Orthogonal Column Latin Hypercubes and Their Application in Computer Experiments
Author(s): Kenny Q. Ye
Source: *Journal of the American Statistical Association*, Vol. 93, No. 444 (Dec., 1998), pp. 1430-1439
Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association
Stable URL: http://www.jstor.org/stable/2670057

Orthogonal Column Latin Hypercubes and Their Application in Computer Experiments

Kenny Q. YE

Latin hypercubes have been frequently used in conducting computer experiments. In this paper, a class of orthogonal Latin hypercubes that preserves orthogonality among columns is proposed. Applying an orthogonal Latin hypercube design to a computer experiment benefits the data analysis in two ways. First, it retains the orthogonality of traditional experimental designs. The estimates of linear effects of all factors are uncorrelated not only with each other, but also with the estimates of all quadratic effects and bilinear interactions. Second, it can facilitate nonparametric fitting procedures, because one can select good space-filling designs within the class of orthogonal Latin hypercubes according to selection criteria.

KEY WORDS: Computer model; Linear regression; Optimal design; Response surface design.

1. INTRODUCTION

In scientific and engineering research, physical experimentation is often very expensive and time consuming. Because many physical systems can be described by mathematical equations, scientists are able to find numerical solutions of those equations to simulate the systems. Such applications have increased rapidly in the past decade. In contrast to a physical experiment, this kind of experiment is called a computer experiment, and the corresponding computer program a computer model.

Given an input vector $\mathbf{x}$, the response $Y(\mathbf{x})$ of a computer model is often deterministic, so that considerations such as blocking and randomization for physical experiments are irrelevant. Also, in computer experiments, changing the levels of variables is only a matter of setting different numbers for the input, whereas in physical experiments, taking more levels of variables often requires an additional cost of making prototypes and a more elaborate and time-consuming implementation of the experiment. Therefore the differences between computer experiments and traditional physical experiments call for different considerations in design and analysis methods for computer experiments.

1.1 A Cooling System Computer Experiment

At a major manufacturing corporation, engineers use a computer model to simulate the cooling systems of an injection molding process. The computer model contains six input variables, three geometric parameters of the cooling systems and three operating parameters. Two responses are of interest: Δ*temperature,* which measures uniformity of the material temperature, and $T(40)$, the time it takes for the material temperature to reach 40°C. Because it is very costly to build a cooling system and very difficult to measure the temperature of liquid material being injected into a mold, a comparable physical experiment would be very expensive. A good design for the experiment should allow experimenters to fit different parametric and nonparametric models and be efficient for fitting these models. An orthogonal Latin hypercube (OLH) design was chosen for this experiment. It takes advantages of response surface designs, optimal designs, and Latin hypercube designs. I revisit this experiment later in Section 5.

1.2 Optimal Designs

Sacks, Schiller, and Welch (1989) used stochastic processes to model the computer experiments and proposed integrated mean squared error (IMSE) and maximum mean squared error (MMSE) as criteria in selecting designs. Using stochastic processes as the prior distribution of $Y(\mathbf{x})$, Currin, Mitchell, Morris, and Ylvisaker (1991) gave a Bayesian interpretation of the same stochastic model and proposed entropy as a selection criterion. The purpose of these criteria is to improve the precision of prediction. This model and the corresponding analysis method are direct extensions of the kriging (Bayesian kriging) method, which has been widely used in geostatistics. It remains unclear as to what extent the choice of the correlation functions that define the stochastic processes will affect the design.

An alternative approach that does not depend on modeling is to sample points uniformly from the experimental region. Such designs are often called space-filling designs. A primary reason for using space-filling designs is that they can improve interpolation methods. A good measure for uniformity is the Kolmogorov–Smirnov distance between the empirical cumulative distribution function (cdf) of a design and the cdf of the uniform distribution on the experimental region, which is also called discrepancy. Fang and Wang (1994) proposed so-called uniform designs based on low-discrepancy sequences. In practice, the discrepancy is very difficult to evaluate in high dimensions. Other space-filling criteria are maximin and minimax distance designs proposed by Johnson, Moore, and Ylvisaker (1990). The maximin distance design maximizes the minimum distance between points; the minimax distance design minimizes the maximum distance between design points. Johnson et al.

Kenny Q. Ye is Assistant Professor, Department of Applied Math and Statistics, State University of New York, Stony Brook, NY 11794. This research is part of author's Ph.D. dissertation at the University of Michigan and is supported in part by National Science Foundation grant DMS-9704649. The author thanks C. F. Jeff Wu, Rahul Mukerjee, Michael Hamada, the associate editor, and an anonymous referee for insightful comments on early versions of this article. The author is also grateful to Joanne Wood and J. Michael Garrett of the Ford Motor Company.

Journal of the American Statistical Association
December 1998, Vol. 93, No. 444, Theory and Methods

(1990) also found that maximin distance designs are special cases of maximum entropy designs when the prior correlation among design points goes to 0.

1.3 Latin Hypercube Designs

An $n \times d$ Latin hypercube consists of d permutations of $S_n = \{1, 2, \ldots, n\}$. A Latin hypercube design takes row vectors as the experimental points in an d-dimensional space. One-dimensional projections of a Latin hypercube design are evenly spaced and have the lowest possible discrepancy.

Latin hypercube sampling was first proposed by McKay, Beckman, and Conover (1979) in their pathbreaking paper on computer experiments. Since then, some efforts have been made to improve Latin hypercube designs for computer experiments along three different directions.

Independently, Owen (1992) and Tang (1993) proposed orthogonal array (OA)-based Latin hypercubes. An $n \times m$ matrix $\mathbf{A}$ with entries from a set of $s \geq 2$ symbols is called an OA of strength r if each $n \times r$ submatrix of $\mathbf{A}$ contains all possible s^r row vectors with the same frequency. Geometrically, r-dimensional projections of a strength-r-OA–based Latin hypercube are all stratified, so it has a better-space filling property.

Some other researchers have looked for optimal Latin hypercubes. Among them, Park (1994) studied optimal Latin hypercube designs based on the IMSE criterion; Morris and Mitchell (1995) investigated optimal Latin hypercubes based on the entropy and maximin distance criteria; and Schonlau, Hamada, and Welch (1996) proposed the use of Latin hypercubes that maximize the minimum distance between points (i.e., the covariate vectors) in low-dimensional projections.

Finally, Iman and Conover (1982) and Owen (1994) proposed using Latin hypercubes with small off-diagonal correlations. Although there are other motivations, it is clear that such designs for computer experiments would benefit linear regression models, as the estimates of linear effects of the input variables are only slightly correlated. Tang (1998) further proposed controlling not only the correlations between linear effects, but also higher-order polynomial effects. In all three works, algorithms were used to obtain such Latin hypercubes.

1.4 Orthogonal Latin Hypercubes

The correlation between two vectors $\mathbf{u} = [u_1, u_2, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, \ldots, v_n]$ is defined as $[\sum(v_i - \bar{v})(u_i - \bar{u})]/\sqrt{[\sum(u_i - \bar{u})^2 \sum(v_i - \bar{v})^2]}$, where $\bar{u} = 1/n \sum u_i$ and $\bar{v} = 1/n \sum v_i$. In this article, a class of Latin hypercubes for which every pair of columns has zero correlation is constructed. Tables 1 and 2 show two such Latin hypercubes, called OLHs. A fully algebraic construction method for OLH is presented in Section 2. Properties of OLHs are presented in Section 3. Based on some algebraic properties, a class of OLHs can be generated using a simple transformation so that optimal designs within this class can be obtained through computer search. In Section 4 a computer algorithm is proposed to search for optimal OLH designs. Finally, the cooling system computer experiment using an OLH design is revisited in Section 5.

Table 1. A 5 × 2 Orthogonal Latin Hypercube

1	−2
2	1
0	0
−1	2
−2	−1

Table 2. A 9 × 4 Orthogonal Latin Hypercube

1	−2	4	3
2	1	3	−4
3	−4	−2	−1
4	3	−1	2
0	0	0	0
−4	−3	1	−2
−3	4	2	1
−2	−1	−3	4
−1	2	−4	−3

2. CONSTRUCTION OF ORTHOGONAL LATIN HYPERCUBES

One can construct an OLH when its number of rows n is a power of 2 or a power of 2 + 1. For $n = 2^m$ or $2^m + 1$, an OLH with $2m - 2$ columns will be constructed. Before laying out the construction method, I would like to give a brief review on permutation groups. (For more information on group theory and permutation groups, see Hungerford 1974.)

A *permutation of a set* $S = \{1, 2, \ldots, N\}$ is a rearrangement of the elements of the set. Rigorously speaking, it is a mapping from S to S that is both one-to-one and onto. Let σ and τ be two permutations of S. It can be shown that their composition $\sigma\tau$ is also a permutation of S. The collection of all permutations of S is then a group under composition. The composition in such a group is also called a product of permutations. Let σ be a permutation of S that maps j to i_j, which is represented by

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & N \\ i_1 & i_2 & \cdots & i_N \end{pmatrix}.$$

Equivalently, one can use a matrix $\mathbf{P}$ to represent σ in which $P_{j i_j} = 1$ and rest of the elements are 0. Therefore, $\mathbf{P}$ is called a *permutation matrix.* Let $\mathbf{e} = [1, 2, \ldots, N]^T$ be a vector. The product of $\mathbf{P}$ and $\mathbf{e}, \mathbf{Pe} = [i_1, i_2, \ldots, i_N]^T$, is called a permutation of $\mathbf{e}$. It can be shown that the permutation matrix of $\sigma\tau$ is $\mathbf{P}_\sigma \mathbf{P}_\tau$, the regular matrix product of the permutation matrices of σ and τ.

Let $i_1, i_2, \ldots, i_r$ be distinct elements of S. Then $(i_1 i_2 \cdots i_r)$ denotes the permutation that maps $i_1 \mapsto i_2, i_2 \mapsto i_3, \ldots, i_{r-1} \mapsto i_r$, and $i_r \mapsto i_1$ and maps every other element of S to itself. Then $(i_1 i_2, \ldots, i_r)$ is called an r-cycle; a 2-cycle is called a *transposition.* It can be shown that every nonidentity permutation is a product of transpositions.

Next, OLHs with $n = 2^m + 1$ runs will be constructed, and then modified to OLHs with 2^m runs.

2.1 $n = 2^m + 1$, $m \geq 2$

To illustrate the construction method more clearly, denote n levels of each column by $\{-2^{m-1}, \ldots, -1, 0, 1, \ldots, 2^{m-1}\}$ instead of $\{1, \ldots, n\}$. Thus two columns $\mathbf{u}$ and $\mathbf{v}$ have zero correlation if $\mathbf{u}^T\mathbf{v} = 0$; that is, $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

For $m = 1$, it is easy to verify that there is no pair of permutations of [1, 2, 3] that are orthogonal. For $m = 2$ and $m = 3$, a 5×2 OLH and a 9×4 OLH are shown in Tables 1 and 2. Notice that in these two OLHs, the bottom half of the OLH is precisely the mirror image of the top half.

Now consider the general case of $n = 2^m + 1$, for which $2m - 2$ mutually orthogonal columns are constructed. Denote the top half of the Latin hypercube by $\mathbf{T}$, which is a $2^{m-1} \times (2m-2)$ matrix. The bottom half is taken as the mirror image of the top half, and a center point is added to complete the OLH.

The remainder of the section is devoted to the construction of the matrix $\mathbf{T}$.

Definition The $\mathbf{M}$ *matrix* is the matrix of which each entry is the absolute value of the corresponding entry in $\mathbf{T}$. The $\mathbf{S}$ *matrix* is the matrix of which each entry is taken as 1 or -1 according to whether the corresponding entry in $\mathbf{T}$ is positive or negative.

Therefore, $\mathbf{T}$ is the elementwise product (or called Hadamard product) of the $\mathbf{M}$ matrix and the $\mathbf{S}$ matrix.

First, I construct the $\mathbf{M}$ matrix. Tables 1 and 2 show that each column of the $\mathbf{M}$ matrix is a permutation of $\{1, 2, \ldots, 2^{m-1}\}$. In the group of permutations on $\{1, 2, \ldots, 2^{m-1}\}$, define the permutations $\{A_k; k = 1, \ldots, m-1\}$ in terms of the composition of the transpositions,

$$A_k = \prod_{j=1}^{2^{m-k-1}} \left\{ \prod_{i=1}^{2^{k-1}} ((j-1)2^k + i \quad j2^k + 1 - i) \right\},$$

where $\prod$ represents the composition of permutations. An equivalent definition of A_k using permutation matrix is then

$$\mathbf{A_k} = \underbrace{\mathbf{I} \otimes \cdots \otimes \mathbf{I}}_{m-1-k} \otimes \underbrace{\mathbf{R} \otimes \cdots \otimes \mathbf{R}}_{k}, \tag{1}$$

where $\mathbf{I}$ is the 2×2 identity matrix,

$$\mathbf{R} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

and $\otimes$ is the Kronecker product.

Example 1. For $m = 4$,

$$A_1 = (12)(34)(56)(78),$$
$$A_2 = (14)(23)(58)(67),$$

and

$$A_3 = (18)(27)(36)(45)$$

in terms of the composition of transpositions. Let $\mathbf{e} = [1, 2, 3, 4, 5, 6, 7, 8]^T$. Then

$$\mathbf{A}_1\mathbf{e} = (\mathbf{I} \otimes \mathbf{I} \otimes \mathbf{R})\mathbf{e} = [2, 1, 4, 3, 6, 5, 8, 7]^T,$$

$$\mathbf{A}_2\mathbf{e} = (\mathbf{I} \otimes \mathbf{R} \otimes \mathbf{R})\mathbf{e} = [4, 3, 2, 1, 8, 7, 6, 5]^T,$$

and

$$\mathbf{A}_3\mathbf{e} = (\mathbf{R} \otimes \mathbf{R} \otimes \mathbf{R})\mathbf{e} = [8, 7, 6, 5, 4, 3, 2, 1]^T.$$

It can be seen from the example that A_1 is the permutation that switches every two neighboring elements, A_2 flips every four neighboring elements, A_3 flips every eight neighboring elements, and so on. An important fact is that the subgroup G generated by $\{A_k, k = 1, \ldots, m-1\}$ is commutative and its elements are reflexive.

Theorem 1. The subgroup G generated by $\{A_k\}$ is commutative.

Theorem 2. All elements in G are reflexive; that is

$$\sigma(\sigma(j)) = j \tag{2}$$

for all $1 \leq j \leq n$. It implies that $\sigma^2 = e$, where e is the permutation that maps every element into itself. Note that $\sigma^2 = e$ is equivalent to $\sigma^{-1} = \sigma$.

Columns of the $\mathbf{M}$ matrix are permutations of $[1, 2, \ldots, 2^{m-1}]$. They consist of

$$\{\mathbf{e}, \mathbf{A}_i\mathbf{e}, \mathbf{A}_{m-1}\mathbf{A}_j\mathbf{e}; i = 1, \ldots, m-1, j = 1, \ldots, m-2\},$$

where $\mathbf{e} = [1, 2, \ldots, 2^{m-1}]^T$. With a properly chosen $\mathbf{S}$ matrix, columns of $\mathbf{T}$, the Hadamard product of the $\mathbf{M}$ and $\mathbf{S}$ matrices, are mutually orthogonal.

Next, consider the construction of the $\mathbf{S}$ matrix that attaches the sign to each element in $\mathbf{T}$. For $k = 1, \ldots, m-1$, define the vector $\mathbf{a}_k$ as

$$\mathbf{a}_k = \mathbf{B}_1 \otimes \mathbf{B}_2 \otimes \cdots \otimes \mathbf{B}_{m-1}, \tag{3}$$

where

$$\mathbf{B}_k = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \qquad \mathbf{B}_i = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

for $i \neq k$. Denote $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ by $\mathbf{1}$, $\left[\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right]$ by $\mathbf{a}$. Notice that $\{\mathbf{a}_k, k = 1, \ldots, m-1\}$ are the $m-1$ column vectors that generate the 2^{m-1} factorial design. They generate a group under the elementwise product of vectors (also called the Schur product).

Example 1 (Continued). In the case of $m = 4$,

$$a_1 = \mathbf{1} \otimes \mathbf{1} \otimes \mathbf{a} = [-1, +1, -1, +1, -1, +1, -1, +1]^T,$$

$$a_2 = \mathbf{1} \otimes \mathbf{a} \otimes \mathbf{1} = [-1, -1, +1, +1, -1, -1, +1, +1]^T,$$

and

$$a_3 = \mathbf{a} \otimes \mathbf{1} \otimes \mathbf{1} = [-1, -1, -1, -1, +1, +1, +1, +1]^T.$$

To complete the construction of $\mathbf{T}$, columns of the $\mathbf{S}$ matrix are chosen to be

$$\{\mathbf{1}, \mathbf{a}_i, \mathbf{a}_1\mathbf{a}_{j+1}; i = 1, \ldots, m-1, j = 1, \ldots, m-2\}.$$

Because $\mathbf{T}$ is the Hadamard product of the $\mathbf{M}$ and the $\mathbf{S}$ matrices, the columns of $\mathbf{T}$ consist of

$$\{\mathbf{1} \times \mathbf{e}, \mathbf{a}_i \times \mathbf{A}_i\mathbf{e}, (\mathbf{a}_1\mathbf{a}_j) \times (\mathbf{A}_j\mathbf{A}_{m-1}\mathbf{e}); \quad i = 1, \ldots, m-1, j = 1, \ldots, m-2\}, \quad (4)$$

where "×" is the elementwise product. Therefore, (a) for column $\mathbf{A}_i\mathbf{e}$ in the $\mathbf{M}$ matrix, the corresponding column in the $\mathbf{S}$ matrix is $\mathbf{a}_i$, for $i = 1, \ldots, m-1$; (b) For column $\mathbf{A}_i\mathbf{A}_{m-1}\mathbf{e}$ in the $\mathbf{M}$ matrix, the corresponding column in the $\mathbf{S}$ matrix is $\mathbf{a}_1\mathbf{a}_{i+1}$, for $i = 1, \ldots, m-2$; and (c) For $\mathbf{e} = [1, 2, \ldots, 2^{m-1}]$ in the $\mathbf{M}$ matrix, the corresponding column in the $\mathbf{S}$ matrix is $\mathbf{1} = [+1, +1, \ldots, +1]$.

In the next section I prove that each pair of columns of $\mathbf{T}$ is orthogonal.

Example 1 (Continued). In the case of $m = 4$, the $\mathbf{M}$ matrix is

$$\begin{array}{cccccc}
\mathbf{e} & A_1 & A_2 & A_3 & A_1A_3 & A_2A_3 \\
1 & 2 & 4 & 8 & 7 & 5 \\
2 & 1 & 3 & 7 & 8 & 6 \\
3 & 4 & 2 & 6 & 5 & 7 \\
4 & 3 & 1 & 5 & 6 & 8 \\
5 & 6 & 8 & 4 & 3 & 1 \\
6 & 5 & 7 & 3 & 4 & 2 \\
7 & 8 & 6 & 2 & 1 & 3 \\
8 & 7 & 5 & 1 & 2 & 4
\end{array}$$

The corresponding $\mathbf{S}$ matrix is

$$\begin{array}{cccccc}
I & a_1 & a_2 & a_3 & a_1a_2 & a_1a_3 \\
+1 & -1 & -1 & -1 & +1 & +1 \\
+1 & +1 & -1 & -1 & -1 & -1 \\
+1 & -1 & +1 & -1 & -1 & +1 \\
+1 & +1 & +1 & -1 & +1 & -1 \\
+1 & -1 & -1 & +1 & +1 & -1 \\
+1 & +1 & -1 & +1 & -1 & +1 \\
+1 & -1 & +1 & +1 & -1 & -1 \\
+1 & +1 & +1 & +1 & +1 & +1
\end{array}$$

Therefore, $\mathbf{T}$ is the elementwise product of the $\mathbf{M}$ matrix and the $\mathbf{S}$ matrix as follows:

$$\begin{array}{cccccc}
1 & -2 & -4 & -8 & 7 & 5 \\
2 & 1 & -3 & -7 & -8 & -6 \\
3 & -4 & 2 & -6 & -5 & 7 \\
4 & 3 & 1 & -5 & 6 & -8 \\
5 & -6 & -8 & 4 & 3 & -1 \\
6 & 5 & -7 & 3 & -4 & 2 \\
7 & -8 & 6 & 2 & -1 & -3 \\
8 & 7 & 5 & 1 & 2 & 4
\end{array}.$$

By adding a center point and the mirror image of the matrix just given, a 17×6 OLH is obtained; see Table 3.

2.2 $n = 2^m$, $m \geq 2$

Given an OLH with $n = 2^m + 1$ runs as constructed previously, by removing its center point $(0, 0, \ldots, 0)$ and rescaling the levels to be equidistant, one can construct an OLH with $n = 2^m$ runs. For example, one can modify the 9×4 OLH in Table 2 to an 8×4 OLH shown in Table 4.

3. PROPERTIES OF ORTHOGONAL LATIN HYPERCUBES

In this section I prove the orthogonal properties of the Latin hypercubes as constructed in Section 2. I also demonstrate that a class of OLHs can be generated by two simple operations.

3.1 Orthogonality

In regression analysis, one would like independent variables in a regression model to be orthogonal to each other, so that the estimates of the coefficients are uncorrelated. Consider a qth order polynomial model with m predictors,

$$Y = \beta_0 + \sum \beta_i x_i + \sum_{i \leq j \leq m} \beta_{ij} x_i x_j + \cdots + \sum_{i_1 \leq i_2 \leq \cdots \leq i_q \leq m} \beta_{i_1 i_2 \cdots i_q} x_{i_1} x_{i_2} \cdots x_{i_q} + \varepsilon.$$

Define β_i to be the linear effects of x_i, β_{ii} to be the quadratic effects of x_i, and β_{ij} to be the bilinear interaction effects of x_i and x_j for $i \neq j$. The OLH designs ensure the independence of estimates of linear effects of each variable. Moreover, the estimates of the quadratic effects and bilinear interaction effects are uncorrelated with the estimates of the linear effects.

Theorem 3. For the Latin hypercubes constructed in Section 2, each column is orthogonal to the others.

Table 3. A 17 × 6 Orthogonal Latin Hypercube

1	−2	−4	−8	7	5
2	1	−3	−7	−8	−6
3	−4	2	−6	−5	7
4	3	1	−5	6	−8
5	−6	−8	4	3	−1
6	5	−7	3	−4	2
7	−8	6	2	−1	−3
8	7	5	1	2	4
0	0	0	0	0	0
−1	2	4	8	−7	−5
−2	−1	3	7	8	6
−3	4	−2	6	5	−7
−4	−3	−1	5	−6	8
−5	6	8	−4	−3	1
−6	−5	7	−3	4	−2
−7	8	−6	−2	1	3
−8	−7	−5	−1	−2	−4

Table 4. An 8 × 4 Orthogonal Latin Hypercube

.5	−1.5	3.5	2.5
1.5	.5	2.5	−3.5
2.5	−3.5	−1.5	−.5
3.5	2.5	−.5	1.5
−3.5	−2.5	.5	−1.5
−2.5	3.5	1.5	.5
−1.5	−.5	−2.5	3.5
−.5	1.5	−3.5	−2.5

Proof. See the Appendix.

Theorem 4. For the Latin hypercubes constructed in Section 2, the elementwise square of each column is orthogonal to all the columns in the Latin hypercube.

Proof. Let $\mathbf{X}$ be a column of a Latin hypercube. Because the bottom half of the OLH is the mirror image of the top half, $\mathbf{X}_i = -\mathbf{X}_{n+1-i}$ for $i = 1, \ldots, n$. For its elementwise square $\mathbf{SX}, (\mathbf{SX})_i = (\mathbf{X}_i)^2 = (-\mathbf{X}_{n+1-i})^2 = (\mathbf{SX})_{n+1-i}$ for $i = 1, \ldots, n$. Hence $\sum(\mathbf{SX})_i\mathbf{X}_i = 0$, with $\mathbf{SX}$ and $\mathbf{X}$ orthogonal to each other. Let $\mathbf{Y}$ be another column of the Latin hypercube, $\mathbf{Y}_i = -\mathbf{Y}_{n+1-i}$ for $i = 1, \ldots, n$. Therefore, $\sum(\mathbf{SX})_i\mathbf{Y}_i = 0$; that is, $\mathbf{SX}$ and $\mathbf{Y}$ are orthogonal to each other.

Theorem 5. For the Latin hypercubes constructed in Section 2, the elementwise product of every two columns is orthogonal to all columns in the Latin hypercube.

Proof. For any two columns $\mathbf{X}$ and $\mathbf{Y}$ of the Latin hypercube, $\mathbf{X}_i = -\mathbf{X}_{n+1-i}, \mathbf{Y}_i = -\mathbf{Y}_{n+1-i}$ for $i = 1, \ldots, n$. For their elementwise product $\mathbf{XY}, (\mathbf{XY})_i = \mathbf{X}_i\mathbf{Y}_i = (-\mathbf{X}_{n+1-i})(-\mathbf{Y}_{n+1-i}) = (\mathbf{XY})_{n+1-i}$. Therefore, $\mathbf{XY}$ is orthogonal to all of the columns in the Latin hypercube.

Theorems 4 and 5 are derived from the symmetrical nature of the design. They guarantee that the estimates of quadratic effects and bilinear interaction effects are uncorrelated with the estimates of linear effects. However, the estimates of quadratic and bilinear interaction effects are correlated with each other. If one assumes that linear effects of each variable are more likely to be important and that only a small number of quadratic and bilinear interaction effects are significant, then OLHs are good designs for fitting second-order polynomial models.

3.2 A Class of Orthogonal Latin Hypercubes

Note that the orthogonality of the Latin hypercubes constructed in Section 2 does not depend on the numerical values of the levels. This allows one to generate a new OLH by replacing $\mathbf{e}$ by one of its permutations. This can be used to define a class of OLHs. Also, reversing the signs of columns of an OLH will not change the orthogonality.

Table 5. A 17 × 6 Maximin (L_1) Distance Orthogonal Latin Hypercube

	A	B	C	D	E	F
1	2	−1	−8	−6	7	4
2	1	2	−3	−7	−6	−5
3	3	−8	1	−5	−4	7
4	8	3	2	−4	5	−6
5	4	−5	−6	8	3	−2
6	5	4	−7	3	−8	1
7	7	−6	5	1	−2	−3
8	6	7	4	2	1	8
9	0	0	0	0	0	0
10	−6	−7	−4	−2	−1	−8
11	−7	6	−5	−1	2	3
12	−5	−4	7	−3	8	−1
13	−4	5	6	−8	−3	2
14	−8	−3	−2	4	−5	6
15	−3	8	−1	5	4	−7
16	−1	−2	3	7	6	5
17	−2	1	8	6	−7	−4

Table 6. A 17 × 6 Maximin (L_2) Distance Orthogonal Latin Hypercube

	A	B	C	D	E	F
1	5	−1	−4	−3	8	6
2	1	5	−2	−8	−3	−7
3	2	−4	1	−7	−6	8
4	4	2	5	−6	7	−3
5	6	−7	−3	4	2	−5
6	7	6	−8	2	−4	1
7	8	−3	7	1	−5	−2
8	3	8	6	5	1	4
9	0	0	0	0	0	0
10	−3	−8	−6	−5	−1	−4
11	−8	3	−7	−1	5	2
12	−7	−6	8	−2	4	−1
13	−6	7	3	−4	−2	5
14	−4	−2	−5	6	−7	3
15	−2	4	−1	7	6	−8
16	−1	−5	2	8	3	7
17	−5	1	4	3	−8	−6

Theorem 6. Theorems 3, 4, and 5 still hold if signs of any subset of the columns are reversed in the OLHs.

Proof. This is easily seen by the definition and symmetrical nature of the designs.

Theorem 7. Replacing $\mathbf{e}$ in (4) by one of its permutations generates a new Latin hypercube that satisfies the properties in Theorems 3, 4, and 5.

Reversing signs of a subset of columns is equivalent to flipping the design points over a hyperplane. Because the rectangular (L_1) distance, Euclidean (L_2) distance, and other commonly used distances are invariant to this change, the space-filling property of the design should remain the same. Therefore, one is less interested in such changes. On the other hand, applying permutations on the $\mathbf{M}$ matrix would generate designs with different space-filling properties.

4. OPTIMAL ORTHOGONAL LATIN HYPERCUBE DESIGNS

Although an OLH design has zero correlations between input variables, it does not necessarily have a good space-filling property, which is desirable for data analysis methods such as residual plots in regression diagnostics and nonparametric surface fitting. By Theorems 6 and 7, a class of OLHs is generated by sign reversal and permutation of the OLH constructed in Section 2. Thus an optimal OLH design can be selected using design criteria such as IMSE, MMSE, entropy, minimax and maximin distance, and discrepancy.

For an OLH constructed in the previous sections, criteria such as IMSE, MMSE, entropy, and minimax and maximin distance are all invariant to the sign reversal of columns, because these criteria are all based on the distance between the design points. I used a simple algorithm that permutes only the $\mathbf{M}$ matrix to search optimal designs within the class of OLHs based on one of the aforementioned criteria:

Step 1. Randomly select a permutation of $\mathbf{e}, \mathbf{e_1}$. Construct an OLH S using (4) as the starting design; evaluate its criterion cr. Without loss of generality, assume that cr is going to be maximized.

Step 2. Generate $[2^{m-1}(2^{m-1}+1)]/2$ permutations of $\mathbf{e_1}$ by transposing i and j for $i \neq j \in \{1, 2, \ldots, 2^{m-1}\}$. Evaluate the criterion for the $[2^{m-1}(2^{m-1}+1)]/2$ Latin hypercubes; choose S_1 to be the Latin hypercubes with the best criterion cr'.

Step 3. If $cr' > cr$, then set $S = S_1$ and $cr = cr'$; go to step 2. If $cr' \leq cr$, then stop the search.

Step 4. Repeat steps 1–3 with r different starting Latin hypercubes.

Example 2. Maximum Distance Orthogonal Latin Hypercubes. Maximin distance is widely considered as a good criterion for spacing filling, particularly for designs with moderately large number of runs. It avoids the problem of choosing the correlation functions for model-based criteria such as IMSE, MMSE, and entropy. One can choose from many kinds of distances, however. Rectangular (L_1) distance and Euclidean (L_2) distance are the two most commonly used in practice.

For $n = 17$, maximin distance OLHs were obtained using 100 different starting OLHs. The results based on the L_1 and L_2 distance are given in Tables 5 and 6. Their projection to two dimensions are shown in Figures 1 and 2.

5. AN COMPUTER EXPERIMENT: COOLING SYSTEM OF INJECTION MOLDING

In an injection molding process, hot liquid material is injected into a mold and then forms to the desired shape. To cool down the hot molding material, cooling systems are often used in injection molding processes. A good cooling system should reduce the material temperature quickly and uniformly. Because it is very difficult to measure the material temperature and to build prototypes of cooling systems, physical experiments would be very costly. At a major manufacturing corporation, engineers use a computer model to simulate the cooling systems of an injection molding pro-

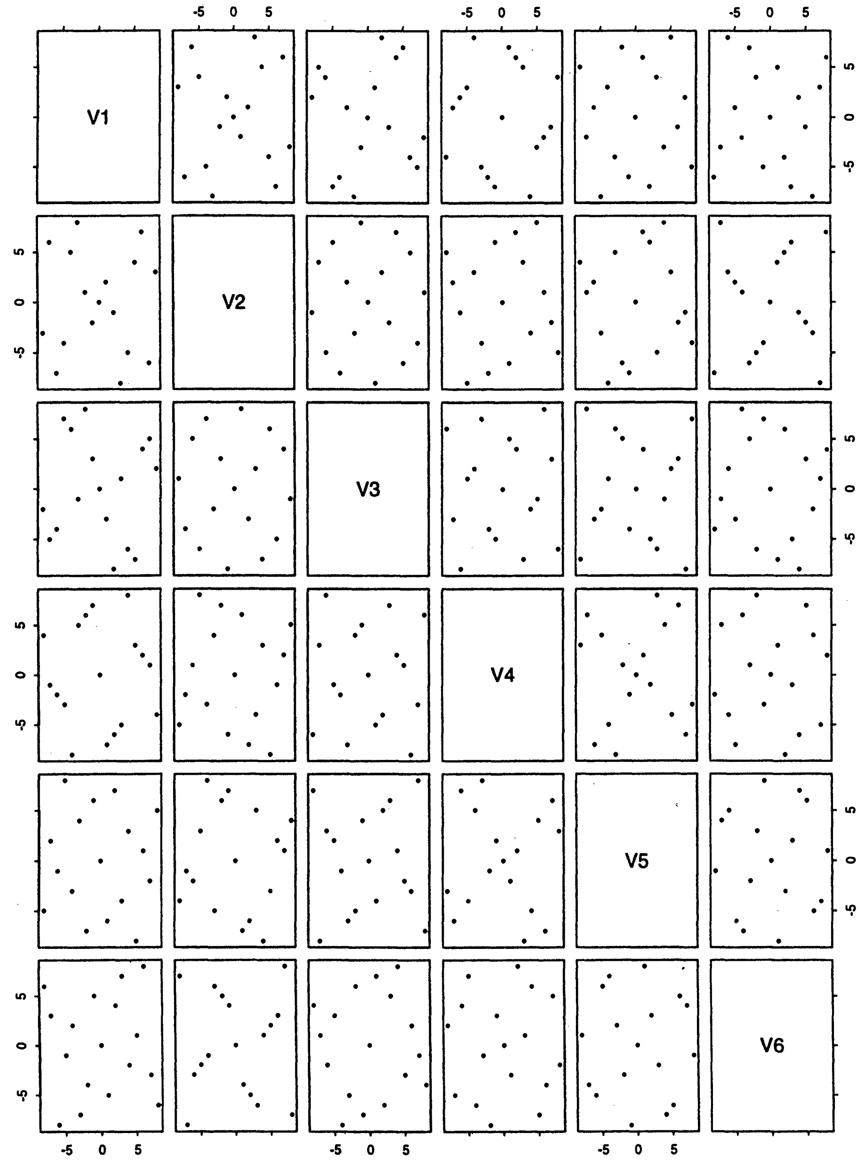


Figure 1. Projection to Two Dimensions of a Maximin (L_1) Distance Orthogonal Latin Hypercube.

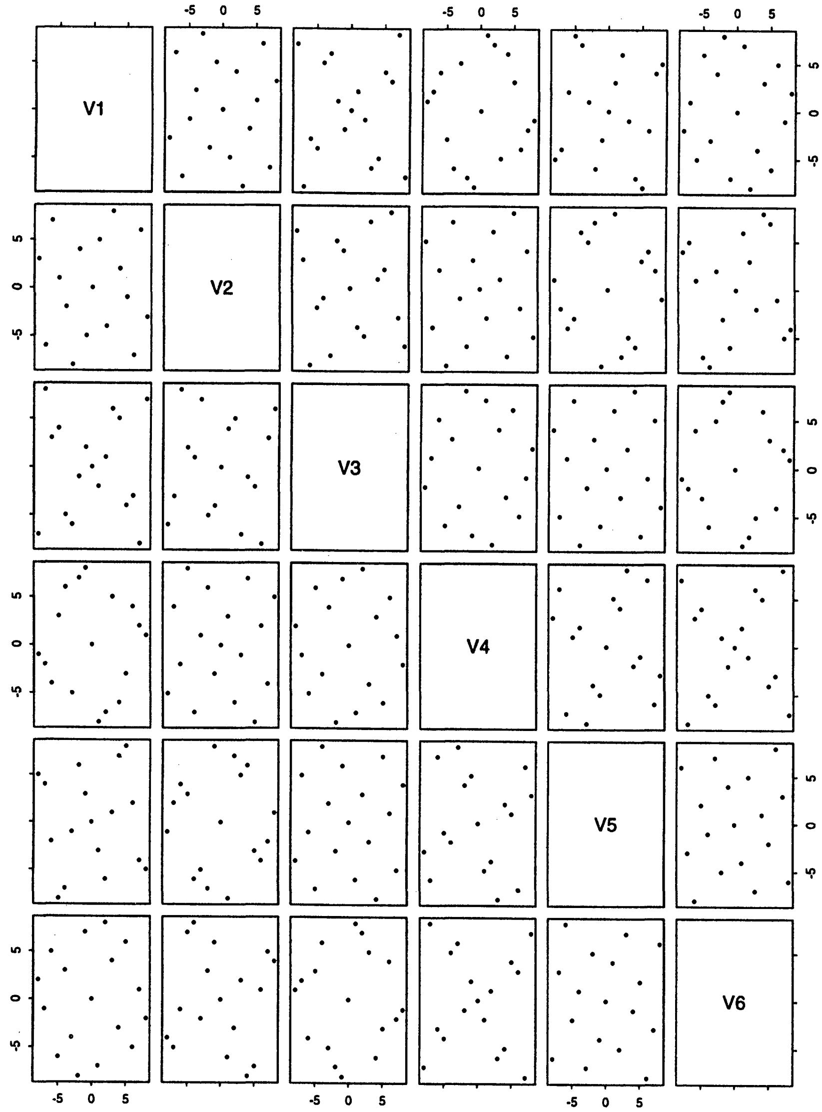


Figure 2. Projection to Two Dimensions of a Maximin (L_2) Distance Orthogonal Latin Hypercube.

cess. The computer model contains six input variables: distance between cooling pipes (A), diameter of cooling pipes (B), distance to mold (C), flow rate (D), coolant temperature (E), and percentage of ethylene glycol (F). Two responses, Δ*temperature* and $T(40)$, are of interest. The former is a measurement of quality; the latter, a measurement of productivity.

A 17-run OLH design as shown in Table 7 was adopted for the computer experiment. Because variable F has only six levels in the model, 17 levels of the corresponding column in the Latin hypercube are collapsed to six levels as follows:

$$\{-8,-7,-6\} \rightarrow -7.5, \{-5,-4,-3\} \rightarrow -4.5, \{-2,-1,0\} \rightarrow -1.5$$

$$\{8,7,6\} \rightarrow 7.5, \{5,4,3\} \rightarrow 4.5, \{2,1\} \rightarrow 1.5.$$

Therefore the correlations between F and the other variables are not exactly 0 but remain less than .05. The design and the results of the computer experiment are shown in Table 7. The experiment domain is $[-1,1]^6$. The real scale of the input variables is omitted, because this is proprietary information.

Two regression models are built for the two responses separately. Residual plots, variable selection methods, and other tools in regression analysis were used to select the final models. Note that the error is due to model bias, not to sampling variation. The regression model for Δtemperature includes linear and the quadratic effects of A and C, the bilinear interaction of A and C, and the linear effect of E. The corresponding R^2 is 99.4%. The regression equation is

$$\Delta\text{temperature} = 2.73 + 4.36A + .965A^2 - 3.58C + 2.58C^2 - 4.06AC - .505E. \quad (5)$$

The regression model for $T(40)$ includes linear effects of all six variables, the quadratic effect of C, and the bilinear interaction of A and C; the corresponding R^2 is 99.3%. The

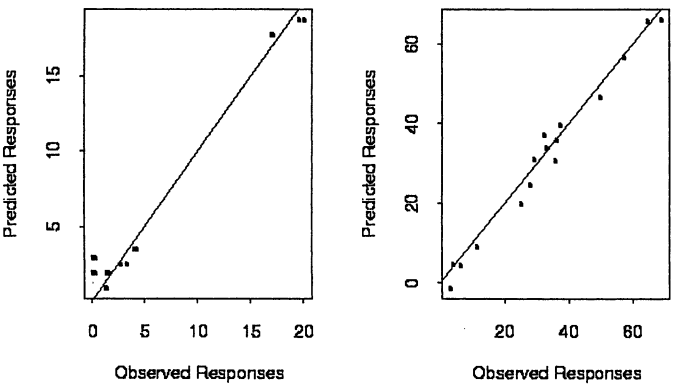
Table 7. The Design and the Responses of the Cooling System Computer Experiment: A 17-Run OLH Design and a 16-Run Fractional Factorial Design

Run	T40	ΔTemp	A	B	C	D	E	F
1	41.87	.1	−1	−.75	.5	.125	.625	−1
2	32.44	0	−.875	−.375	.125	−.5	−.25	1
3	28.92	0	−.75	1	.25	.625	−.125	−.6
4	14.59	1.39	−.625	−.125	−.375	−.75	−1	−.2
5	13.66	5	−.5	.25	−1	.875	−.375	.2
6	50.98	.3	−.375	.875	.625	−.25	.5	1
7	37.17	4.5	−.25	−.5	−.75	.375	.875	.6
8	30.31	6.7	−.125	.625	−.875	−1	.75	−.6
9	33.28	3.09	0	0	0	0	0	−.2
10	33.89	1.6	.125	−.625	.875	1	−.75	.6
11	29.61	2.8	.25	.5	.75	−.375	−.875	−.6
12	18.43	9.2	.375	−.875	−.625	.25	−.5	−1
13	53.38	2.09	.5	−.25	1	−.875	.375	−.2
14	56.65	3.4	.625	.125	.375	.75	1	.2
15	40.39	7.8	.75	−1	−.25	−.625	.125	.6
16	35.99	8.1	.875	.375	−.125	.5	.25	−1
17	20.31	12.9	1	.75	−.5	−.125	−.625	1
1	32.19	1.31	−1	−1	−1	1	1	1
2	11.35	19.6	1	−1	−1	−1	−1	1
3	29.26	1.4	−1	1	−1	−1	1	−1
4	3.37	20	1	1	−1	1	−1	−1
5	25.17	.11	−1	−1	1	1	−1	−1
6	64.72	3.3	1	−1	1	−1	1	−1
7	28.08	.4	−1	1	1	−1	−1	1
8	68.99	2.7	1	1	1	1	1	1
9	33.16	4	1	1	1	−1	−1	−1
10	49.88	.4	−1	1	1	1	1	−1
11	37.17	4.3	1	−1	1	1	−1	1
12	57.47	.11	−1	−1	1	−1	1	1
13	36.18	17	1	1	−1	−1	1	1
14	4.13	1.59	−1	1	−1	1	−1	1
15	35.78	17.11	1	−1	−1	1	1	−1
16	6.35	1.4	−1	−1	−1	−1	−1	−1

regression equation is

$$T(40) = 35.0 + 3.43A - 1.38B + 12.6C - 1.11D + 14.7E + 2.68F - 3.45C^2 + 3.78AC. \quad (6)$$

In the model for $T(40)$, the effects involved with A, C, and E are much bigger than the other effects. Therefore, the other variables could be considered practically insignificant.



(a) (b)

Figure 3. Cooling System Application: Predicted Responses Versus Observed Responses. (a) Delta temperature; (b) T(40).

A 16-run 2^{6-2}_{IV} design was used to conduct a computer experiment before the 17-run OLH design. It is impossible to estimate the quadratic effects using this 16-run two-level experiment. However, it can be used to verify that models (5) and (6) are not overfitted. Figure 3 shows the scatterplots of the predicted responses according to (5) and (6) against the observed responses at these 16 points. Notice that these 16 points are at the corners of the experimental region. The figure shows that the regression models describe the system very well in the experimental region.

6. DISCUSSION

A class of orthogonal Latin hypercubes (OLHs) has been constructed algebraically, and optimal designs based on some criteria can be searched within this class using a simple algorithm. Such designs have two advantages over the Latin hypercube designs proposed by Iman and Conover (1982), Owen (1992), and Tang (1994). First, no complicated algorithms are used to achieve small correlations between input variables; instead, the correlations are guaranteed to be 0. Second, other criteria can be used in the choice designs within the class. One the other hand, a disadvantage of the OLH is that the run size is less flexible, and a large number of runs may be needed to handle a large number of input variables. However, this disadvantage is not always as severe for computer experiments as it can be for physical experiments. Computer experiments are sometimes much cheaper and faster to run than physical experiments, and a moderately large number of runs are often affordable. Some computer experiments have used up to tens of thousands of runs.

There is no standard statistical model for a computer experiment, and little is known about which parametric, nonparametric, or semiparametric model would fit the data well before the data are collected. Therefore, designs for computer experiments should facilitate different methods of multivariate data analysis. Traditional factorial designs are not favored in computer experiments, because they are not versatile in fitting different statistical models. However, in my experience, low-order polynomial models work well in many cases, as in the cooling system experiment in Section 5. In these cases, OLH designs have the same desirable properties as traditional factorial design. On the other hand, if polynomial models do not fit the data well, then the OLH designs still facilitate modern multivariate analysis methods such as Kriging, MARS, GAM, LOESS, and so on.

APPENDIX: PROOFS OF THEOREMS

Lemma A.1. Let $\mathbf{A}_1, \mathbf{B}_1, \mathbf{A}_2, \mathbf{B}_2$ be $m \times n, p \times q, n \times l$, and $q \times r$ matrices. Then $(\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) = (\mathbf{A}_1\mathbf{A}_2) \otimes (\mathbf{B}_1\mathbf{B}_2)$.

Proof. See the work of Rao (1973, chap. 1).

Proof of Theorem 1

One need only prove that the matrix products of permutation matrices are commutative. All of the elements of G are Kronecker products of $\mathbf{R}$ and $\mathbf{I}$. Using Lemma A.1, it is sufficient to observe that $\mathbf{RI} = \mathbf{IR}, \mathbf{R}^2 = \mathbf{R}^2$, and $\mathbf{I}^2 = \mathbf{I}^2$.

Proof of Theorem 2

It is equivalent to prove $\mathbf{P}^2 = \mathbf{I}$, where $\mathbf{P}$ is the permutation matrix associated with σ. Using Lemma A.1, this is true for $\mathbf{A}_i$'s, because $\mathbf{R}^2 = \mathbf{I}^2 = \mathbf{I}$ and $\otimes_{i=1}^{m-1}\mathbf{I} = I_{2^{m-1}\times 2^{m-1}}$. Using Theorem 1, it is true for any permutation matrix generated by $\{\mathbf{A}_i\}$.

Proof of Theorem 3

To prove this, some algebraic structure must first be established. Let G_l be the group of all permutations of $\{1, 2, \ldots, 2^l\}$ and let H_l be the group generated by $\{-\mathbf{1}, \mathbf{a}_k, k = 1, 2, \ldots, l\}$, where $\mathbf{a}_k$'s are defined in (3) and $-\mathbf{1}$ is the vector with all of the 2^l elements being -1. Equivalently, G_l is the group of all $2^l \times 2^l$ permutation matrices. The operation of the group H_l is the elementwise product. The group generated by $\{\mathbf{a}_i, i = 1, 2, \ldots, l\}$ is a subgroup of H_l.

Define an action of G_l on H_l as follows:

Definition. For $\sigma \in \mathsf{G}_l$, and $\mathbf{h} \in \mathsf{H}_l$,

$$\sigma(\mathbf{h})[i] = \mathbf{h}[\sigma^{-1}(i)], \tag{A.1}$$

where $(\cdot)[i]$ is the ith element of a vector. That is, $\sigma(\mathbf{h})$ is a vector whose ith element is the $\sigma^{-1}(i)$th element of $\mathbf{h}$. Let $\mathbf{G}$ be the permutation matrix of σ. Then (A.1) is equivalent to

$$\sigma(\mathbf{h}) = \mathbf{G}\mathbf{h}. \tag{A.2}$$

The properties of this action are described in the following lemmas.

Lemma A.2. For $\sigma_1, \sigma_2 \in \mathsf{G}, \mathbf{h} \in \mathsf{H}$, and $(\sigma_1\sigma_2)(\mathbf{h}) = \sigma_1(\sigma_2(\mathbf{h}))$.

Proof. Using (A.2),

$$(\sigma_1\sigma_2)(\mathbf{h}) = (\mathbf{G}_1\mathbf{G}_2)\mathbf{h} = \mathbf{G}_1(\mathbf{G}_2\mathbf{h}) = \sigma_1(\sigma_2(\mathbf{h})).$$

Lemma A.3. For $\sigma \in \mathsf{G}, \mathbf{h}_1, \mathbf{h}_2 \in \mathsf{H}$, and $\sigma(\mathbf{h}_1 \times \mathbf{h}_2) = \sigma(\mathbf{h}_1) \times \sigma(\mathbf{h}_2)$. As a special case, $\sigma(-\mathbf{h}) = -\sigma(\mathbf{h})$, where $-\mathbf{h}$ stands for $-\mathbf{1} \times \mathbf{h}$.

Proof.

$$\begin{aligned}\sigma(\mathbf{h}_1 \times \mathbf{h}_2)[i] &= (\mathbf{h}_1 \times \mathbf{h}_2)[(\sigma^{-1}(i)] \\ &= \mathbf{h}_1[\sigma^{-1}(i)]\mathbf{h}_2[\sigma^{-1}(i)] \\ &= \sigma(\mathbf{h}_1)[i]\sigma(\mathbf{h}_2)[i].\end{aligned}$$

Lemma A.4. For $k \geq j, A_k(\mathbf{a}_j) = -\mathbf{a}_j$; for $k < j, A_k(\mathbf{a}_j) = \mathbf{a}_j$, where $\{A_k\}$ and $\{\mathbf{a}_j\}$ are defined in (1) and (3).

Proof. Using Lemma A.1, from (1), (3), and (A.2), $A_k(\mathbf{a}_j)$ is the Kronecker product of $\mathbf{Ra} = -\mathbf{a}, \mathbf{R1} = \mathbf{1}, \mathbf{Ia} = \mathbf{a}$, and $\mathbf{I1} = \mathbf{1}$. Also, by the definition of Kronecker product, it is easy to see that

$$\mathbf{B}_1 \otimes (-\mathbf{B}_2) = -(\mathbf{B}_1 \otimes \mathbf{B}_2). \tag{A.3}$$

For $k \geq j, A_k(\mathbf{a}_j)$ has exactly one $\mathbf{Ra}$ in the Kronecker product. Therefore, from (A.3), $A_k(a_j) = -a_j$. For $k < j, A_k(a_j)$ has no $\mathbf{Ra}$ in the Kronecker product, implying that $A_k(a_j) = a_j$.

The next two lemmas are essential to the construction of OLHs.

Lemma A.5. Suppose that σ is a reflexive permutation and $\mathbf{h} \in \mathsf{H}$. Let $\mathbf{P}$ be the corresponding permutation matrix of σ. If

$$\sigma(\mathbf{h}) = -\mathbf{h}, \tag{A.4}$$

then $\mathbf{e} = [1, 2, \ldots, 2^{m-1}]$ and $\mathbf{v} = \mathbf{Pe} \times \mathbf{h}$ are orthogonal; that is, $\mathbf{v}^T\mathbf{e} = 0$.

Proof. First, by the definition of permutation matrix, $P_{ij} = 1$ if and only if $\sigma(i) = j$; otherwise, $P_{ij} = 0$. Because σ is reflexive, $\sigma(i) = j_i$, and $\sigma(j_i) = i$, $\mathbf{P}$ is symmetric; that is, $\mathbf{P} = \mathbf{P}^T$. Also, because $\sigma(\sigma(i)) = i, \mathbf{P}^2 = \mathbf{I}$. Replacing $\mathbf{v}$ by $\mathbf{Pe} \times \mathbf{h}$,

$$\begin{aligned}\mathbf{v}^T\mathbf{e} &= (\mathbf{Pe} \times \mathbf{h})^T\mathbf{e} \\ &= \sum_{i=1}^{m} (\mathbf{Pe})[i]\mathbf{h}[i]\mathbf{e}[i].\end{aligned} \tag{A.5}$$

Because $\mathbf{P}^T\mathbf{P} = \mathbf{P}^2 = \mathbf{I}$,

$$\begin{aligned}\mathbf{v}^T\mathbf{e} &= (\mathbf{Pe} \times \mathbf{h})^T\mathbf{e} \\ &= (\mathbf{Pe} \times \mathbf{h})^T(\mathbf{P}^T\mathbf{P})\mathbf{e} \\ &= (\mathbf{P}^2\mathbf{e} \times \mathbf{P}h)^T\mathbf{Pe} \\ &= (\mathbf{e} \times (-\mathbf{h}))^T\mathbf{Pe} \\ &= \sum_{i=1}^{m} -\mathbf{e}[i]\mathbf{h}[i](\mathbf{Pe})[i].\end{aligned} \tag{A.6}$$

Therefore, from (A.5) and (A.6), $\mathbf{v}^T\mathbf{e} = -\mathbf{v}^T\mathbf{e}$. Hence $\mathbf{v}^T\mathbf{e} = 0$.

Lemma A.6. Suppose that $\sigma_1, \sigma_2 \in \mathsf{G}$ are reflexive and commutative; $\mathbf{h}_1, \mathbf{h}_2 \in \mathsf{H}$. If

$$(\sigma_1\sigma_2)(\mathbf{h}_1 \times \mathbf{h}_2) = -(\mathbf{h}_1 \times \mathbf{h}_2), \tag{A.7}$$

then $\mathbf{t_1} = \mathbf{P}_1\mathbf{e} \times \mathbf{h}_1$ and $\mathbf{t_2} = \mathbf{P_2e} \times \mathbf{h}_2$ are orthogonal, where $\mathbf{P}_1$ and $\mathbf{P}_2$ are the corresponding permutation matrices of σ_1 and σ_2.

Proof. A similar approach as in the proof of Lemma A.5 is used. First,

$$\begin{aligned}\mathbf{t_1^T t_2} &= (\mathbf{P}_1\mathbf{e} \times \mathbf{h}_1)^T(\mathbf{P}_2\mathbf{e} \times \mathbf{h}_2) \\ &= \sum_{i=1}^{m} (\mathbf{P}_1\mathbf{e})[i]\mathbf{h}_1[i](\mathbf{P}_2\mathbf{e})[i]\mathbf{h}_2[i].\end{aligned} \tag{A.8}$$

Because σ_1 and σ_2 are commutative (i.e., $\mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_2\mathbf{P}_1$),

$$\begin{aligned}\mathbf{t_1^T t_2} &= (\mathbf{P}_1\mathbf{e} \times \mathbf{h}_1)^T(\mathbf{P}_2\mathbf{e} \times \mathbf{h}_2) \\ &= (\mathbf{P}_1\mathbf{e} \times \mathbf{h}_1)^T(\mathbf{P}_1\mathbf{P}_2)^T(\mathbf{P}_1\mathbf{P}_2)(\mathbf{P}_2\mathbf{e} \times \mathbf{h}_2) \\ &= (\mathbf{P}_2\mathbf{e} \times (\mathbf{P}_1\mathbf{P}_2)\mathbf{h}_1)^T(\mathbf{P}_1\mathbf{e} \times (\mathbf{P}_1\mathbf{P}_2)\mathbf{h}_2) \\ &= \sum_{i=1}^{m} (\mathbf{P}_2\mathbf{e})[i](\mathbf{P}_1\mathbf{e})[i](\mathbf{P}_1\mathbf{P}_2(\mathbf{h}_1 \times \mathbf{h}_2))[i] \\ &= \sum_{i=1}^{m} (\mathbf{P}_2\mathbf{e})[i](\mathbf{P}_1\mathbf{e})[i](-\mathbf{h}_1 \times \mathbf{h}_2)[i] \\ &= -\sum_{i=1}^{m} (\mathbf{P}_2\mathbf{e})[i](\mathbf{P}_1\mathbf{e})[i]\mathbf{h}_1[i]\mathbf{h}_2[i].\end{aligned} \tag{A.9}$$

Again, combining (A.8) and (A.9), $\mathbf{t_1^T t_2} = 0$.

Now I prove the theorem for $n = 2^m + 1, m \geq 2$. Extension to the case $n = 2^m$ is obvious. Because the bottom half is the mirror image of the top half, it is sufficient to prove that the columns in $\mathbf{T}$, the top half matrix, as constructed in (4) are orthogonal with each other.

As discussed earlier, each column $\mathbf{t}_i$ of $\mathbf{T}$ corresponds to a column g_i in the $\mathbf{M}$ matrix, which is a permutation of $\mathbf{e}$, and a column $\mathbf{h}_i$ in the $\mathbf{S}$ matrix, which is an element in H_{m-1}. Let $\mathbf{t}_i = \mathbf{g}_i \times \mathbf{h}_i$ be the elementwise product of $\mathbf{g}_i$ and $\mathbf{h}_i$. What remains to be proven is that $\{(\mathbf{g}_i, \mathbf{h}_i)\}$ given in (4) satisfy the conditions in Lemmas A.5 and A.6. From Theorems 1 and 2, G is commutative and its elements are reflexive. The proof that $\{(\mathbf{g}_i, \mathbf{h}_i)\}$ satisfy (A.4) or (A.7) is divided into five cases.

Case 1. $\mathbf{g}_1 = \mathbf{e}, \mathbf{g}_2 = \mathbf{A_j}, 1 \le j \le m-1$. The corresponding elements in H are $\mathbf{h}_1 = \mathbf{1}$ and $\mathbf{h}_2 = \mathbf{a}_j$. From Lemma 2.4, $A_j(\mathbf{a}_j) = -\mathbf{a}_j$.

Case 2. $g_1 = \mathbf{e}, g_2 = \mathbf{A_j A_{m-1}}, 1 \le j < m-2$. The corresponding elements in H are $h_1 = \mathbf{1}$ and $\mathbf{h}_2 = \mathbf{a}_1\mathbf{a}_{j+1}$. Hence, by Lemmas A.3, A.4, and A.5,

$$\begin{aligned}
&(A_jA_{m-1})(\mathbf{a}_1\mathbf{a}_{j+1}) = (A_jA_{m-1})(\mathbf{a}_1\mathbf{a}_{j+1})\\
&\quad = A_j(A_{m-1}(\mathbf{a}_1\mathbf{a}_{j+1})) = A_j(A_{m-1}(\mathbf{a}_1)A_{m-1}(\mathbf{a}_{j+1}))\\
&\quad = A_j(-\mathbf{a}_1 \times -\mathbf{a}_{j+1}) = A_j(\mathbf{a}_1\mathbf{a}_{j+1})\\
&\quad = A_j(\mathbf{a}_1)A_j(\mathbf{a}_{j+1})\\
&\quad = -\mathbf{a}_1\mathbf{a}_{j+1}.
\end{aligned}$$

Case 3. $\mathbf{g}_1 = \mathbf{A_i}, \mathbf{g}_2 = \mathbf{A_j}, i \ne j$. The corresponding elements in H are $\mathbf{h}_1 = \mathbf{a}_i$, and $\mathbf{h}_2 = \mathbf{a}_j$. From Lemmas A.3, A.4, and A.5,

$$\begin{aligned}
&(A_iA_j)(\mathbf{a}_i\mathbf{a}_j) = (A_iA_j)(\mathbf{a}_i\mathbf{a}_j)\\
&\quad = A_iA_j(\mathbf{a}_i)A_iA_j(\mathbf{a}_j) = A_jA_i(\mathbf{a}_i)A_iA_j(\mathbf{a}_j)\\
&\quad = A_j(-\mathbf{a}_i)A_i(-\mathbf{a}_j) = A_i(\mathbf{a}_j)A_j(\mathbf{a}_i)\\
&\quad = -\mathbf{a}_i\mathbf{a}_j.
\end{aligned}$$

Case 4. $\mathbf{g}_1 = \mathbf{A_i A_{m-1}}, \mathbf{g}_2 = \mathbf{A_j A_{m-1}}, 1 \le i \ne j < m-2$. The corresponding elements in H are $\mathbf{h}_1 = \mathbf{a}_1\mathbf{a}_{i+1}$ and $\mathbf{h}_2 = \mathbf{a}_1\mathbf{a}_{j+1}$. From Lemmas A.3, A.4, and A.5,

$$\begin{aligned}
&((A_iA_{m-1})A_jA_{m-1})(\mathbf{a}_1\mathbf{a}_{i+1}\mathbf{a}_1\mathbf{a}_{j+1}) = (A_iA_j)(\mathbf{a}_{i+1}\mathbf{a}_{j+1})\\
&\quad = A_iA_j(\mathbf{a}_{i+1})A_iA_j(\mathbf{a}_{j+1}) = A_j(\mathbf{a}_{i+1})A_i(\mathbf{a}_{j+1})\\
&\quad = -\mathbf{a}_{i+1}\mathbf{a}_{j+1} = -(\mathbf{a}_1\mathbf{a}_{i+1}\mathbf{a}_1\mathbf{a}_{j+1}).
\end{aligned}$$

Case 5. $\mathbf{g}_1 = \mathbf{A_i}, \mathbf{g}_2 = \mathbf{A_j A_{m-1}}, 1 \le i \le m-1, 1 \le j < m-2$. The corresponding elements in H are $h_1 = a_i$ and $h_2 = a_1a_{j+1}$. From Lemmas A.3, A.4, and A.5,

$$\begin{aligned}
&(A_iA_jA_{m-1})(\mathbf{a}_i\mathbf{a}_1\mathbf{a}_{j+1})\\
&\quad = (A_iA_jA_{m-1})(\mathbf{a}_i\mathbf{a}_{j+1})(A_iA_jA_{m-1})(\mathbf{a}_1)\\
&\quad = -\mathbf{a}_1(A_iA_jA_{m-1})(\mathbf{a}_i\mathbf{a}_{j+1})\\
&\quad = -\mathbf{a}_1(A_iA_j)(A_{m-1}(\mathbf{a}_i)A_{m-1}(\mathbf{a}_{j+1}))\\
&\quad = -\mathbf{a}_1(A_iA_j)(\mathbf{a}_i\mathbf{a}_j)\\
&\quad = -\mathbf{a}_1(A_iA_j)(\mathbf{a}_i)(A_iA_j)(\mathbf{a}_{j+1})\\
&\quad = -\mathbf{a}_1A_j(\mathbf{a}_i)A_i(\mathbf{a}_{j+1})\\
&\quad = -\mathbf{a}_1\mathbf{a}_{j+1}\mathbf{a}_i.
\end{aligned}$$

Proof of Theorem 7

Orthogonality between any pair of columns follows, because the proof of Lemmas A.5 and A.6 still hold if one replaces $\mathbf{e}$ by its permutation $\mathbf{w}$. Properties as stated in Theorems 4 and 5 follows from the symmetry of the design.

[Received May 1997. Revised June 1998.]

REFERENCES

Currin, C., Mitchell, T., Morris, D., and Ylvisaker, D. (1991), "Bayesian Prediction of Deterministic Functions, With Applications to the Design and Analysis of Computer Experiments," *Journal of the American Statistical Association*, 86, 953–963.

Fang, K. T., and Wang, Y. (1994), *Number-Theoretic Methods in Statistics*, London: Chapman and Hall.

Hungerford, T. W. (1974), *Algebra*, New York: Springer-Verlag.

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.

Johnson, M., Moore, L., and Ylvisaker, D. (1990), "Minimax and Maximin Distance Designs," *Journal of Statistical Planning and Inference*, 26, 131–148.

McKay, M., Beckman, R., and Conover, W. (1979), "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code," *Technometrics*, 21, 239–246.

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

Owen, A. B. (1992), "Orthogonal Arrays for Computer Experiments, Integration and Visualization," *Statistica Sinica*, 2, 439–452.

——— (1994), "Controlling Correlations in Latin Hypercube Samples," *Journal of the American Statistical Association*, 89, 1517–1522.

Park, J.-S. (1994), "Optimal Latin-Hypercube Designs for Computer Experiments," *Journal of Statistical Planning and Inference*, 39, 95–111.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications* (2nd ed.), New York: Wiley.

Sacks, J., Schiller, S., and Welch, W. (1989), "Designs for Computer Experiments," *Technometrics*, 31, 41–47.

Sacks, J., Welch, W., Mitchell, T., and Wynn, P. (1989), "Design and Analysis of Computer Experiments," *Statistical Science*, 4, 409–435.

Schonlau, M., Hamada, M., and Welch, W. (1996), "Identifying Parametric Nonlinear Models for Computer Codes," I.I.Q.P. Research Report, University of Waterloo.

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

——— (1998), "Selecting Latin Hypercubes Using Correlation Criteria," *Statistica Sinica*, 8, 965–977.