is a perpendicular section through the tool.

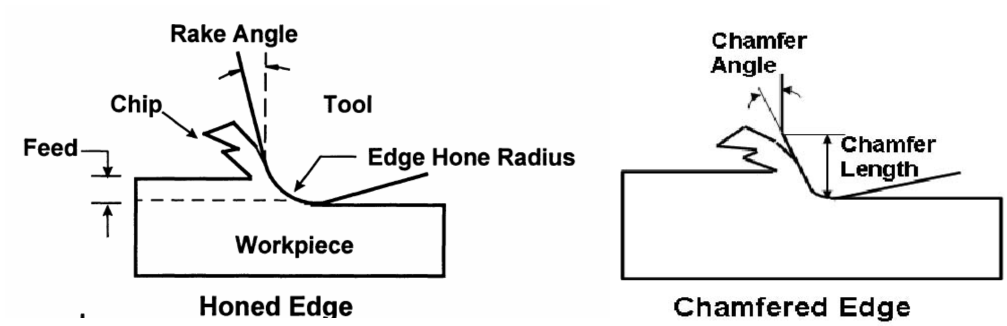

Figure 3: Illustration of hone and chamfer tool edges.

Table 1: Factors and their ranges in the hard turning experiment

Type of Factor	Notation	Factor	Ranges
Branching factor	z_1	Cutting edge shape	hone or chamfer
Nested factors	$v_1\|z_1$=chamfer	Angle (degree)	$17 \sim 20$
	$v_2\|z_1$=chamfer	Length (μm)	$115 \sim 140$
	$v_1\|z_1$=hone	None	None
	$v_2\|z_1$=hone	None	None
Shared factors	x_1	Cutting edge radius (μm)	$5 \sim 25$
	x_2	Rake angle (degree)	$-15 \sim -5$
	x_3	Tool nose radius (mm)	$0.4 \sim 1.6$
	x_4	Cutting speed (m/min)	$120 \sim 240$
	x_5	Feed (mm/rev)	$0.05 \sim 0.15$
	x_6	Depth of cut (mm)	$0.1 \sim 0.25$

Table 2: An Example of Branching Latin hypercube design

run	z_1	v_1	x_1	x_2
1	1	1	4	1
2	1	2	3	8
3	1	3	8	5
4	1	4	2	3
5	2	1	7	2
6	2	2	1	6
7	2	3	6	7
8	2	4	5	4

in eight runs. A possible design of experiment is shown in Table 2. We can see that the shared factors have the one-dimensional balance, because they take eight different levels in the experiment. The nested factor has a one-dimensional balance within each level of the branching factor. Note that v_1 represents two different factors, one when $z_1 = 1$ and the other when $z_1 = 2$. Therefore, $v_1 = 1$ in run 1 is not the same as $v_1 = 1$ in run 5. In the next section, we discuss some general strategies for designing such experiments using LHDs.

It is well known that not all LHDs are "good". Most of the research in this area has focused on finding "good" LHDs based on some optimal design criteria. See Iman and Conover (1982), Tang (1993), Owen (1994), Morris and Mitchell (1995), Tang (1998), Ye (1998), Ye, Li, and Sudjianto (2000), Jin, Chen, and Sudjianto (2005), and Joseph and Hung (2008). We need to extend those optimal design criteria for experiments with branching and nested factors. Take for example the case of maximin LHD proposed by Morris and Mitchell (1995). Here the optimal design criterion is to maximize the inter-site distance among the experimental points (runs). Now with the branching and nested factors, the notion of "distance" does not exist for all factors. Branching factors are qualitative and thus they can not be measured by distances. Moreover, for nested factors, the notion of "distance" exists only if the corresponding levels of the branching factor are the same. Another major aspect that makes the design of experiment different from the usual designs is the importance of the interaction between branching and nested factors. As noted before, these interactions are usually not negligible. Therefore, if

Table 3: Illustration of the naive strategy

run	z_1	$v_1^{z_1} \cdots v_{m_1}^{z_1}\ x_1 \quad \cdots \quad x_t$
1	1	
$\vdots$	$\vdots$	LHD(n_0, $m_1 + t$)
$\vdots$	1	
$\vdots$	2	
$\vdots$	$\vdots$	LHD(n_0, $m_1 + t$)
$2n_0$	2	

any of the main effects is highly correlated with one of these interactions, then that main effect will be misspecified. Thus, the optimal design criteria should be modified to capture the branching-by-nested interaction effects.

The remaining of the paper is organized as follows. In Section 2, we discuss some general strategies for designing experiments with branching and nested factors and introduce the concept of branching LHD. In Section 3, we discuss three different criteria for finding an optimal branching LHD. The analysis of experiments with branching and nested factors is discussed in Section 4. We illustrate the proposed methods using the hard turning experiment in Section 5. Some concluding remarks and future research directions are given in Section 6.

2. BRANCHING LATIN HYPERCUBE DESIGNS

In general, an LHD with N runs and p factors, denoted by LHD(N, p), can be generated using a random permutation of $\{1, 2, \cdots, N\}$ for each factor. However, as discussed before, this cannot be done if the experiment involves branching and nested factors.

Consider a simple case where the branching factor z_1 has only two levels and m_1 factors are nested within each level of the branching factor. Thus, there are m_1 nested factors ($v_1^{z_1}$, $\cdots$, $v_{m_1}^{z_1}$), where each of the nested factor stands for two different factors depending on the two levels of the branching factor. In addition, there are t more shared quantitative factors. Now we discuss some strategies for constructing LHDs with branching and nested factors.

A naive strategy is to choose an LHD for the nested and shared factors and repeat it

Table 4: Branching Latin hypercube design with one branching factor

run	z_1	$v_1^{z_1} \cdots v_{m_1}^{z_1}$	$x_1 \ \cdots \ x_t$
1	1		
$\vdots$	$\vdots$	LHD(n_1, m_1)	
$\vdots$	1		LHD($2n_1$, t)
$\vdots$	2		
$\vdots$	$\vdots$	LHD(n_1, m_1)	
$2n_1$	2		

under each level of the branching factor. That is, first we choose an LHD(n_0, $m_1 + t$) that can accommodate the nested and shared factors. This will then be repeated for the two levels of the branching factor. The resulting design is shown in Table 3. This is easy to construct. Moreover, we can easily choose optimal LHDs for the nested and shared factors using the existing methods. In addition, because all of the combinations of nested and shared factors are repeated at each level of the branching factor, we can estimate the interactions involving branching factor. However, there are two drawbacks. One is that if we project the design matrix onto one of those shared factors (x_1, $\cdots$, x_t), there are some replications. Hence, the design points are not spread out as uniformly as they could be. The other drawback is that, the run size of these designs can be quite large.

The foregoing problems with the naive strategy can be easily overcome by using one LHD for all of the shared factors. As a result, the design points are spread out more uniformly in the experimental region and the run size required is comparably smaller. As an example, for $m_1 = 3$ and $t = 5$, the run size of the naive approach ($2n_0$ in Table 3) should be at least 16 (because $n_0 \geq 8$). It can be reduced to 6 ($2n_1 \geq 6$) by the new design illustrated in Table 4. Following the terminology used by Phadke (1989), we name a design with this structure a *branching Latin hypercube design* (BLHD).

Now consider a more general case with q branching factors denoted by $\boldsymbol{z} = (z_1, \cdots, z_q)$. Assume that all of them are qualitative by nature. For each branching factor z_u, there are k_u levels and under each of these different levels, there are m_u nested factors. Note that in general, the number of factors nested under each level of a branching factor can be different. For example, in the hard turning experiment there are two factors nested under chamfer but none under hone. However, for notational simplicity, we assume the number of nested factors to be the same (for a given branching factor) and develop the construction of BLHD. Later we explain how it can be extended to deal with unequal number of nested factors.

Denote the nested factors by $\boldsymbol{v}^{z_u} = (v_1^{z_u}, \cdots, v_{m_u}^{z_u})'$, $1 \leq u \leq q$. Again note that, each nested factor corresponds to different factors depending on the branching factor and its level. That is why we use a superscript to denote the branching factor level. In addition to the branching and nested factors, there are t shared quantitative factors $\boldsymbol{x} = (x_1, \cdots, x_t)'$. Let $\boldsymbol{v} = ((\boldsymbol{v}^{z_1})', \cdots, (\boldsymbol{v}^{z_q})')'$, and $\boldsymbol{w} = (\boldsymbol{x}', \boldsymbol{z}', \boldsymbol{v}')'$ represents all of the p factors involved in the experiment, where $p = t + q + \sum_{u=1}^{q} m_u$. A N-run BLHD can then be represented by $\boldsymbol{W} = (\boldsymbol{w}_1, \cdots, \boldsymbol{w}_N)'$. In general, it consists of three parts. The first part is a design for branching factors. Because branching factors are qualitative factors, we can choose an orthogonal array of appropriate size depending on the number of levels of each branching factor. The second part consists of LHDs for the nested factors. Choose LHD(n_u, m_u) for the m_u nested factors under the branching factor z_u, $u = 1, 2, \cdots, q$. The third part is a LHD(N, t) for all of the shared factors $\boldsymbol{x}$. If there are k_u levels for branching factor z_u, where $1 \leq u \leq q$, it is clear that $N = k_1 n_1 = k_2 n_2 = \cdots = k_q n_q$. Thus, we have one orthogonal array for the branching factors, q LHDs for the nested factors, and one LHD for the shared factors. These designs can be assembled to obtain a BLHD.

As an example, consider the case with two branching factors (z_1 and z_2) each at two levels. There are m_1 nested factors under z_1 and m_2 nested factors under z_2. Furthermore, there are t shared factors. Table 5 illustrates a N-run BLHD for this example. The first part is a 4-run orthogonal array for those two branching factors. For the second part, we choose LHD(n_1, m_1) for the nested factors under z_1. Similarly, LHD(n_2, m_2) is chosen for the nested factors under z_2. This LHD is divided into two halves and distributed among the two levels of z_2 as shown in the Table. The third part consists of a LHD(N, t) for the t shared factors.

As in the case with LHDs, not all BLHDs are good. We need to use some optimal design criteria to choose the best BLHD. This is discussed in the next section.

Table 5: Branching Latin hypercube design with two branching factors

run	z_1	$v_1^{z_1} \cdots v_{m_1}^{z_1}$	z_2	$v_1^{z_2} \cdots v_{m_2}^{z_2}$	$x_1 \cdots x_t$
1	1	LHD(n_1, m_1)	1	LHD(n_2, m_2) first half	
⋮	1		2	LHD(n_2, m_2) first half	LHD(N, t)
⋮	2	LHD(n_1, m_1)	1	LHD(n_2, m_2) second half	
N	2		2	LHD(n_2, m_2) second half	

3. OPTIMAL BRANCHING LATIN HYPERCUBE DESIGNS

As discussed in the introduction, there are several approaches for finding a good LHD. We may think it is enough to use one of those approaches to generate $q+1$ optimal LHDs for the nested and shared factors and assemble them to obtain the BLHD. However, such an assembly of optimal LHDs may not lead to an optimal BLHD. Moreover, we need to make sure that in a BLHD, the correlation between the branching-by-nested interaction and any other main effect is small. In this section, we propose three optimal design criteria for finding good BLHDs.

3.1 Maximin BLHD

Morris and Mitchell (1995) proposed to find LHDs that maximize the minimum inter-site distance. Let $\boldsymbol{g}$ and $\boldsymbol{h}$ be two design points (or sites or runs). Consider the distance measure $d(\boldsymbol{g},\boldsymbol{h}) = \{\sum_{j=1}^{p} |g_j - h_j|^{\varsigma}\}^{1/\varsigma}$, in which $\varsigma = 1$ and $\varsigma = 2$ correspond to the rectangular and Euclidean distance, respectively. For simplicity, rectangular distance ($\varsigma = 1$) is considered for the rest of the paper. For a given LHD, define a distance list $(D_1, D_2, \cdots, D_M)$ in which the elements are the distinct values of inter-site distances, sorted from the smallest to the largest. Let J_i be the number of pairs of design points in the design separated by D_i. Then a design is called a maximin design if it sequentially maximizes D_i's and minimizes J_i's in the following order: $D_1, J_1, D_2, J_2, \cdots, D_M, J_M$. A scalar-valued function which can be used to rank competing designs in such a way that the maximin design receives the highest ranking is given by

$$\phi_\lambda = (\sum_{i=1}^{M} J_i D_i^{-\lambda})^{1/\lambda} = \left(\sum_{\boldsymbol{g}\neq\boldsymbol{h}} d(\boldsymbol{g},\boldsymbol{h})^{-\lambda}\right)^{1/\lambda}, \qquad (1)$$

where λ is a positive integer.

Extension of the maximin criterion to BLHDs is not straightforward. Different from LHDs where all factors can be measured by distances, BLHDs have branching factors which have no notion of distance and nested factors where definition of distance depend on the corresponding branching factors. Because of different roles of factors, instead of calculating all the pairwise distances over all factors, we need to consider branching and nested factors separately from those of shared factors. First note that for BLHDs, not all factors are divided into same number of levels as that in LHDs: there are k_u levels for branching factor z_u, n_u levels for nested factors $\boldsymbol{v}^{z_u}$, and N levels for $\boldsymbol{x}$. Therefore, before calculating the distances, the design matrix should be scaled to $(-1, 1)$.

Start with a simple case where $q = 1$ and $m_1 = 1$. Thus, there are $t + 2$ factors in the experiment. Assume that the last two factors are branching factor z_1 and nested factor $v_1^{z_1}$, respectively. We need to define two types of inter-site distances. The first type of distance focuses on all of the shared factors. It is the distance projection onto the t-dimensional space ($\boldsymbol{x}$), which can be defined by $d_x(\boldsymbol{g}, \boldsymbol{h}) = \sum_{j=1}^{t} |g_j - h_j|$, where $\boldsymbol{g} = (g_1, \cdots g_{t+2})$ and $\boldsymbol{h} = (h_1, \cdots, h_{t+2})$ are $(t + 2)$-dimensional design points. There are a total of $\binom{N}{2}$ distances. The second type of distance takes into account of the branching and nested factors by considering distances within each level of branching factors. The objective here is to spread out the design points for each level of branching factors. To do so, for each level $z_{1,i}$ of the branching factor, where $1 \leq i \leq k_1$, distances are calculated based on $\boldsymbol{x}$ and $v_1^{z_1}$. Define the second type of distance by $d_B(\boldsymbol{g}, \boldsymbol{h}) = \sum_{l=1}^{t} |g_l - h_l| + |g_{t+2} - h_{t+2}|$. One can easily obtain $d_B(\boldsymbol{g}, \boldsymbol{h}) = d_{v_1}(\boldsymbol{g}, \boldsymbol{h}) + d_x(\boldsymbol{g}, \boldsymbol{h})$, where $d_{v_1}(\boldsymbol{g}, \boldsymbol{h}) = |g_{t+2} - h_{t+2}|$. The second type of inter-site distances are calculated only for those within the same level of the branching factor ($g_{t+1} = h_{t+1} = z_{1,i}$, for some $1 \leq i \leq k_1$) and thus there are $\binom{n_1}{2} k_1$ of them.

Note that, the first type of distance measure consists of t dimensions, while the second consists of $t+1$ dimensions. After standardizing with respect to their dimensions, the maximin distance criterion can be extended to BLHDs by defining a distance list based on these $\binom{N}{2} + \binom{n_1}{2} k_1$ standardized inter-site distances. Furthermore, as in (1), the scalar-valued function can

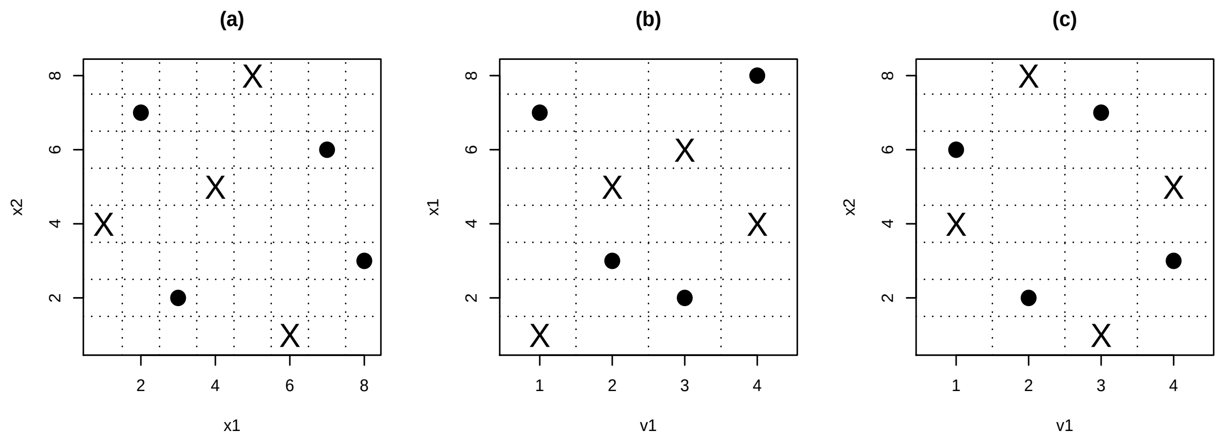

Figure 4: Maximin BLHD. "X" stands for $z_1 = 1$ and solid points stands for $z_1 = 2$.

be defined as

$$\phi_\lambda = \left(\sum_{\boldsymbol{g} \neq \boldsymbol{h}} \left[\frac{t}{d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda + \sum_{i=1}^{k_1} \sum_{g_{t+1}=h_{t+1}=z_{1,i}} \left[\frac{1+t}{d_{v_1}(\boldsymbol{g}, \boldsymbol{h}) + d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda \right)^{1/\lambda}, \tag{2}$$

where $\sum_{g_{t+1}=h_{t+1}=z_{1,i}} d_{v_1}(\boldsymbol{g}, \boldsymbol{h})$ is the sum of $\binom{n_1}{2}$ pairwise distances in which $\boldsymbol{g}$ and $\boldsymbol{h}$ have the same level of branching factor and $\sum_{\boldsymbol{g} \neq \boldsymbol{h}} d_x(\boldsymbol{g}, \boldsymbol{h}) = \sum_{i=1}^{k_1} \sum_{g_{t+1}=h_{t+1}=z_{1,i}} d_x(\boldsymbol{g}, \boldsymbol{h}) + \sum_{g_{t+1} \neq h_{t+1}} d_x(\boldsymbol{g}, \boldsymbol{h})$.

To illustrate this idea, consider the simple example in Table 2. Assume that the branching factor z_1 is qualitative. The optimal design found by the modified maximin criterion (2) (with $\lambda = 15$) is $x_1 = \{1, 5, 6, 4, 7, 3, 2, 8\}$, $x_2 = \{4, 8, 1, 5, 6, 2, 7, 3\}$, and z_1 and v_1 remain the same as in Table 2. This maximin BLHD is plotted in Figure 4, where "X" represent those design points with $z_1 = 1$ and solid points stands for those with $z_1 = 2$. The first part in (2) tries to maximize the inter-site distances in the space of x_1 and x_2 (Figure 4(a)), in which the "X" and solid points are not distinguished. Whereas the second part in (2) tries to maximize the inter-site distances in the space of x_1, x_2, and v_1 (Figures 4(b) and (c)). Moreover, because these distances are calculated only within the same level of the branching factor, the inter-site distances among the "X" points and among the solid points are maximized.

If we were to use only the first part in (2) as the criterion, then the optimal design would be space-filling only over the shared factors. The design points can be quite structured with respect to the branching and nested factors. This can be seen in Figure 5. It is clear

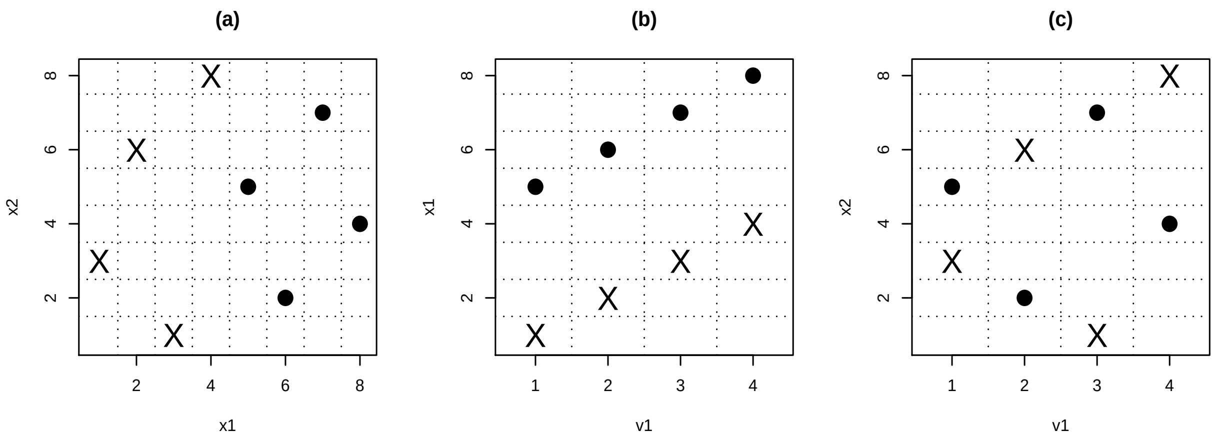

Figure 5: Maximin for shared factors

that although the design points are evenly spread out in the (x_1, x_2) space (Figure 5(a)), experiments for $z_1 = 1$ concentrate on lower level of x_1 and experiments for $z_1 = 2$ concentrate on higher level of x_1. Furthermore, the nested factor v_1 is highly correlated with x_1. This clearly shows the importance of the new criterion in (2). We should also point out that the two designs for the shared factors (Figure 4(a) and Figure 5(a)) are isomorphic. However, these two isomorphic designs are not equally good when we include the branching factor. The new criterion could clearly identify that Figure 4(a) gives a better design than Figure 5(a).

Now consider the general case with q branching factors $\boldsymbol{z}$ and the corresponding nested factors $\boldsymbol{v}$. Assume that for each branching factor z_u, there are k_u levels denoted by $z_{u,i}$, $1 \leq i \leq k_u$. Let $\boldsymbol{g}$ and $\boldsymbol{h}$ are design points with p dimensions. Let $\delta_u = (t + \sum_{l=1}^{u}(m_{l-1}+1)+1)$, such that the δ_uth factor is the uth branching factor, and the $[\delta_u + 1]$th to $[\delta_u + m_u]$th factors are the corresponding nested factors. As an extension of (2), the maximin criterion for BLHDs can be written as

$$\phi_\lambda = \left(\sum_{\boldsymbol{g} \neq \boldsymbol{h}} \left[\frac{t}{d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda + \sum_{u=1}^{q} \sum_{i=1}^{k_u} \sum_{g_{\delta_u} = h_{\delta_u} = z_{u,i}} \left[\frac{m_u + t}{d_{v_u}(\boldsymbol{g}, \boldsymbol{h}) + d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda \right)^{1/\lambda}, \tag{3}$$

where $d_x(\boldsymbol{g}, \boldsymbol{h}) = \sum_{j=1}^{t} |g_j - h_j|$ and the distance measure for the u-th branching factor is denoted by $d_{v_u}(\boldsymbol{g}, \boldsymbol{h}) = \sum_{l=\delta_u+1}^{\delta_u+m_u} |g_l - h_l|$. This criterion is general and includes some interesting special cases.

Case 1: If $q = 0$, and $m_u = 0$ for all u, then this would lead to the standard LHD(N, t). Up to a constant, the maximin criterion (3) would be the same as (1) proposed by Morris and Mitchell (1995).

Case 2: If there is no nested factor corresponding to the branching factors, that is $m_u = 0$, for all u, one can think about this as an experiment with q qualitative factors $\boldsymbol{z}$ and t quantitative factors $\boldsymbol{x}$. As a special case of (3), the maximin criterion for experimental design with quantitative and qualitative factors can be written as

$$\phi_\lambda = \left(\sum_{\boldsymbol{g} \neq \boldsymbol{h}} \left[\frac{t}{d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda + \sum_{u=1}^{q} \sum_{i=1}^{k_u} \sum_{g_{\delta_u} = h_{\delta_u} = z_{u,i}} \left[\frac{t}{d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda \right)^{1/\lambda} \tag{4}$$

Case 3: Another special case is that when $t = 0$. In this situation, experiments consist branching factors and nested factors but no shared factors. Because $d_x(\boldsymbol{g}, \boldsymbol{h}) = 0$, for all $\boldsymbol{g}$ and $\boldsymbol{h}$, (3) can be simplified to

$$\phi_\lambda = \left(\sum_{u=1}^{q} \sum_{i=1}^{k_u} \sum_{g_{\delta_u} = h_{\delta_u} = z_{u,i}} \left[\frac{m_u}{d_v(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda \right)^{1/\lambda}.$$

3.2 Minimum Correlation BLHD

Apart from space-filling, another important issue regarding experimental designs is how to construct them such that the significant factors can be correctly identified. For LHDs, it can be achieved by minimizing the pairwise correlation among factors (Iman and Conover, 1982; Owen, 1994; Tang, 1998). Owen (1994) proposed a performance measure ρ^2 for evaluating the goodness of an LHD with respect to pairwise correlations. For LHD(N, t),

$$\rho^2 = \frac{\sum_{i=2}^{t} \sum_{j=1}^{i-1} \rho_{ij}^2}{t(t-1)/2}, \tag{5}$$

where ρ_{ij} is the linear correlation between columns i and j.

Different from LHDs, orthogonality among main effects is not enough in BLHDs. In BLHDs, it is equally important to consider the branching-by-nested interactions. Therefore, we propose a modified correlation criterion, which minimizes the correlations among the main effects of all factors as well as those between main effect of a shared factor and a branching-by-nested interaction effect. To do so, we first enlarge the BLHDs by including two-factor

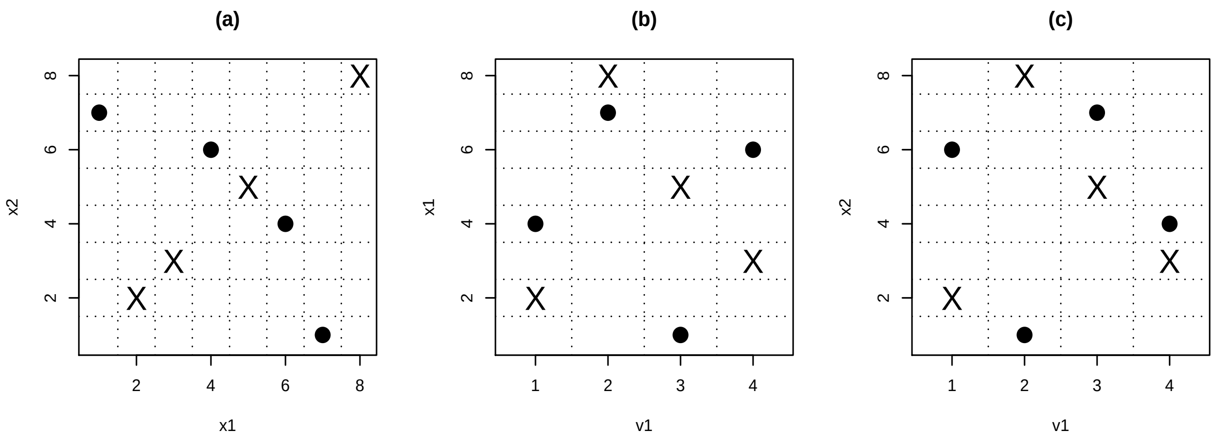

Figure 6: Minimum correlation BLHD

interactions which represent the branching-by-nested interaction. There are m_u such interactions for each branching factor z_u, where m_u is the number of nested factors. Therefore, the total number of branching-by-nested interactions would be s, where $s = \sum_{u=1}^{q} m_u$. Thus, the new criterion for BLHDs is given by

$$\rho^2 = \frac{\sum_{i=2}^{p} \sum_{j=1}^{i-1} \rho_{ij}^2 + \sum_{i=1}^{t} \sum_{j=1}^{s} \tilde{\rho}_{ij}^2}{(p(p-1)/2) + st}, \tag{6}$$

where ρ_{ij}^2 is the linear correlation between columns i and j in the design and $\tilde{\rho}_{ij}^2$ is the linear correlation between x_i and jth branching-by-nested interaction.

Consider the same example used in Table 2. The optimal design that minimizes ρ^2 in (6) is $x_1 = (2, 8, 5, 3, 4, 7, 1, 6)$, $x_2 = (2, 8, 5, 3, 6, 1, 7, 4)$, and z_1 and v_1 remain the same as in Table 2. The design is plotted in Figure 6.

3.3 Orthogonal-Maximin BLHD

Maximizing minimum inter-site distances does not ensure minimizing pairwise correlations and vice-versa. Therefore, Joseph and Hung (2008) proposed a multi-objective criterion for LHD that combines the maximin distance and the minimum correlation criteria. This criterion becomes even more important in the case of BLHDs, because it is important to ensure small correlations between the shared factors and the branching-by-nested interactions besides enuring good space-filling properties. To extend the result in Joseph and Hung (2008) to BLHD, we should scale ϕ_λ and ρ^2 to the same range so that some meaningful weights can be assigned in the multi-objective function. The following result gives the lower and upper bounds for ϕ_λ, which can be used for scaling it to $[0, 1]$. Here we only consider the case of a single branching factor. The result can be extended to include more than one branching factor, but the expressions become more complicated.

PROPOSITION 1. *If there is only one branching factor, then*

$$\phi_{\lambda,L} \leq \phi_\lambda \leq \phi_{\lambda,U},$$

where

$$\phi_{\lambda,L} = 3\left[\sum_{i=1}^{k_1} \sum_{g_{\delta_1}=h_{\delta_1}=z_{1,i}} 2^{\frac{1}{p+1}}(m_1+t)^{\frac{p}{p+1}} + \sum_{g_{\delta_1}\neq h_{\delta_1}} t^{\frac{p}{p+1}}\right]^{\frac{(\lambda+1)}{\lambda}} \left[t(N^2-1) + \sum_{i=1}^{k_1} m_1(n_1^2-1)\right]^{-1},$$

and

$$\phi_{\lambda,U} = \frac{N}{2}\left[\sum_{i=1}^{k_1}\sum_{j=1}^{n_1-1} \frac{(n_1-j)(t+m_1)^\lambda}{j^\lambda(t+k_1m_1)^\lambda} + \sum_{j=1}^{N-1}\frac{N-j}{j^\lambda}\right]^{1/\lambda}.$$

Thus, the multi-objective criterion is to minimize

$$\psi_\lambda = w\rho^2 + (1-w)\frac{\phi_\lambda - \phi_{\lambda,L}}{\phi_{\lambda,U} - \phi_{\lambda,L}}. \tag{7}$$

We usually take $w = .5$ and call the design that minimizes this criterion as orthogonal-maximin BLHD.

The design matrix in Table 2 is an orthogonal-maximin BLHD. It is plotted in Figure 7. Comparisons of the optimal designs found by the forgoing three criteria are provided in Table 6. It can be seen that the correlation between the shared factor and the branching-by-nested interaction (denoted by INT) is 0 in the case of minimum correlation design. However, the points are much closer compared to the maximin BLHD. The orthogonal-maximin BLHD provides a good compromise between them.

Because of the combinatorial nature of the optimization problem, finding the optimal BLHD for large dimensions is a difficult task. Several methods such as simulated annealing (Morris and Mitchell 1995), columnwise-pairwise algorithm (Ye, Li, and Sudjianto 2000),

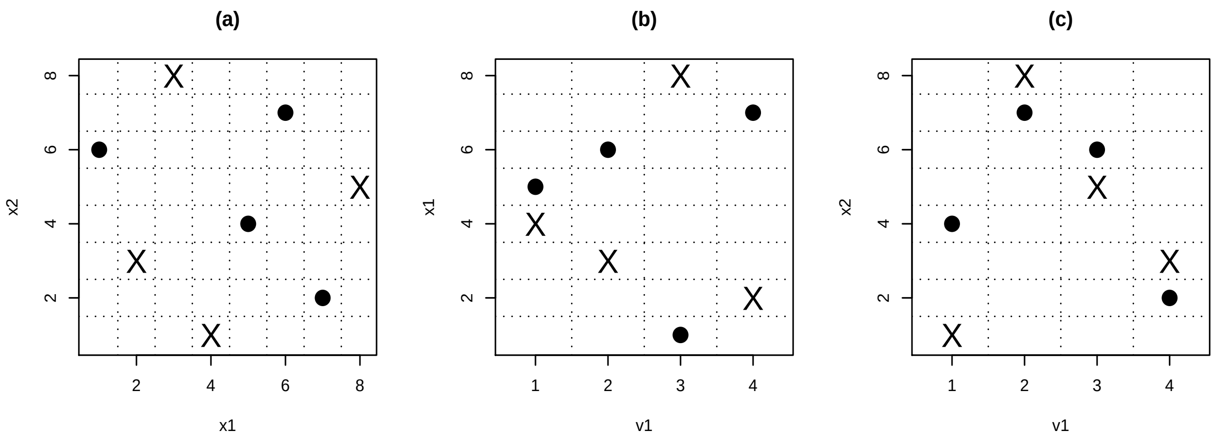

Figure 7: Orthogonal-maximin BLHD

Table 6: Comparison of different BLHDs

	Maximin Distance	Minimum Correlation	Orthogonal-Maximin
ϕ_λ	2.36	4.35	2.37
$D_1(J_1)$	$\frac{1}{2}$ (12)	$\frac{1}{4}$ (3)	$\frac{1}{2}$ (13)
ρ^2	0.0287415	0.000283	0.01445
(cor(x_1, INT), cor(x_2, INT))	(0.195,0)	(0,0)	(-0.098,0)

enhanced stochastic evolutionary algorithm (Jin, Chen, and Sudjianto 2005), and the modified simulated annealing (Joseph and Hung 2008) are proposed in the literature for finding the optimal LHD. These methods can be easily adapted for finding the optimal BLHD as well. A C++ code based on Joseph and Hung's algorithm is available from the authors upon request.

4. KRIGING WITH BRANCHING AND NESTED FACTORS

In this section we explain how branching and nested factors can be incorporated in kriging. Although similar extensions can be made on other methods such as linear regression, we focus here on kriging because of its popularity in the analysis of computer experiments (Sacks et al. 1989). The ordinary kriging model is given by $Y(\boldsymbol{w}) = \mu + Z(\boldsymbol{w})$, where $Z(\boldsymbol{w})$ is a weak stationary stochastic process with mean 0 and covariance function $\sigma^2\psi$ and $\boldsymbol{w}$

$\in \mathbb{R}^p$. The correlation function is defined as $\text{cor}\{Y(\boldsymbol{w}_1), Y(\boldsymbol{w}_2)\} = \psi(\boldsymbol{w}_1, \boldsymbol{w}_2)$. Usually, a product correlation structure is assumed for the correlation function. Consider the example in Table 2. The correlation function between two points $\boldsymbol{w}_1 = (x_{11}, x_{12}, z_{11}, v_{11}^{z_{11}})$ and $\boldsymbol{w}_2 = (x_{21}, x_{22}, z_{21}, v_{21}^{z_{21}})$ can be described as a product of correlation functions of each factor (ψ_i). In the usual cases, a common correlation function is chosen for each factor. However, this cannot be done in the present problem because of the different types of factors.

For the shared factors, a Gaussian correlation function may be used (Santner et al. 2003):

$$\psi_i(x_{1i}, x_{2i}) = \exp\{-\alpha_i(x_{1i} - x_{2i})^2\},$$

whereas for a branching factor, an isotropic correlation function may be used (Joseph and Delaney 2007):

$$\xi_1(z_{11}, z_{21}) = \exp\left\{ -\theta_1 I_{[z_{11} \neq z_{21}]} \right\}.$$

Here α_i and θ_1 are correlation parameters and I_A is an indicator function which takes value 1 when A is true and 0 otherwise.

A new correlation function needs to be developed for nested factors. Assume that they are quantitative factors. We cannot use the Gaussian correlation function here because a nested factor represents different factors depending on the level of the branching factor. Therefore, it is not reasonable to use one correlation parameter for a given nested factor. Instead, they should be different depending on the level of branching factor. For the example in Table 2, the correlation function for $v_1^{z_1}$ can be defined as following. If two points have the same level in the branching factor, for example $z_{11} = z_{21} =$ "1", then the correlation function will be $\exp\{-\gamma_1(v_{11}^{z_{11}} - v_{21}^{z_{21}})^2\}$. Similarly, if $z_{11} = z_{21}=$"2", the correlation function will be $\exp\{-\gamma_2(v_{11}^{z_{11}} - v_{21}^{z_{21}})^2\}$. If the two points do not have the same level in the branching factor (i.e., $z_{11} \neq z_{21}$), then correlation should be determined by the branching factor not the nested factor. Hence, in this case, the correlation function for nested factor will be equal to 1. Succinctly, the correlation function for the nested factor can be defined as

$$\varpi_1(v_{11}^{z_{11}}, v_{21}^{z_{21}}) = \exp\left\{ -\sum_{j=1}^{2} \gamma_j (v_{11}^{z_{11}} - v_{21}^{z_{21}})^2 I_{[z_{11}=z_{21}=j]} \right\}. \tag{8}$$

We can easily extend this to a more general situation. Let there are q branching factors $z_1, \cdots, z_q$. For each branching factor z_u, there are m_u nested factors. Assume $\boldsymbol{x}_1 = (x_{11}, \cdots x_{1t})'$, $\boldsymbol{z}_1 = (z_{11}, \cdots, z_{1q})'$, $\boldsymbol{v}_1^{z_u} = (v_{11}^{z_u}, \cdots, v_{1\,m_u}^{z_u})'$, and $\boldsymbol{v}_1 = ((\boldsymbol{v}_1^{z_1})', \cdots, (\boldsymbol{v}_1^{z_q})')$. Similarly, $\boldsymbol{x}_2 = (x_{21}, \cdots x_{2t})'$, $\boldsymbol{z}_2 = (z_{21}, \cdots, z_{2q})'$, $\boldsymbol{v}_2^{z_u} = (v_{21}^{z_u}, \cdots, v_{2m_u}^{z_u})'$, and $\boldsymbol{v}_2 = ((\boldsymbol{v}_2^{z_1})', \cdots, (\boldsymbol{v}_2^{z_q})')$. Given any two design points $\boldsymbol{w}_1 = (\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{v}_1)$ and $\boldsymbol{w}_2 = (\boldsymbol{x}_2, \boldsymbol{z}_2, \boldsymbol{v}_2)$, the correlation function can be written as

$$\mathrm{cor}(Y(\boldsymbol{w}_1), Y(\boldsymbol{w}_2)) = \left(\prod_{i=1}^{t} \psi_i(\boldsymbol{x}_1, \boldsymbol{x}_2)\right) \prod_{u=1}^{q} \left[\xi_u(\boldsymbol{z}_1, \boldsymbol{z}_2)\left(\prod_{j=1}^{m_u} \varpi_u(v_{1j}^{z_{1u}}, v_{2j}^{z_{2u}})\right)\right], \tag{9}$$

where $\psi_i(\boldsymbol{x}_1, \boldsymbol{x}_2) = \exp\{-\alpha_i(x_{1i} - x_{2i})^2\}$ is the correlation function for the shared factors, $\xi_u(\boldsymbol{z}_1, \boldsymbol{z}_2) = \exp\Big\{ - \theta_u I_{[z_{1u} \neq z_{2u}]}\Big\}$ is the correlation function for the branching factors, and

$$\varpi_u(v_{1i}^{z_{1u}}, v_{2i}^{z_{2u}}) = \exp\left\{ - \sum_{j=1}^{k_u} \gamma_{uij}(v_{1i}^{z_{1u}} - v_{2i}^{z_{2u}})^2 I_{[z_{1u} = z_{2u} = z_{u,j}]}\right\}, \tag{10}$$

is the correlation function for the nested factors. Note that $z_{u,j}$, $1 \leq j \leq k_u$ are the k_u levels for each branching factor z_u. Thus, we obtain

$$\begin{aligned} \mathrm{cor}(Y(\boldsymbol{w}_1), Y(\boldsymbol{w}_2)) = \; & \exp\Big\{ - \textstyle\sum_{i=1}^{t} \alpha_i(x_{1i} - x_{2i})^2 \\ & - \textstyle\sum_{u=1}^{q} \Big[\theta_u I_{[z_{1u} \neq z_{2u}]} + \sum_{i=1}^{m_u} \sum_{j=1}^{k_u} \gamma_{uij}(v_{1i}^{z_{1u}} - v_{2i}^{z_{2u}})^2 I_{[z_{1u} = z_{2u} = z_{u,j}]}\Big]\Big\}. \end{aligned} \tag{11}$$

Denote the correlation parameters by $\boldsymbol{\Theta} = (\boldsymbol{\alpha}', \boldsymbol{\theta}', \boldsymbol{\gamma}')$, where $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_t)'$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_q)'$, and $\boldsymbol{\gamma} = (\gamma_{111}, \cdots \gamma_{qm_qk_q})'$. We can estimate these parameters from data and obtain the ordinary kriging predictor. This will be explained with an example in the next section.

5. HARD TURNING EXPERIMENT

The objective of our experiment is to optimize a hard turning process with respect to cutting forces. Hard turning is a metal cutting process that produces machined parts out of hard materials with good dimensional accuracy, surface finish, and surface integrity. Minimizing cutting forces will help reduce the power requirements, elastic distortion of the workpiece, and tool wear; thus reducing the manufacturing cost and improving the quality of the machined part.

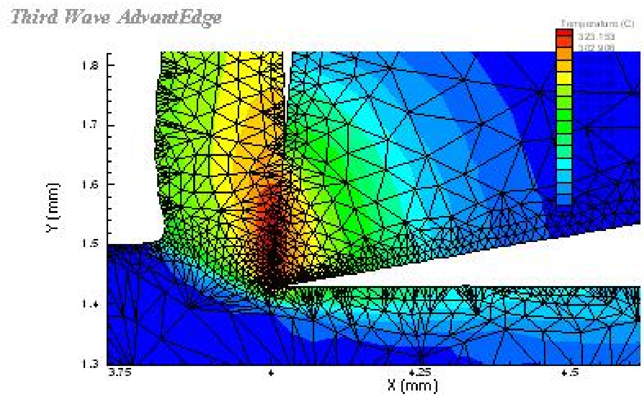

Figure 8: Finite element mesh and temperature distribution

Nine factors are selected for experimentation, which include one branching factor, two nested factors, and six shared factors. The factors and their ranges are shown in Table 1. A 30-run orthogonal-maximin BLHD is generated by using the modified simulated annealing algorithm proposed by Joseph and Hung (2008). The optimal design matrix is given in Table 7. The branching factor (cutting edge shape, z_1), is labeled "1" and "2" to stand for chamfer and hone, respectively. Two nested factors (v_1 and v_2) are nested within the cutting edge shape. Recall that, if the cutting edge is chamfer, v_1 stands for chamfer angle and v_2 stands for chamfer land length, otherwise there is no factor.

The experiments are performed using a highly sophisticated finite element based machining simulation software *AdvantEdge*. This software models the underlying physics of metal cutting as a thermo-mechanical plastic deformation process and captures various material and geometric nonlinearities of the process. The theoretical basis of the simulation model can be found in Marusich and Ortiz (1995). The simulations are computationally intensive and require hours of running time for producing a single output (about 12 to 24 hours). The simulation outputs are deterministic and incorporate all the factors in Table 1. Various responses are produced by the software such as temperature, residual stresses, and forces. A finite element mesh and temperature distribution is shown in Figure 8. In this article, we chose to analyze only the resultant cutting force (y). The data are given in Table 7.

This is an example where the number of nested factors are not the same for different levels of the branching factor. For simplicity of notations, we had assumed them to be the same in Section 3. Therefore, we should explain, how the criteria can be modified to deal with unequal number of nested factors. This is easy to do. Consider first the maximin criterion ϕ_λ. In (3), use $m_1 = 2$ for the first 15 runs and use $m_1 = 0$ for the last 15 runs. Now consider ρ^2 in (6). The pairwise correlations involving a nested factor should be calculated using the first 15 runs. Moreover, because there are no factors when branching factor level is "2", we do not need to consider the branching-by-nested interaction.

Since the cutting forces are positive, we first apply a log transformation before fitting the ordinary kriging model. We also normalize all of the factor settings in Table 7 into $[-1, 1]$. The 30 design points after normalization are denoted by $\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_{30}\}$, where for all $1 \leq j \leq 30$, $\boldsymbol{w}_j = (x_{j1}, \cdots, x_{j6}, z_{j1}, v_{j1}^{z_{j1}}, v_{j2}^{z_{j1}})'$, $z_{j1} = -1$ represents the chamfer edge and $z_{j1} = 1$ represents the hone edge. The parameters in the kriging model can be estimated as (Santner, William, and Notz, 2003)

$$
\begin{aligned}
\hat{\boldsymbol{\Theta}} &= arg \min_{\boldsymbol{\Theta}} N \log \hat{\sigma^2} + \log |\boldsymbol{\Psi}|, \\
\hat{\mu} &= (\mathbf{1}'\boldsymbol{\Psi}^{-1}\mathbf{1})^{-1}\mathbf{1}'\boldsymbol{\Psi}^{-1}\boldsymbol{y}, \\
\hat{\sigma}^2 &= \frac{1}{N}(\boldsymbol{y} - \hat{\mu}\mathbf{1})'\boldsymbol{\Psi}^{-1}(\boldsymbol{y} - \hat{\mu}\mathbf{1}),
\end{aligned}
$$

where $\mathbf{1}$ is a vector of 1's having length 30, $\boldsymbol{y} = (y_1, \cdots, y_{30})'$, and $\boldsymbol{\Psi}$ is a 30×30 matrix whose njth element is

$$
\begin{aligned}
\exp\Big\{ & - \textstyle\sum_{i=1}^{6} \hat{\alpha}_i (x_{ni} - x_{ji})^2 - \hat{\theta}_1 \mathbf{I}_{[z_{n1} \neq z_{j1}]} \\
& - \hat{\gamma}_{111}(v_{n1}^{z_{n1}} - v_{j1}^{z_{j1}})\mathbf{I}_{[z_{n1} = z_{j1} = -1]} - \hat{\gamma}_{121}(v_{n2}^{z_{n1}} - v_{j2}^{z_{j1}})\mathbf{I}_{[z_{n1} = z_{j1} = -1]} \Big\}.
\end{aligned}
$$

We obtain

$$\hat{\alpha} = (0.09, 0.01, 0.03, 0.01, 0.94, 1.08)', \ \hat{\theta} = \hat{\theta}_1 = 0.13, \ \hat{\gamma} = (\hat{\gamma}_{111}, \hat{\gamma}_{121})' = (0.14, 0.01)', \hat{\mu} = 5.1.$$

Note that we do not need to estimate γ_{112} and γ_{122} in this example, because there is no factor nested within hone edge. Thus, ordinary kriging predictor is given by (see, e.g., Joseph 2006)

$$\hat{y}(\boldsymbol{w}) = 5.1 + \hat{\boldsymbol{\psi}}(\boldsymbol{w})'\hat{\boldsymbol{\Psi}}^{-1}(\boldsymbol{y} - 5.1\mathbf{1}), \tag{12}$$

Table 7: Orthogonal-maximin BLHD and data for the hard turning experiment

Run	z_1	v_1	v_2	x_1	x_2	x_3	x_4	x_5	x_6	y
1	1	1	6	15	23	7	9	18	10	162.1
2	1	2	11	25	3	25	14	25	19	284.9
3	1	3	3	4	20	18	18	5	26	160.3
4	1	4	14	9	6	6	27	7	17	121.1
5	1	5	8	16	8	21	2	2	1	104.6
6	1	6	1	17	10	5	25	19	25	241.9
7	1	7	12	29	26	15	5	14	12	195.4
8	1	8	5	26	16	30	22	15	6	159.5
9	1	9	15	7	13	26	7	11	27	241.6
10	1	10	10	1	29	20	23	6	5	88.33
11	1	11	2	20	21	27	10	20	29	320.4
12	1	12	7	8	11	14	4	29	21	218.8
13	1	13	13	22	9	1	24	27	9	193.5
14	1	14	4	10	2	24	28	13	13	198.6
15	1	15	9	28	25	13	17	3	28	155.1
16	2			19	5	9	1	8	20	164.4
17	2			14	28	17	6	21	24	323.6
18	2			6	17	4	16	12	4	109.1
19	2			11	1	12	15	4	8	115.4
20	2			27	22	8	30	24	16	254.8
21	2			21	14	23	19	10	22	217.0
22	2			23	18	22	12	28	3	243.7
23	2			3	27	3	3	26	14	131.5
24	2			13	15	19	29	16	30	258.7
25	2			24	12	2	11	1	18	109.3
26	2			18	24	28	8	17	2	174.8
27	2			12	30	11	26	9	11	157.0
28	2			2	4	16	13	30	15	133.1
29	2			30	7	10	20	23	7	210.1
30	2			5	19	29	21	22	23	273.3

where $\boldsymbol{w} = (x_1, \cdots, x_6, z_1, v_1^{z_1}, v_2^{z_1})' \in [-1,1]^9$ and $\hat{\boldsymbol{\psi}}(\boldsymbol{w})$ is a vector of length 30 with the jth element

$$\exp\left\{ -\sum_{i=1}^{6} \hat{\alpha}_i (x_i - x_{ji})^2 - \hat{\theta}_1 \mathbf{I}_{[z_1 \neq z_{j1}]} - \hat{\gamma}_{111}(v_1^{z_1} - v_{j1}^{z_{j1}})\mathbf{I}_{[z_1 = z_{j1} = -1]} - \hat{\gamma}_{121}(v_2^{z_1} - v_{j2}^{z_{j1}})\mathbf{I}_{[z_1 = z_{j1} = -1]} \right\}.$$

To understand the effects of the factors, we apply the sensitivity analysis technique on the ordinary kriging predictor (see Welch et al. 1992). The main effects plot is shown in Figure 9 (a). We can see that the cutting edge radius (x_1), feed (x_5), depth of cut (x_6), and chamfer angle (v_1) have significant effects on the cutting force. We also found a significant interaction between cutting edge radius and depth of cut (Figure 9 (b)). The depth of cut has a positive effect on the cutting forces, but this effect is more significant when the cutting edge radius is smaller. This can be explained physically as follows. For a small cutting edge radius, an increase in depth of cut produces an increase in material deformation through shear and consequently its effect on the force is more significant. For larger cutting edge radius values, the contribution of ploughing of material around the cutting edge to the cutting force is more pronounced and consequently an increase in depth of cut does not produce as significant a change in the cutting force.

The optimal setting of the factors can be found by minimizing the ordinary kriging predictor in (12). We obtain $(x_1, x_2, x_3, x_4, x_5, x_6, z_1, v_1, v_2) = (-1.00, -0.76, 0.69, 0.70, -0.66, -0.70, -1, 0.05, 0.16)$. In their original scales, the optimal setting for the shared factors is $(x_1, x_2, x_3, x_4, x_5, x_6) = (5, -13.80, 1.41, 222, 0.067, 0.123)$ and the optimal cutting edge geometry is chamfer with angle 18.74 degree and length 128.13 microns. The resultant cutting force predicted under this setting is 81 N, which is much smaller than the observed forces in the experiment. We also performed a new experiment at the optimal setting and obtained the resultant force as 79 N. This confirms the validity of the optimal setting obtained from our model.

6. CONCLUSIONS

Design and analysis of experiments with branching and nested factors have surprisingly received scant attention in the literature. One possible reason for this could be that the

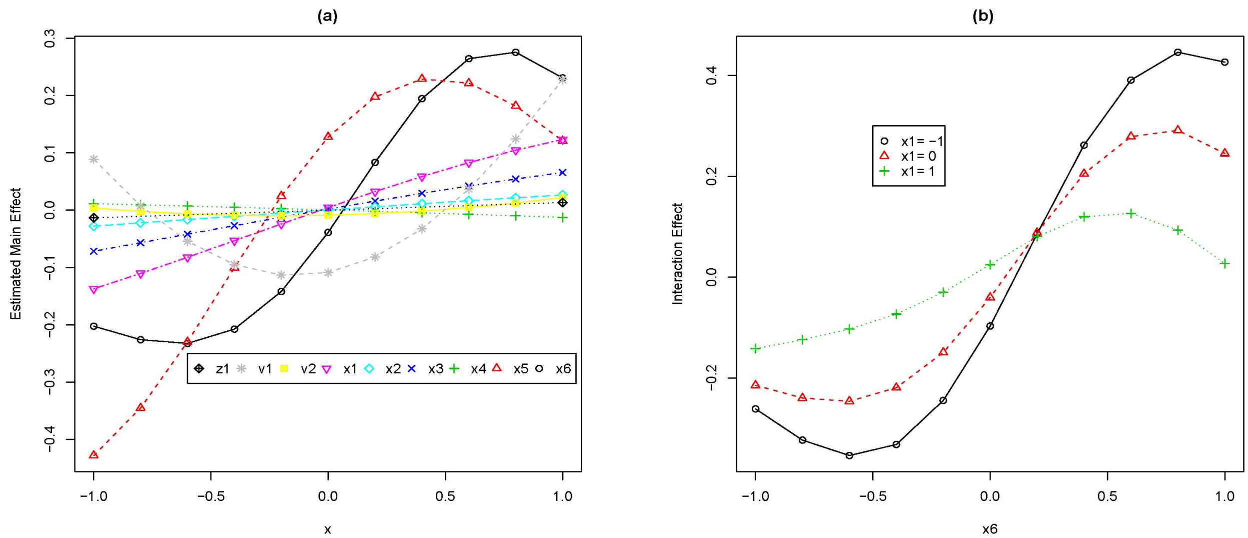

Figure 9: (a) Main effects plot and (b) interaction between x_1 and x_6.

experiments can be performed in two stages, i.e., in the first stage perform an experiment with the branching factors and shared factors. Because there are no nested factors, this experiment can be designed using the existing methods. Now the investigator can analyze the data and find out the best level of the branching factor. Then, a second stage of the experiments can be performed using only the nested factors under the optimum level of the branching factor. The design for this experiment can also be easily obtained using the existing methods. Although this two-stage approach is quite intuitive, the final results may not be optimal. This is because, a different level of the branching factor may be the true optimum, but could not be identified in the first stage of the experiment because the nested factors under that branching level was not set at their optimal levels. This problem can be avoided using branching designs. It allows us to find the optimal settings of the branching factors, nested factors, and shared factors simultaneously.

Taguchi (1987) and Phadke (1989) have reported several case studies on experiments using branching designs, but their approach is not general enough to apply to more complex experiments such as a computer experiment. The optimality properties of their approach using orthogonal arrays are also not known. In this work, we have proposed branching Latin hypercube designs that is suitable for a computer experiment when it involves branching and nested factors. The optimal choice of such designs are also discussed. The approach was successfully applied to the optimization of a machining process.

Although the primary focus was on Latin hypercube designs, some issues regarding the use of orthogonal arrays and its applications in physical experiments are also important. Research on these topics is currently ongoing and will be reported elsewhere.

ACKNOWLEDGMENTS

This research was supported by U. S. National Science Foundation grants CMMI-0654369 and CMMI-0620259. We are grateful to the Editor, an AE and four referees for their helpful comments and suggestions.

APPENDIX: PROOF OF PROPOSITION 1

If there is one branching factor ($q = 1$), a BLHD will include k_1 small LHD(n_1, m_1) for the k_1 levels of the branching factor and a LHD(N, t) for the shared factors. ϕ_λ can be written as

$$\phi_\lambda = \left(\sum_{\boldsymbol{g} \neq \boldsymbol{h}} \left[\frac{t}{d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda + \sum_{i=1}^{k_1} \sum_{g_{\delta_1} = h_{\delta_1} z_{1,i}} \left[\frac{m_1 + t}{d_{v_1}(\boldsymbol{g}, \boldsymbol{h}) + d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda \right)^{1/\lambda}$$

As shown in Joseph and Hung (2008), for a given LHD(N,t), the average inter-site distance (rectangular measure) is $N(N^2 - 1)t/6$, which is a constant. With these constraints, finding a lower bound for ϕ_λ can be formulated as a constraint minimization problem:

$$\begin{aligned} &\min \ \phi_\lambda \\ &\text{subject to} \quad \textstyle\sum_{g_{\delta_1} = h_{\delta_1} = z_{1,i}} d_{v_1}(\boldsymbol{g}, \boldsymbol{h}) = \frac{m_1 n_1 (n_1^2 - 1)}{6}, 1 \leq i \leq k_1 \\ &\qquad\qquad\quad \textstyle\sum_{\boldsymbol{g} \neq \boldsymbol{h}} d_x(\boldsymbol{g}, \boldsymbol{h}) = \frac{tN(N^2 - 1)}{6}, \end{aligned} \tag{13}$$

where $\sum_{g_{\delta_1} = h_{\delta_1} = z_{1,i}} d_{v_1}(\boldsymbol{g}, \boldsymbol{h})$ is the sum of inter-site distances for those smaller LHD(n_1,m_1) and $\sum_{\boldsymbol{g} \neq \boldsymbol{h}} d_x(\boldsymbol{g}, \boldsymbol{h})$ is the sum of inter-site distances for LHD(N,t). Since

$$\phi_\lambda \geq \phi_\lambda^* = \left(\sum_{g_{\delta_1} \neq h_{\delta_1}} \left[\frac{t}{d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda + \sum_{i=1}^{k_1} \sum_{g_{\delta_1} = h_{\delta_1} = z_{1,i}} 2 \left[\frac{m_1 + t}{d_{v_1}(\boldsymbol{g}, \boldsymbol{h}) + d_x(\boldsymbol{g}, \boldsymbol{h})} \right]^\lambda \right)^{1/\lambda},$$

the lower bound can be found by minimizing ϕ_λ^* with the same constraints in (13). Hence, the lower bound can be solved by using the Lagrange multiplier method. For the upper bound of ϕ_λ, result for BLHD is a simple extension of that for LHD. Therefore, they can be proved by the same argument used in Joseph and Hung (2008).

REFERENCES

Fang, K. T., Li, R., and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*, CRC Press, New York.

Hicks, C. R. and Turner K. V. (1999). *Fundamental Concepts in the Design of Experiments*, 5th ed., Oxford: Oxford University Press, New York.

Iman, R. L. and Conover, W. J. (1982), "A Distribution-free Approach to Inducing Rank Correlation Among Input Variables," *Communications in Statistics Part B–Simulation and Computation*, 11, 311-334.

Jin, R., Chen, W. and Sudjianto, A. (2005), "An Efficient Algorithm for Constructing Optimal Design of Computer Experiments," *Journal of Statistical Planning and Inference*, to appear.

Joseph, V. R. (2006), "Limit Kriging," *Technometrics*, 48, 458-466.

Joseph, V. R. and Delaney, J. D. (2007), "Functional Induced Priors for the Analysis of Experiments," *Technometrics*, 49, 1-11.

Joseph, V. R. and Hung, Y. (2008), "Orthogonal-maximin Latin Hypercube Designs," *Statistica Sinica*, to appear. (available at http://www.isye.gatech.edu/statistics/papers).

Marusich, T. D. and Ortiz, M. (1995), "Modeling and Simulation of High-Speed Machining," *International Journal for Numerical Methods in Engineering*, 38, 3675-3694.

McKay, M. D., Beckman, R. J. and Conover, W. J. (1979), "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, 21, 239-245.

Montgomery, D. C. (2004). *Design and Analysis of Experiments*, Wiley, New York.

Morris, M. D. and Mitchell, T. J. (1995), "Exploratory Designs for Computer Experiments," *Journal of Statistical Planning and Inference*, 43, 381-402.

Owen, A. B. (1994), "Controlling Correlations in Latin Hypercube Samples," *Journal of American Statistical Association*, 89, 1517-1522.

Phadke, M. S. (1989), *Quality Engineering Using Robust Design*, Englewood Cliffs, NJ: Prentice Hall.

Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P. (1989), "Design and Analysis of Computer Experiments," *Statistical Science*, 4, 409-423.

Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, Springer, New York.

Taguchi, G. (1987), *Systems of Experimental Design, Vol.1 & Vol 2*, White Plains, New York: Unipub/Kraus International.

Tang, B. (1993), "Orthogonal Array-based Latin Hypercubes," *Journal of American Statistical Association*, 88, 1392-1397.

Tang, B. (1998), "Selecting Latin Hypercubes Using Correlation Criteria," *Statistica Sinica*, 8, 965-978.

Welch, W. J., Buck, R. J., Sacks, J., Wynn, H. P., Mitchell, T. J., and Morris, M. D. (1992), "Screening, Predicting, and Computer Experiments," *Technometrics*, 34, 15-25.

Ye, K.Q. (1998), "Orthogonal Column Latin Hypercubes and their Application in Computer Experiments," *Journal of American Statistical Association*, 93, 1430-1439.

Ye, K. Q., Li, W. and Sudjianto, A. (2000), "Algorithmic Construction of Optimal Symmetric Latin Hypercube Designs," *Journal of Statistical Planning and Inference*, 90, 145-159.