On Reduction of Finite Sample Variance by Extended Latin Hypercube Sampling

Nobuaki Hoshino
Graduate School of Economics, University of Tokyo
and
Akimichi Takemura
Faculty of Economics, University of Tokyo

January, 1997

Abstract

McKay, Conover and Beckman (1979) introduced Latin hypercube sampling (LHS) for reducing variance of Monte Carlo simulations. More recently Owen (1992a) and Tang (1993) generalized LHS using orthogonal arrays. In the Owen's class of generalized LHS, we define extended Latin hypercube sampling of strength m (henceforth denoted as ELHS(m)), such that ELHS(1) reduces to LHS. We first derive explicit formula for the finite sample variance of ELHS(m) by detailed investigation of combinatorics involved in ELHS(m).

Based on this formula, we give a sufficient condition for variance reduction by ELHS(m), generalizing similar result of McKay, Conover and Beckman (1979) for $m = 1$. Actually our sufficient condition for $m = 1$ contains the sufficient condition by McKay, Conover and Beckman (1979) and thus strengthens their result.

1 INTRODUCTION

Monte Carlo simulation is often used to evaluate the expectation of a statistic $W = g(X_1, \ldots, X_K)$, which is not analytically tractable. In usual Monte Carlo simulations simple random sampling (SRS) is used to generate sample points. SRS is widely applicable because of its simplicity. However its sampling variance is often large and many replications are needed to achieve desired precision. Therefore methods for reducing sampling variance of SRS are of great importance.

One way of reducing variance of SRS is to scatter the sample points more uniformly over the sample space than SRS. This is the basic idea of various techniques known as quasi-Monte Carlo methods. See Niederreiter (1992) for a review. Uniformity can be achieved by stratifying the sample space. LHS can be interpreted as a method for stratifying each univariate margin simultaneously. A natural extension is to stratify each m-variate margins simultaneously, which can be achieved by the sampling design based on orthogonal arrays. Owen (1992a) and independently Tang (1993) proposed orthogonal array (OA) based sampling. They showed that OA based sampling can improve LHS substantially. We will define a class of extended Latin hypercube sampling ELHS(m) of strength m, $m \leq K$, such that ELHS(1) is equivalent to LHS.

The reduction of variance by LHS and ELHS is closely related to ANOVA (Analysis of Variance) decomposition of the statistic W. Stein (1987) showed that LHS asymptotically filters out main effects of W. Therefore LHS asymptotically achieves variance reduction for any statistic W. Similarly, OA based sampling asymptotically filters some higher order interactions as well, and hence asymptotically achieves further variance reduction for any statistic. However for the finite sample case, LHS and its generalizations do not necessarily lead to variance reduction due to combinatorial complications.

In order to investigate reduction of finite sample variance we first derive explicit expression of the finite sample variance of ELHS(m) in terms of the ANOVA decomposition of the finite sample cell mean function of $W = g(X_1, \ldots, X_K)$. Although Owen (1992a) describes finite sample variance of OA based sampling with the aid of results by Patterson (1954), Patterson's results are stated without proof. Using the expression of finite sample variance of ELHS(m) we derive a sufficient condition for the reduction of the variance of ELHS(m) over SRS. Our sufficient condition is given in terms of m-variate monotonicity of the statistic $g(X_1, \ldots, X_K)$. For the case of $m = 1$, our sufficient condition requires that g is monotone in any $K - 1$ variables out of K variables $X_1, \ldots, X_K$. The sufficient condition given by McKay, Conover and Beckman (1979) requires that g is monotone in each $X_i, i = 1, \ldots, K$. Thus our result for $m = 1$ strengthens the result of McKay, Conover and Beckman (1979).

The organization of the paper is as follows. In Section 2 we define ELHS(m) and introduce appropriate notational conventions. In Section 3 we derive explicit expression for the finite sample variance of ELHS(m). Based on this expression, we give a sufficient condition for reduction of finite sample variance of ELHS(m) over SRS in Section 4. In Section 5 we perform some numerical simulations to confirm our theoretical results. In Section 6 we make some additional comments on LHS and OA based sampling.

2 CONSTRUCTION OF EXTENDED LATIN HYPERCUBE SAMPLING

We consider evaluation of the expectation of a statistic $W = g(X_1, \ldots, X_K)$:

$$\mu = E(W) = E[g(X_1, \ldots, X_K)],$$

where $(X_1, \ldots, X_K) \in R^K$. We assume that $X_1, \ldots, X_K$ are independent continuous random variables with known distribution functions F_j with the density function f_j, $j = 1, \ldots, K$. $\boldsymbol{X} = (X_1, \ldots, X_K)$ has the joint distribution function $F = F_1 \cdots F_K$ with the joint density function $f = f_1 \cdots f_K$.

Suppose that the evaluation of this expectation is not analytically tractable and we use Monte Carlo methods. In our extended Latin hypercube sampling defined below, the sample points are generated in two steps. For the first step we partition the sample space R^K into N^K cells of equal probability $1/N^K$ and choose N^m, $m \leq K$, cells out of N^K cells using random orthogonal array. For the second step the actual sample points are generated according to the conditional distribution on the chosen cells.

We now describe the first step. A $\lambda N^m \times K$ matrix D, with elements taken from a set of N symbols, is called an orthogonal array of strength m ($m \leq K$), size λN^m, K constraints, N levels, frequency λ, if in any $\lambda N^m \times m$ submatrix of D each of the all possible $1 \times m$ row vectors occurs the same number λ of times. Such an array is denoted by OA($\lambda N^m, K, N, m$). Without essential loss of generality, we only consider the case $\lambda = 1$ as in the original LHS of McKay, Conover and Beckman (1979).

Orthogonal array is a natural generalization of orthogonal Latin squares. Plackett and Burman (1946) generate orthogonal arrays of strength 2 by combining mutually orthogonal Latin squares. Rao (1947) formulates the concept of orthogonal arrays in general form and gives the lower bound of N for fixed m, K. Bose and Bush (1952) give sufficient condition for existence of orthogonal arrays, since they do not always exist.

Lemma 1 *Let N be a prime or prime power. An orthogonal array $OA(N^m, K, N, m)$ exists if there exists a $K \times m$ matrix of elements from the Galois field $GF(N)$ such that every $m \times m$ submatrix is of rank m.*

Under Lemma 1 we arbitrarily choose and fix an orthogonal array and denote it by D_0. The ij element of D_0 is denoted by d_{ij}. We call D_0 a *generator array*. Actually we will show that the selection of D_0 does not affect the variance of ELHS(m). We generate random orthogonal array by random permutations of elements of D_0. For simplicity of notation we take the set of N symbols as

$$Z_N = \{1, 2, 3, \ldots, N\}.$$

Let $\mathcal{S}_N$ denote the symmetric group of Z_N, i.e. the set of permutations of $\{1, \ldots, N\}$. Let $\pi \in \mathcal{S}_N$, then $(\pi(1), \ldots, \pi(N))$ is a particular permutation of $(1, \ldots, n)$. $\mathcal{S}_N$ has $N!$ elements. Consider the uniform distribution on $\mathcal{S}_N$, where each permutation of $\mathcal{S}_N$ has the equal probability $1/N!$. We choose K permutations $\pi_j, j = 1, \ldots, K$, independently and uniformly from $\mathcal{S}_N$ and we apply $\pi_j \in S_N$ to the jth column of D_0, $j = 1, \ldots, K$. The ij element of the resulting array is $\pi_j(d_{ij})$.

In addition to the above randomization of the elements of each column of D_0 we also consider permutation of the rows of D_0 as in Tang (1993). This randomization is needed only for the sake of clear argument and can be omitted in practice. Note that the above columnwise randomization does not guarantee the exchangeability of the rows of the resulting array. This lack of exchangeability of the rows can be overcome by randomly permuting the rows of the array D_0. Let $\tilde{\pi} \in \mathcal{S}_{N^m}$ be a uniform random permutation of $1, 2, \ldots, N^m$. $\tilde{\pi}$ is chosen independently of $\pi_1, \ldots, \pi_K$. Let the ij element of $\Pi(D_0)$ be defined by $\pi_j(d_{\tilde{\pi}(i)j})$.

Combining random permutations of the elements of each column of D_0 and the rows of D_0, we generate an orthogonal array $\Pi(D_0)$ stochastically. This constitutes the first step of generating the sample points.

When $\pi_j, j = 1, \ldots, K$, and $\tilde{\pi}$ are given we denote

$$z_{ij} = \pi_j(d_{\tilde{\pi}(i)j}), \qquad \boldsymbol{z}_i = (z_{i1}, z_{i2}, \ldots, z_{iK}) \in Z_N^K.$$

$\boldsymbol{z}_i$ is the ith row of $\Pi(D_0)$. Let $F_j^{-1}(u) = \inf\{x | F_j(x) \geq u\}$ be the quantile function and let

$$P_j^{(z_{ij})} = F_j^{-1}\left(\left(\frac{z_{ij} - 1}{N}, \frac{z_{ij}}{N}\right]\right) \subset R, \quad j = 1, \ldots, K.$$

Let $\boldsymbol{z}_i$ correspond to the following cell in the sample space R^K:

$$P_1^{(z_{i1})} \times \cdots \times P_K^{(z_{iK})}.$$

Note that these cells have the same probability $1/N^K$ under the joint distribution F of $(X_1, \ldots, X_K)$.

For the second step we generate random vector $(X_{i1}, \ldots, X_{iK})$ in the cell $P_1^{(z_{i1})} \times \cdots \times P_K^{(z_{iK})}$ according to the conditional distribution of F on the cell. Let $0 < U_{ij} \leq$

$1, i = 1, \ldots, N^m, j = 1, \ldots, K$, be independent uniform random variables. Then X_{ij} can be generated as

$$X_{ij} = F_j^{-1}\{(z_{ij} - U_{ij})/N\}.$$

We now define our estimator T_{EL} of μ based on ELHS(m) by

$$T_{EL} = \frac{1}{N^m} \sum_{i=1}^{N^m} g(X_{i1}, \ldots, X_{iK}).$$

Obviously T_{EL} is an unbiased estimator of μ. Note that T_{EL} is invariant with respect to the permutation of the rows of D_0. Therefore in practice we do not need the added permutation $\tilde{\pi}$ of the rows of D_0. Furthermore we denote the usual estimator of μ based on SRS by

$$T_R = \bar{W}.$$

We are interested in the comparison of variances of T_{EL} and T_R.

3 FINITE SAMPLE VARIANCE OF ELHS

3.1 ANOVA decomposition of the cell mean function

In order to investigate the variance of T_{EL} we introduce the *cell mean function* and its ANOVA decomposition. The cell $P_1^{(z_1)} \times \cdots \times P_K^{(z_K)}$ is indexed by $\boldsymbol{z} = (z_1, z_2, \ldots, z_K) \in Z_N^K$. Abusing the notation we simply denote the cell as $\boldsymbol{z} = (z_1, z_2, \ldots, z_K) \in Z_N^K$. Suppose that a sample point $\boldsymbol{X}$ is obtained from a cell $\boldsymbol{z}$. We call the conditional expectation $E[g(\boldsymbol{X})|\boldsymbol{z}]$ the cell mean function and denote it by

$$\mu_{\boldsymbol{z}} = E[g(\boldsymbol{X})|\boldsymbol{z}] = \int g(\boldsymbol{X}) \prod_{i=1}^{K} f_i^{(z_i)}(X_i) d\boldsymbol{X},$$

where

$$f_i^{(z_i)}(x) = \begin{cases} N \cdot f_i(x) & \text{if } \quad x \in P_i^{(z_i)} \\ 0 & \text{otherwise.} \end{cases}$$

Let $F_i^{(z_i)} = \int_{-\infty}^{x} f_i^{(z_i)}(u) du$. We denote the random variable having the distribution function $F_i^{(z_i)}$ by $X_i^{(z_i)}$ and the random vector having the distribution function $F_1^{(z_1)} \cdots F_K^{(z_K)}$ by $\boldsymbol{X}^{(\boldsymbol{z})}$.

We denote the usual ANOVA decomposition for the cell mean function by

$$\begin{aligned} \mu_{z_1 z_2 \ldots z_K} - \mu &= \sum_{i=1}^{K} \alpha_1(z_i) + \sum_{i_1 < i_2} \alpha_2(z_{i_1}, z_{i_2}) \\ &\quad + \sum_{i_1 < i_2 < i_3} \alpha_3(z_{i_1}, z_{i_2}, z_{i_3}) + \cdots \\ &\quad + \alpha_K(z_{i_1}, \ldots, z_{i_K}), \end{aligned}$$

where

$$\alpha_1(z_i) = \frac{1}{N^{K-1}} \sum_{z_1=1}^{N} \cdots \sum_{z_{i-1}=1}^{N} \sum_{z_{i+1}=1}^{N} \cdots \sum_{z_K=1}^{N} (\mu_{z_1 z_2 \cdots z_K} - \mu),$$

$$\alpha_2(z_{i_1}, z_{i_2}) = \frac{1}{N^{K-2}} \sum_{\substack{j=1 \\ j \neq i_1, i_2}}^{K} \sum_{z_j=1}^{N} (\mu_{z_1 z_2 \cdots z_K} - \mu - \alpha_1(z_{i_1}) - \alpha_1(z_{i_2})),$$

$$\alpha_3(z_{i_1}, z_{i_2}, z_{i_3}) = \frac{1}{N^{K-3}} \sum_{\substack{j=1 \\ j \neq i_1, i_2, i_3}}^{K} \sum_{z_j=1}^{N} (\mu_{z_1 z_2 \cdots z_K} - \mu - \alpha_1(z_{i_1}) - \alpha_1(z_{i_2}) - \alpha_1(z_{i_3}) - \alpha_2(z_{i_1}, z_{i_2}) - \alpha_2(z_{i_1}, z_{i_3}) - \alpha_2(z_{i_2}, z_{i_3})),$$

and we continue this process to α_K. Summation over z_i is always taken from 1 to N and from now on we omit the range $1 \leq z_i \leq N$ from the summation signs.

We denote the sum of squares of sth order interaction effects by

$$\varphi_s^2 = \sum_{i_1 < \cdots < i_s} \sum_{z_{i_1}} \cdots \sum_{z_{i_s}} \alpha_s(z_{i_1}, \ldots, z_{i_s})^2 \cdot \frac{1}{N^s}.$$

Then $\mathrm{Var}(T_R)$ can be written as

$$\begin{aligned} \mathrm{Var}(T_R) &= \frac{1}{N^m} \{\mathrm{Var}(E[\boldsymbol{X}^{(\boldsymbol{z})}]) + E[\mathrm{Var}(\boldsymbol{X}^{(\boldsymbol{z})})]\} \\ &= \frac{1}{N^m} \sum_{s=1}^{K} \varphi_s^2 + \nabla_r, \end{aligned}$$

where

$$\nabla_r = \frac{1}{N^m} E_{\boldsymbol{z}}[\mathrm{Var}(\boldsymbol{X}^{(\boldsymbol{z})})].$$

See Section 5.2 of Serfling (1980).

3.2 The Variance of Estimators Using ELHS

In this subsection we derive the following expression of the variance of T_{EL}.

Theorem 1

$$\mathrm{Var}(T_{EL}) = \frac{1}{N^m} \sum_{s=m+1}^{K} \varphi_s^2 \{1 - (1-N)^{1-s} \cdot \sum_{u=0}^{m-1} (-N)^u \binom{s-1}{u}\} + \nabla_r. \qquad (1)$$

Concerning the binomial coefficient, from now on we use the notation

$$\binom{a}{b} = 0 \qquad \text{if } a < b \text{ or } b < 0, \qquad (2)$$

where a, b are integers. Then for $1 \leq s \leq m$

$$\sum_{u=0}^{m-1} (-N)^u \binom{s-1}{u} = \sum_{u=0}^{s-1} (-N)^u \binom{s-1}{u} = (1-N)^{s-1}.$$

Therefore (1) can alternatively be written as

$$\mathrm{Var}(T_{EL}) = \frac{1}{N^m} \sum_{s=1}^{K} \varphi_s^2 \{1 - (1-N)^{1-s} \cdot \sum_{u=0}^{m-1} (-N)^u \binom{s-1}{u}\} + \nabla_r. \qquad (3)$$

We will prove Theorem 1 in this form.

Note that in (3) interaction effects up to the order m, i.e. $\varphi_s, s = 1, \ldots, m$, are canceled out. This has to be the case for ELHS(m). In ELHS(m) we stratify m-variate margin, such that for $s \le m$ all elements of Z_N^s appear in each combination of s axes equal number of times. In view of $\sum_{z_{i_1}} \cdots \sum_{z_{i_s}} \alpha_s(z_{i_1}, \ldots, z_{i_s}) = 0$, interaction effects of cell mean function up to the order m vanish for each realization of T_{EL}.

The rest of this section is devoted to the proof of Theorem 1. Our proof is given in the form of Lemma 2 through Lemma 5.

If the cell selections are independent, the two step generation process of the last subsection amounts to SRS. Therefore the difference in SRS and ELHS comes from the restriction on selection of the cells. Let

$$(\boldsymbol{y}, \boldsymbol{z}) \in Z_N^K \times Z_N^K$$

be two cell indices corresponding to two rows of the random orthogonal array $\Pi(D_0)$. For definiteness we let $\boldsymbol{y}$ be the first row and $\boldsymbol{z}$ be the second row of $\Pi(D_0)$. Because of the restriction imposed by the property of the orthogonal array, $\boldsymbol{y}$ and $\boldsymbol{z}$ are not independent.

Using the exchangeability of the rows of $\Pi(D_0)$ we have

$$\begin{aligned}
\mathrm{Var}(T_{EL}) &= \frac{1}{N^{2m}} \mathrm{Var}(W_1 + \cdots + W_{N^m}) \\
&= \frac{1}{N^{2m}} N^m \mathrm{Var}(W_1) + \frac{1}{N^{2m}} \binom{N^m}{2} 2 \,\mathrm{Cov}(W_1, W_2) \\
&= \frac{1}{N^m} \mathrm{Var}(W_1) + \frac{N^m - 1}{N^m} \mathrm{Cov}(W_1, W_2) \\
&= \mathrm{Var}(T_R) + \frac{N^m - 1}{N^m} \cdot E_{\boldsymbol{y},\boldsymbol{z}}[E(W_1 W_2 | \boldsymbol{y}, \boldsymbol{z}) - E(W_1 | \boldsymbol{y}, \boldsymbol{z}) E(W_2 | \boldsymbol{y}, \boldsymbol{z})] \\
&= \mathrm{Var}(T_R) + \frac{N^m - 1}{N^m} \cdot (E[\mu_{\boldsymbol{y}} \mu_{\boldsymbol{z}}] - E[\mu_{\boldsymbol{y}}] E[\mu_{\boldsymbol{z}}]) \\
&= \mathrm{Var}(T_R) + \frac{N^m - 1}{N^m} \cdot (E[\mu_{\boldsymbol{y}} \mu_{\boldsymbol{z}}] - \mu^2) \\
&= \mathrm{Var}(T_R) + \frac{N^m - 1}{N^m} \cdot \mathrm{Cov}_m(\mu_{\boldsymbol{y}}, \mu_{\boldsymbol{z}}), \qquad (4)
\end{aligned}$$

where $\mathrm{Cov}_m(\mu_{\boldsymbol{y}}, \mu_{\boldsymbol{z}}) = E[(\mu_{\boldsymbol{y}} - \mu)(\mu_{\boldsymbol{z}} - \mu)]$ is the covariance of two cell mean functions under the first step of ELHS(m).

Let

$$\#\{\boldsymbol{y}, \boldsymbol{z}\} = \sum_{i=1}^{K} \delta_{y_i z_i}$$

denote the number of overlapping indices of $\boldsymbol{y}$ and $\boldsymbol{z}$, where

$$\delta_{y_i z_i} = \begin{cases} 1, & y_i = z_i, \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$w(h) = P(\#\{\boldsymbol{y}, \boldsymbol{z}\} = h)$$

be the probability that $\boldsymbol{y}$ and $\boldsymbol{z}$ have h common indices under the first step of ELHS(m). The probabilities $w(h)$ is given in the following lemma.

Lemma 2 *For* $0 \leq h \leq m-1$

$$w(h) = \frac{\binom{K}{h} \sum_{t=0}^{m-h-1} (-1)^t (N^{m-h-t} - 1) \binom{K-h}{t}}{N^m - 1}. \tag{5}$$

$w(h) = 0$ *for* $m \leq h \leq K$.

Proof We first show that $w(h) = 0$ for $m \leq h \leq K$. Note that in an orthogonal array $\mathrm{OA}(N^m, K, N, m)$, the rows of of each $N^m \times m$ submatrix are all distinct. This implies two rows of $\mathrm{OA}(N^m, K, N, m)$ can only have less than m elements in common. Therefore $w(h) = 0$ for $m \leq h \leq K$.

Now let $h < m$. Fix the first row as $\boldsymbol{y} = (1, 1, \ldots, 1)$ and consider the conditional probability

$$P(\#\{\boldsymbol{y}, \boldsymbol{z}\} = h \mid \boldsymbol{y} = (1, 1, \ldots, 1)).$$

Let D be a particular orthogonal array whose first row is $\boldsymbol{y} = (1, 1, \ldots, 1)$. With the remaining $N^m - 1$ rows of D, we want to know the number of rows that have just h axes indexed as 1. Let us count the number of rows $\boldsymbol{z}$ of D such that $z_1 = z_2 = \cdots = z_h = 1$. There are $N^{m-h} - 1$ rows of this form. Using an inclusion-exclusion formula, we can count the number of rows that satisfy $z_1 = \cdots = z_h = 1, z_{h+1} \neq 1, \ldots, z_K \neq 1$ among these. Note that the first row (1,1,...,1) has to be subtracted from the count. Therefore the number is

$$\begin{aligned}
&(N^{m-h} - 1) - \binom{K-h}{1} (N^{m-h-1} - 1) + \cdots + (-1)^{m-h-1} \binom{K-h}{m-h-1} (N-1) \\
&= \sum_{t=0}^{m-h-1} (-1)^t (N^{m-h-t} - 1) \binom{K-h}{t}.
\end{aligned} \tag{6}$$

This times $\binom{K}{h}$ yields the total number of rows that have just h axes indexed as 1. Note that the number in (6) is the same for all orthogonal arrays $\mathrm{OA}(N^m, K, N, m)$ whose first row is $\boldsymbol{y} = (1, 1, \ldots, 1)$.

Because of the random permutation $\tilde{\pi}$ of the rows of D_0, $\boldsymbol{z}$ is equally likely to be any of the remaining $N^m - 1$ rows. Therefore this conditional probability is expressed

$$P(\#\{\boldsymbol{y}, \boldsymbol{z}\} = h \mid \boldsymbol{y} = (1, 1, \ldots, 1)) = \frac{\binom{K}{h} \sum_{t=0}^{m-h-1} (-1)^t (N^{m-h-t} - 1) \binom{K-h}{t}}{N^m - 1}.$$

Now this conditional probability does not depend on the fact that we have fixed $\boldsymbol{y} = (1, 1, \ldots, 1)$. Any other particular value of the vector $\boldsymbol{y}$ leads to the same conditional probability. Therefore the conditional probability is equal to the unconditional probability $w(h)$. Q.E.D.

Define

$$Q(h) = \{(\boldsymbol{y}, \boldsymbol{z}) \mid \#\{\boldsymbol{y}, \boldsymbol{z}\} = h\} \subset Z_N^K \times Z_N^K.$$

$w(h)$ of (5) gives the probability $P((\boldsymbol{y}, \boldsymbol{z}) \in Q(h)) = P(Q(h))$. The number of elements belonging to $Q(h)$ is

$$|Q(h)| = N^K \binom{K}{h} (N-1)^{K-h}. \tag{7}$$

It is obvious that the conditional distribution of $(\boldsymbol{y}, \boldsymbol{z})$ given $(\boldsymbol{y}, \boldsymbol{z}) \in Q(h)$ is uniform over $Q(h)$. Furthermore by symmetry $E(\mu_{\boldsymbol{y}}|Q(h)) = E(\mu_{\boldsymbol{z}}|Q(h)) = \mu$. Therefore

$$\mathrm{Cov}(\ \mu_{\boldsymbol{y}}, \mu_{\boldsymbol{z}} \mid Q(h)\ ) = \frac{1}{|Q(h)|} \sum_{Q(h)} (\mu_{\boldsymbol{y}} - \mu)(\mu_{\boldsymbol{z}} - \mu).$$

This conditional covariance of $\mu_{\boldsymbol{y}}$ and $\mu_{\boldsymbol{z}}$ is evaluated as follows.

Lemma 3

$$E[\ (\mu_{\boldsymbol{y}} - \mu)(\mu_{\boldsymbol{z}} - \mu) \mid Q(h)\ ] = \sum_{s=1}^{K} \varphi_s^2 \cdot c(s,h) \cdot \frac{1}{N^K \binom{K}{h} (N-1)^{K-h}},$$

where

$$c(s,h) = \sum_{j=0}^{K-h} \binom{h+j}{h} \binom{K-s}{h-s+j} N^{2K-h-j} (-1)^j.$$

Proof For $l = 1, \ldots, K$ we denote

$$\sum_{(l)} = \sum_{i_1 < \cdots < i_l} \sum_{\{(\boldsymbol{z},\boldsymbol{y}) | z_{i_1} = y_{i_1}, \ldots, z_{i_l} = y_{i_l}\}},$$

where $\boldsymbol{z} = (z_1, z_2, \ldots, z_K)$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_K)$. We calculate $\sum_{(l)} (\mu_{\boldsymbol{z}} - \mu)(\mu_{\boldsymbol{y}} - \mu)$. First suppose that $z_1 = y_1, \ldots, z_l = y_l$, then

$$\begin{aligned}
&\sum_{z_1} \cdots \sum_{z_K} \sum_{y_1 = z_1} \cdots \sum_{y_l = z_l} \sum_{y_{l+1}} \cdots \sum_{y_K} (\mu_{\boldsymbol{z}} - \mu)(\mu_{\boldsymbol{y}} - \mu) \\
&= \sum_{z_1} \cdots \sum_{z_l} \alpha_l(z_1, \ldots, z_l)^2 N^{2K-2l} \\
&\quad + \sum_{1 \le i_1 < \cdots < i_{l-1} \le l} \sum_{z_{i_1}} \cdots \sum_{z_{i_{l-1}}} \alpha_{l-1}(z_{i_1}, .., z_{i_{l-1}})^2 N^{2K-2l+1} \\
&\quad + \cdots \\
&\quad + \sum_{1 \le i_1 < i_2 \le l} \sum_{z_{i_1}} \sum_{z_{i_2}} \alpha_2(z_{i_1}, z_{i_2})^2 N^{2K-l-2} \\
&\quad + \sum_{i_1=1}^{l} \sum_{z_{i_1}} \alpha_1(z_{i_1})^2 N^{2K-l-1}.
\end{aligned} \tag{8}$$

Same argument applies for the case $z_{i_1} = y_{i_1}, \ldots, z_{i_l} = y_{i_l}$. There are $\binom{K}{l}$ ways of choosing $1 \le i_1 < \cdots < i_l \le K$. Consider the sum over these cases:

$$\sum_{i_1 < \cdots < i_l} \sum_{z_1} \cdots \sum_{z_K} \sum_{y_{i_1} = z_{i_1}} \cdots \sum_{y_{i_l} = z_{i_l}} \sum_{j \ne i_1, \ldots, i_l} \sum_{y_j} (\mu_{\boldsymbol{z}} - \mu)(\mu_{\boldsymbol{y}} - \mu).$$

In each case, (8) shows that the tth order interaction terms appear $\binom{l}{t}$ times. By symmetry the coefficient of $\{\sum_{i_1 < \cdots < i_t} \sum_{z_{i_1}} \cdots \sum_{z_{i_t}} \alpha_t(z_{i_1}, \ldots, z_{i_t})^2 N^{2K-l-t}\}$ is

$$\frac{\binom{K}{l} \cdot \binom{l}{t}}{\binom{K}{t}} = \binom{K-t}{l-t}.$$

Therefore

$$\begin{aligned}
&\sum_{(l)}(\mu_{\boldsymbol{z}} - \mu)(\mu_{\boldsymbol{y}} - \mu) \\
= & \sum_{i_1<\cdots<i_l} \sum_{z_{i_1}} \cdots \sum_{z_{i_l}} \alpha_l(z_{i_1}, \ldots, z_{i_l})^2 N^{2K-2l} \\
& + \sum_{i_1<\cdots<i_{l-1}} \sum_{z_{i_1}} \cdots \sum_{z_{i_{l-1}}} \alpha_{l-1}(z_{i_1}, .., z_{i_{l-1}})^2 N^{2K-2l+1} \binom{K-l+1}{1} \\
& + \cdots \\
& + \sum_{i_1<i_2} \sum_{z_{i_1}} \sum_{z_{i_2}} \alpha_2(z_{i_1}, z_{i_2})^2 N^{2K-l-2} \binom{K-2}{l-2} \\
& + \sum_{i_1} \sum_{z_{i_1}} \alpha_1(z_{i_1})^2 N^{2K-l-1} \binom{K-1}{l-1} \\
= & N^{2K-l} \sum_{s=1}^{l} \varphi_s^2 \binom{K-s}{l-s}. \qquad (9)
\end{aligned}$$

Now we utilize an extended form of inclusion-exclusion principle. By Section 1.2 of Galambos and Simonelli (1996)

$$\sum_{Q(h)} = \sum_{(h)} - \binom{h+1}{1} \sum_{(h+1)} + \binom{h+2}{2} \sum_{(h+2)} + \cdots + (-1)^{K-h} \binom{K}{K-h} \sum_{(K)}.$$

Then by (9)

$$\sum_{Q(h)} (\mu_{\boldsymbol{z}_1} - \mu)(\mu_{\boldsymbol{z}_2} - \mu) = \sum_{l=h}^{K} N^{2K-l} (-1)^{l-h} \binom{l}{l-h} \sum_{s=1}^{l} \varphi_s^2 \binom{K-s}{l-s}. \qquad (10)$$

Let $l = h + j$. (10) equals

$$\sum_{j=0}^{K-h} N^{2K-l} (-1)^j \binom{h+j}{j} \sum_{s=1}^{\min(K,h+j)} \varphi_s^2 \binom{K-s}{h+j-s}.$$

By the notational convention (2) this can be written as

$$\sum_{s=1}^{K} \varphi_s^2 \sum_{j=0}^{K-h} N^{2K-h-j} (-1)^j \binom{K-s}{h-s+j} \binom{h+j}{h} = \sum_{s=1}^{K} \varphi_s^2 \, c(s,h). \qquad (11)$$

By (7) the number of elements belonging to $Q(h)$ is $N^K \binom{K}{h} (N-1)^{K-h}$. Dividing (11) by this yields the conditional covariance. Q.E.D.

Combining Lemma 2 and Lemma 3

$$\mathrm{Cov}_m(\mu_{\boldsymbol{y}}, \mu_{\boldsymbol{z}}) = \sum_{h=1}^{m-1} w(h) \, \mathrm{Cov}_m(\mu_{\boldsymbol{y}}, \mu_{\boldsymbol{z}} \mid Q(h))$$

can be written as

$$\sum_{h=0}^{m-1}\sum_{t=0}^{m-h-1}\sum_{s=1}^{K}\sum_{j=0}^{K-h}\varphi_s^2\times$$
$$(\frac{N^{m-h-t}-1}{N^m-1}(-1)^{t+j}N^{K-h-j}(N-1)^{h-K}\binom{K-h}{t}\binom{h+j}{h}\binom{K-s}{h-s+j}).$$

Then by (4)

$$\begin{aligned}\mathrm{Var}(T_{EL}) &= \mathrm{Var}(T_R)+\frac{1}{N^m}(\sum_{h=0}^{m-1}\sum_{t=0}^{m-h-1}\sum_{s=1}^{K}\sum_{j=0}^{K-h}\varphi_s^2(-1)^{t+j}N^{K-h-j}\times\\ &\quad (N-1)^{h-K}(N^{m-h-t}-1)\binom{K-h}{t}\binom{h+j}{h}\binom{K-s}{h-s+j}).\end{aligned}$$

To evaluate $\mathrm{Var}(T_{EL})$, we concentrate on the coefficient of $\varphi_s^2, s=1,\ldots,K$. That is to say, our problem is to simplify

$$\sum_{h=0}^{m-1}\sum_{t=0}^{m-h-1}\sum_{j=0}^{K-h}(-1)^{t+j}N^{K-h-j}(N-1)^{h-K}(N^{m-h-t}-1)\binom{K-h}{t}\binom{h+j}{h}\binom{K-s}{h-s+j}.$$

Lemma 4

$$\begin{aligned}&\sum_{j=0}^{K-h}(-1)^{t+j}N^{K-h-j}(N-1)^{h-K}(N^{m-h-t}-1)\binom{K-h}{t}\binom{h+j}{h}\binom{K-s}{h-s+j}\\ &= \binom{K-h}{t}\frac{N^{m-t-h}-1}{N-1}(-1)^{t+s}(N-1)^{1-s}\cdot\{\sum_{u=0}^{h}(1-N)^u\binom{K-s}{h-u}\binom{s}{u}\}.\end{aligned}$$

Proof The left hand side equals

$$\sum_{j=0}^{K-h}(-1)^jN^{K-h-j}(N-1)^{h-K+1}\binom{h+j}{h}\binom{K-s}{h-s+j}\frac{N^{m-h-t}-1}{N-1}(-1)^t\binom{K-h}{t}.$$

Therefore it suffices to show

$$\begin{aligned}&\sum_{j=0}^{K-h}(-1)^jN^{K-h-j}(N-1)^{h-K+1}\binom{h+j}{h}\binom{K-s}{h-s+j}\\ &= (-1)^s(N-1)^{1-s}\sum_{u=0}^{h}(1-N)^u\binom{K-s}{h-u}\binom{s}{u}. \qquad (12)\end{aligned}$$

Consider the binomial coefficient on the left hand side of (12).

$$\begin{aligned}\binom{h+j}{h}\binom{K-s}{h-s+j} &= \sum_{u=\max(0,s-j)}^{\min(s,h)}\binom{h+j-s}{h-u}\binom{s}{u}\binom{K-s}{h-s+j}\\ &= \sum_{u=0}^{h}\binom{h+j-s}{h-u}\binom{s}{u}\binom{K-s}{h-s+j}. \qquad (13)\end{aligned}$$

Using the relation

$$\binom{h+j-s}{a}\binom{K-s}{h+j-s}=\binom{K-s}{a}\binom{K-s-a}{h+j-s-a}$$

with $a=h-u$, (13) can be written as

$$\sum_{u=0}^{h}\binom{s}{u}\binom{K-s}{h-u}\binom{K-s-h+u}{u+j-s}.$$

Now we take the sum over j. Then

$$\begin{aligned}
&\sum_{j=0}^{K-h}(-1)^j N^{K-h-j}(N-1)^{h-K+1}\binom{K-s-h+u}{u+j-s}\\
&= (N-1)^{h-K+1}\sum_{j=s-u}^{K-h}(-1)^j N^{K-h-j}\binom{K-s-h+u}{u+j-s}\\
&= (N-1)^{h-K+1}(-1)^{s-u}(N-1)^{K-s-h+u}\\
&= (N-1)^{1-s+u}(-1)^{s-u}.
\end{aligned}$$

Accordingly the left hand side of (12) is written as

$$\sum_{u=0}^{h}(N-1)^{1-s}(1-N)^u(-1)^s\binom{s}{u}\binom{K-s}{h-u}$$

and this proves the lemma. Q.E.D.

Since $(-1)^s(N-1)^{1-s}$ is the common term, we omit this term and expand the rest of the terms. Expand $(N^{m-t-h}-1)/(N-1)$ as

$$\frac{N^{m-t-h}-1}{N-1}=\sum_{i=0}^{m-t-h-1}N^i,$$

then

$$\begin{aligned}
&\sum_{h=0}^{m-1}\sum_{t=0}^{m-h-1}\binom{K-h}{t}\frac{N^{m-t-h}-1}{N-1}(-1)^t\cdot\{\sum_{u=0}^{h}(1-N)^u\binom{K-s}{h-u}\binom{s}{u}\}\\
&= \sum_{h=0}^{m-1}\sum_{t=0}^{m-h-1}N^t\{\sum_{i=0}^{m-t-h-1}\binom{K-h}{i}(-1)^i\}\cdot\{\sum_{u=0}^{h}(1-N)^u\binom{K-s}{h-u}\binom{s}{u}\}. \quad (14)
\end{aligned}$$

By induction on c it is easy to show

$$\sum_{b=0}^{c}(-1)^b\binom{a}{b}=(-1)^c\binom{a-1}{c} \quad (15)$$

for nonnegative c. Then (14) is

$$\sum_{h=0}^{m-1}\sum_{t=0}^{m-h-1}N^t\{\binom{K-h-1}{m-t-h-1}(-1)^{m-t-h-1}\}\cdot\{\sum_{u=0}^{h}(1-N)^u\binom{K-s}{h-u}\binom{s}{u}\}. \quad (16)$$

By $(1-N)^u = \sum_{l=0}^{u}(-N)^l\binom{u}{l}$ the second term of (16) can be written as

$$\begin{aligned}\sum_{u=0}^{h}(1-N)^u\binom{K-s}{h-u}\binom{s}{u} &= \sum_{u=0}^{h}\sum_{l=0}^{u}(-N)^l\binom{u}{l}\binom{K-s}{h-u}\binom{s}{u} \\ &= \sum_{l=0}^{h}(-N)^l\sum_{u=l}^{h}\binom{u}{l}\binom{K-s}{h-u}\binom{s}{u}.\end{aligned}$$

Let $v = u - l$, then

$$\begin{aligned}\sum_{u=l}^{h}\binom{u}{l}\binom{K-s}{h-u}\binom{s}{u} &= \sum_{v=0}^{h}\binom{s}{l+v}\binom{K-s}{h-l-v}\binom{l+v}{l} \\ &= \sum_{v=0}^{h}\binom{K-h}{s-l-v}\binom{h-l}{v}\frac{(K-s)!s!}{(K-h)!(h-l)!l!} \\ &= \binom{K-l}{s-l}\frac{(K-s)!s!}{(K-h)!(h-l)!l!} \\ &= \binom{s}{l}\binom{K-l}{h-l}.\end{aligned}$$

Consequently (16) equals

$$\sum_{h=0}^{m-1}\{\sum_{t=0}^{m-h-1}N^t\binom{K-h-1}{m-t-h-1}(-1)^{m-t-h-1}\}\cdot\{\sum_{l=0}^{h}(-N)^l\binom{s}{l}\binom{K-l}{h-l}\}.$$

At this point, we want to know the coefficient of $N^u, u = 0,\ldots,m-1$.

$$\begin{aligned}&\sum_{h=0}^{m-1}\sum_{u=0}^{m-1}\sum_{l=0}^{u}\binom{K-h-1}{m-(u-l)-h-1}\binom{s}{l}\binom{K-l}{h-l}(-1)^{m-(u-l)-h-1+l}N^u \\ &= \sum_{u=0}^{m-1}\sum_{l=0}^{u}\sum_{h=0}^{m-1}\binom{s}{l}\binom{K-l}{h-l}\binom{K-h-1}{m-h-1-(u-l)}(-1)^{m-h-1-(u-l)+l}N^u. \quad (17)\end{aligned}$$

Let $q = 1 + (u - l)$. (17) equals

$$\begin{aligned}&\sum_{u=0}^{m-1}\sum_{l=0}^{u}\sum_{h=0}^{m-1}\binom{s}{l}\binom{K-l}{h-l}\binom{K-h-1}{m-h-q}(-1)^{m-h-q+l}N^u \\ &= \sum_{u=0}^{m-1}\sum_{l=0}^{u}\sum_{h=l}^{m-q}\binom{s}{l}\binom{K-l}{h-l}\binom{K-h-1}{m-h-q}(-1)^{m-h-q}(-1)^lN^u. \quad (18)\end{aligned}$$

Lemma 5 *For $l < m-q,\ q \geq 1,\ l \geq 0,\ K \geq m$*

$$\sum_{h=l}^{m-q}(-1)^{m-h-q}\binom{K-h-1}{m-h-q}\binom{K-l}{h-l} = 1.$$

Proof Note that $K - l \geq 0$, $K - m + q \geq 0$. Then

$$\frac{(1+x)^{K-l}}{(1+x)^{K-m+q}} = (1+x)^{m-q-l}.$$

The coefficient of x^{m-q-l} on the right hand side is obviously 1. We expand the left hand side.

$$\sum_{i,j} \binom{K-l}{i} x^i \binom{K-m+q+j-1}{K-m+q-1} (-1)^j x^j.$$

Consider the case $i + j = m - q - l$. The coefficient of x^{m-q-l} is

$$\sum_{i=0}^{m-q-l} \binom{K-l}{i} \binom{K-m+q-1+(m-q-l-i)}{K-m+q-1} (-1)^{m-q-l-i}.$$

Let $i = h - l$. This leads to

$$\sum_{h=l}^{m-q} \binom{K-1-h}{m-h-q} \binom{K-l}{h-l} (-1)^{m-h-q} = 1.$$

Q.E.D.

Using Lemma 5, (18) equals

$$\sum_{u=0}^{m-1} \sum_{l=0}^{u} \binom{s}{l} (-1)^l N^u = \sum_{u=0}^{m-1} (-N)^u \binom{s-1}{u}$$

by (15). Summarizing the above calculations we have

$$\begin{aligned} \mathrm{Var}(T_{EL}) &= \mathrm{Var}(T_R) + \frac{1}{N^m} \sum_{s=1}^{K} \varphi_s^2 \{(-1)^s (N-1)^{1-s} \cdot \sum_{u=0}^{m-1} (-N)^u \binom{s-1}{u}\} \\ &= \frac{1}{N^m} \sum_{s=1}^{K} \varphi_s^2 \{1 + (-1)^s (N-1)^{1-s} \cdot \sum_{u=0}^{m-1} (-N)^u \binom{s-1}{u}\} + \nabla_r. \end{aligned}$$

This completes the proof of Theorem 1.

4 A SUFFICIENT CONDITION FOR VARIANCE REDUCTION

In this section we obtain a sufficient condition for the variance reduction $\mathrm{Var}(T_{EL}) \leq \mathrm{Var}(T_R)$. McKay, Conover and Beckman (1979) shows that the monotonicity of $g(X_1, \ldots, X_K)$ in each argument $X_i, i = 1, \ldots, K$, is a sufficient condition for the variance reduction at $m = 1$. In the following we clarify their condition using the results of Section 3 and generalize the condition for $m > 1$. Our main result of this section is given in Theorem 3 of Section 4.2.

In view of (3) and (4) $\mathrm{Var}(T_{EL}) \leq \mathrm{Var}(T_R)$ if and only if the covariance term is nonpositive, i.e.,

$$(N^m - 1)\mathrm{Cov}_m(\mu_{\boldsymbol{y}}, \mu_{\boldsymbol{z}}) = \sum_{s=1}^{K} \varphi_s^2 \{(-1)^s (N-1)^{1-s} \cdot \sum_{u=0}^{m-1} (-N)^u \binom{s-1}{u}\} \leq 0. \quad (19)$$

We rewrite (19) as

$$\sum_{u=0}^{m-1}(-N)^u \sum_{s=1+u}^{K} \varphi_s^2(-1)^s(N-1)^{1-s}\binom{s-1}{u} \leq 0. \tag{20}$$

Our objective is to obtain a sufficient condition for (20). Here we prefer investigating the case $m=1$ first and generalizing the result to $m>1$ later.

4.1 LHS

For the case $m=1$ the covariance term is

$$\begin{aligned}\sum_{s=1}^{K}\varphi_s^2(-1)^s(N-1)^{1-s} &= -\sum_{i=1}^{K}\sum_{z_i}\alpha_1(z_i)^2\cdot N^{-1}(N-1)^0 \\ &+\sum_{i_1<i_2}\sum_{z_{i_1}}\sum_{z_{i_2}}\alpha_2(z_{i_1},z_{i_2})^2\cdot N^{-2}(N-1)^{-1} \\ &-\sum_{i_1<i_2<i_3}\sum_{z_{i_1}}\sum_{z_{i_2}}\sum_{z_{i_3}}\alpha_3(z_{i_1},z_{i_2},z_{i_3})^2\cdot N^{-3}(N-1)^{-2} \\ &+\cdots \\ &+(-1)^K\sum_{z_1}\cdots\sum_{z_K}\alpha_K(z_1,\ldots,z_K)^2N^{-K}(N-1)^{1-K}.\end{aligned} \tag{21}$$

Let us consider a function with K arguments. We call it *monotone in the ith argument*, if regarded as a function of the ith argument (with all other arguments held fixed arbitrarily), it is monotone in the usual sense. Direction of the monotonicity (increasing or decreasing) must not depend on the values of the other arguments. This is the same monotonicity condition used in McKay, Conover and Beckman (1979). Their proof of the variance reduction depends on results of Lehmann (1966). Our proof is completely different and we utilize the monotonicity in a different way. For example, "$\mu_{\boldsymbol{z}}$ is monotone in z_1" implies that

$$(\mu_{z_1z_2\ldots z_K}-\mu_{z_1'z_2\ldots z_K})(\mu_{z_1y_2\ldots y_K}-\mu_{z_1'y_2\ldots y_K})\geq 0,$$

for all $z_1, z_1', z_2, \ldots, z_K, y_2, \ldots, y_K$.

Lemma 6 *If $\mu_{\boldsymbol{z}}$ is monotone in z_1, then*

$$\begin{aligned}0 \leq\ & \sum_{z_1}\alpha_1(z_1)^2\cdot N^{-1}(N-1)^0(-1)^0 \\ &+\sum_{1<i}\sum_{z_1}\sum_{z_i}\alpha_2(z_1,z_i)^2\cdot N^{-2}(N-1)^{-1}(-1)^1 \\ &+\sum_{1<i_1<i_2}\sum_{z_1}\sum_{z_{i_1}}\sum_{z_{i_2}}\alpha_3(z_1,z_{i_1},z_{i_2})^2\cdot N^{-3}(N-1)^{-2}(-1)^2 \\ &+\cdots \\ &+\sum_{z_1}\cdots\sum_{z_K}\alpha_K(z_1,\ldots,z_K)^2N^{-K}(N-1)^{1-K}(-1)^{K-1}.\end{aligned}$$

Proof By the monotonicity in z_1

$$(\mu_{z_1 z_2 \ldots z_K} - \mu_{z_1' z_2 \ldots z_K})(\mu_{z_1 y_2 \ldots y_K} - \mu_{z_1' y_2 \ldots y_K}) \geq 0,$$
$$\forall z_1, z_1', z_2, \ldots, z_K, y_2 \neq z_2, \ldots, y_K \neq z_K.$$

Therefore

$$\sum_{z_1} \sum_{z_1'} \sum_{z_2} \cdots \sum_{z_K} \sum_{y_2 \neq z_2} \cdots \sum_{y_K \neq z_K} (\mu_{z_1 z_2 \ldots z_K} - \mu_{z_1' z_2 \ldots z_K})(\mu_{z_1 y_2 \ldots y_K} - \mu_{z_1' y_2 \ldots y_K}) \geq 0. \tag{22}$$

The left hand side of this inequality can be written as

$$\begin{aligned}
&\sum_{z_1} \sum_{z_1'} \sum_{z_2} \cdots \sum_{z_K} \sum_{y_2 \neq z_2} \cdots \sum_{y_K \neq z_K} \\
&\{(\alpha_1(z_1) - \alpha_1(z_1'))^2 \\
&+ \sum_i (\alpha_2(z_1, z_i) - \alpha_2(z_1', z_i))(\alpha_2(z_1, y_i) - \alpha_2(z_1', y_i)) \\
&+ \sum_{i_1 < i_2} (\alpha_3(z_1, z_{i_1}, z_{i_2}) - \alpha_3(z_1', z_{i_1}, z_{i_2}))(\alpha_3(z_1, y_{i_1}, y_{i_2}) - \alpha_3(z_1', y_{i_1}, y_{i_2})) \\
&+ \cdots \\
&+ (\alpha_K(z_1, z_2, .., z_K) - \alpha_K(z_1', z_2, .., z_K))(\alpha_K(z_1, y_2, .., y_K) - \alpha_K(z_1', y_2, .., y_K))\}.
\end{aligned}$$

Cross terms sum up to zero. For $2 \leq t \leq s$

$$\begin{aligned}
&\sum_{y_t \neq z_t} (\alpha_s(z_1, y_2, \ldots, y_s) - \alpha_s(z_1', y_2, \ldots, y_s)) \\
&\quad = \quad (-1)(\alpha_s(z_1, y_2, \ldots, z_t, \ldots, y_s) - \alpha_s(z_1', y_2, \ldots, z_t, \ldots, y_s)).
\end{aligned}$$

Further

$$\sum_{z_1} \sum_{z_1'} (\alpha_s(z_1, \ldots) - \alpha_s(z_1', \ldots))^2 = 2N \sum_{z_1} (\alpha_s(z_1, \ldots))^2.$$

Therefore (22) divided by $2N$ is

$$\begin{aligned}
0 \quad \leq \quad & N^{K-1}(N-1)^{K-1} \sum_{z_1} \alpha_1(z_1)^2 \\
& + N^{K-2}(N-1)^{K-2} \sum_{2 \leq i \leq K} \sum_{z_1} \sum_{z_i} \alpha_2(z_1, z_i)^2(-1) \\
& + N^{K-3}(N-1)^{K-3} \sum_{2 \leq i_1 < i_2 \leq K} \sum_{z_1} \sum_{z_{i_1}} \sum_{z_{i_2}} \alpha_3(z_1, z_{i_1}, z_{i_2})^2(-1)^2 \\
& + \cdots \\
& + N^0(N-1)^0 \sum_{z_1} \cdots \sum_{z_K} \alpha_K(z_1, \ldots, z_K)^2(-1)^{K-1}.
\end{aligned} \tag{23}$$

Multiplying by $N^{-K}(N-1)^{1-K}$ proves the lemma. Q.E.D.

By Lemma 6, under the monotonicity of $\mu_{\boldsymbol{z}}$ in z_1 our new sufficient condition for the variance reduction of LHS is given as

$$0 \quad \leq \quad \sum_{1<i} \sum_{z_i} \alpha_1(z_i)^2 \cdot N^{-1}(N-1)^0(-1)^0$$

$$
\begin{aligned}
&+ \sum_{1<i_1<i_2} \sum_{z_{i_1}} \sum_{z_{i_2}} \alpha_2(z_{i_1}, z_{i_2})^2 \cdot N^{-2}(N-1)^{-1}(-1)^1 \\
&+ \sum_{1<i_1<i_2<i_3} \sum_{z_{i_1}} \sum_{z_{i_2}} \sum_{z_{i_3}} \alpha_3(z_{i_1}, z_{i_2}, z_{i_3})^2 \cdot N^{-3}(N-1)^{-2}(-1)^2 \\
&+ \cdots \\
&+ \sum_{z_2} \cdots \sum_{z_K} \alpha_{K-1}(z_2, \ldots, z_K)^2 \cdot N^{1-K}(N-1)^{2-K}(-1)^{K-2}. \qquad (24)
\end{aligned}
$$

Note that -1 times the right hand side of this inequality corresponds to the covariance of partially averaged cell mean function:

$$
\mu_{\cdot z_2 z_3 \ldots z_K} = \frac{1}{N} \sum_{z_1=1}^{N} \mu_{z_1 z_2 \ldots z_K}.
$$

$\mu_{\cdot z_2 z_3 \ldots z_K}$ can be regarded as a function with $K-1$ arguments and the ANOVA decomposition for $\mu_{\cdot z_2 z_3 \ldots z_K}$ is given as

$$
\begin{aligned}
\mu_{\cdot z_2 \ldots z_K} - \mu &= \sum_{i=2}^{K} \alpha_1(z_i) \\
&+ \sum_{2 \le i_1 < i_2} \alpha_2(z_{i_1}, z_{i_2}) \\
&+ \cdots \\
&+ \alpha_{K-1}(z_2, \ldots, z_K).
\end{aligned}
$$

Let

$$
\boldsymbol{z}^- = (z_2, \ldots, z_K), \quad \boldsymbol{y}^- = (y_2, \ldots, y_K).
$$

Then (24) is equivalent to $\mathrm{Cov}(\mu_{\boldsymbol{z}^-}, \mu_{\boldsymbol{y}^-}) \le 0$ under LHS.

Therefore by assuming the monotonicity in any one axis, we have reduced our problem from K arguments to $K-1$ arguments. Now we focus on the covariance of the partially averaged cell mean function $\mu_{\boldsymbol{z}^-}$. Applying the monotonicity in z_2, we can further reduce the number of arguments. To complete this inductive argument we need the following initial condition.

Lemma 7 *Let $m = 1, K = 2$. Then monotonicity in one axis implies*

$$
\mathrm{Cov}(\mu_{\boldsymbol{z}}, \mu_{\boldsymbol{y}}) \le 0.
$$

Proof

$$
\begin{aligned}
\mathrm{Cov}(\mu_{\boldsymbol{z}}, \mu_{\boldsymbol{y}}) &= -N^{-1}\{\sum_{z_1} \alpha_1(z_1)^2 + \sum_{z_2} \alpha_1(z_2)^2\} \\
&\quad + N^{-2}(N-1)^{-1} \sum_{z_1} \sum_{z_2} \alpha_2(z_1, z_2)^2.
\end{aligned}
$$

As in (23) the monotonicity in z_1 implies

$$
\begin{aligned}
0 &\ge -N^{-1}\{\sum_{z_1} \alpha_1(z_1)^2\} \\
&\quad + N^{-2}(N-1)^{-1} \sum_{z_1} \sum_{z_2} \alpha_2(z_1, z_2)^2.
\end{aligned}
$$

Now the lemma follows from $\sum_{z_2} \alpha_1(z_2)^2 \geq 0$. Q.E.D.

Summarizing the above arguments, we have established the following theorem.

Theorem 2 *If the cell mean function $\mu_{\mathbf{z}}$ is monotone in any $K-1$ out of the K axes, then* $\mathrm{Var}(T_L) \leq \mathrm{Var}(T_R)$.

Note that the ranges of $X_i^{(z_i)}, i = 1, \ldots, K$, are monotone in z_i's. Therefore, if g is monotone in the ith argument then the cell mean function is also monotone in the ith axis. Hence the following corollary is an immediate consequence of Theorem 2.

Corollary 1 *If $g(X_1, \ldots, X_K)$ is monotone in $K-1$ arguments, then* $\mathrm{Var}(T_L) \leq \mathrm{Var}(T_R)$.

The sufficient condition given by McKay, Conover and Beckman (1979) for LHS requires that g is monotone in all K arguments for variance reduction. We have shown that actually $K-1$ monotonicities are sufficient.

4.2 ELHS

To show the variance reduction for $m > 1$, we employ induction on m. Let us begin with $m = 2$. The covariance term is

$$\begin{aligned}
& \sum_{i=1}^{K} \sum_{z_i} \alpha_1(z_i)^2 \cdot N^{-1}(N-1)^0(-1)^1 \\
& + \sum_{i_1<i_2} \sum_{z_{i_1}} \sum_{z_{i_2}} \alpha_2(z_{i_1}, z_{i_2})^2 \cdot N^{-2}(N-1)^{-1}(-1)^2 \\
& + \sum_{i_1<i_2<i_3} \sum_{z_{i_1}} \sum_{z_{i_2}} \sum_{z_{i_3}} \alpha_3(z_{i_1}, z_{i_2}, z_{i_3})^2 \cdot N^{-3}(N-1)^{-2}(-1)^3 \\
& + \cdots \\
& + \sum_{z_1} \cdots \sum_{z_K} \alpha_K(z_1, \ldots, z_K)^2 N^{-K}(N-1)^{1-K}(-1)^K \\
+ \; & \sum_{i_1<i_2} \sum_{z_{i_1}} \sum_{z_{i_2}} \alpha_2(z_{i_1}, z_{i_2})^2 \cdot N^{-1}(N-1)^{-1}(-1)^1 \binom{2-1}{1} \\
& + \sum_{i_1<i_2<i_3} \sum_{z_{i_1}} \sum_{z_{i_2}} \sum_{z_{i_3}} \alpha_3(z_{i_1}, z_{i_2}, z_{i_3})^2 \cdot N^{-2}(N-1)^{-2}(-1)^2 \binom{3-1}{1} \\
& + \cdots \\
& + \sum_{z_1} \cdots \sum_{z_K} \alpha_K(z_1, \ldots, z_K)^2 N^{-K+1}(N-1)^{1-K}(-1)^{K-1} \binom{K-1}{1}. \qquad (25)
\end{aligned}$$

Subtracting (21) from (25), the remainder is

$$\sum_{s=2}^{K} \varphi_s^2 (-1)^s (N-1)^{1-s} (-N) \binom{s-1}{1}. \qquad (26)$$

Supposing that the reduction of variance under LHS holds, the reduction under $m = 2$ is assured if (26) is nonpositive. As in our argument for LHS, we reduce the number of arguments via induction. We formulate an extension of the monotonicity at $m = 2$ as follows. $\mu_{\boldsymbol{z}}$ *is monotone in* (z_1, z_2) if $\forall z_1, \ldots, z_K, z_1', z_2', y_3, \ldots, y_K$

$$\begin{aligned} &(\mu_{z_1 z_2 \cdots z_K} - \mu_{z_1' z_2 \cdots z_K} - \mu_{z_1 z_2' z_3 \cdots z_K} + \mu_{z_1' z_2' z_3 \cdots z_K}) \\ &\times \quad (\mu_{z_1 z_2 y_3 \cdots y_K} - \mu_{z_1' z_2 y_3 \cdots y_K} - \mu_{z_1 z_2' y_3 \cdots y_K} + \mu_{z_1' z_2' y_3 \cdots y_K}) \geq 0. \end{aligned} \tag{27}$$

The reason why we call it "monotone" is that the monotonicity in two axes is related to the sign of the mixed partial derivative. By

$$\frac{\partial^2}{\partial x \partial y} f(x, y) = \lim \frac{f(x + \Delta_x, y + \Delta_y) - f(x, y + \Delta_y) - f(x + \Delta_x, y) + f(x, y)}{\Delta_x \Delta_y},$$

the monotonicity means that for differentiable functions the sign of the partial mixed derivative remains the same regardless of the values of the variables.

We sum up (27) for all $z_1, \ldots, z_K, z_1', z_2', y_3 \neq z_3, \ldots, y_K \neq z_K$. This results in

$$\begin{aligned} 0 \;\leq\; & \sum_{z_1} \sum_{z_2} \alpha_2(z_1, z_2)^2 (N(N-1))^{-1} \\ & - \sum_{2<i} \sum_{z_1} \sum_{z_2} \sum_{z_i} \alpha_3(z_1, z_2, z_i)^2 (N(N-1))^{-2} \\ & + \sum_{2<i_1<i_2} \sum_{z_1} \sum_{z_2} \sum_{z_{i_1}} \sum_{z_{i_2}} \alpha_4(z_1, z_2, z_{i_1}, z_{i_2})^2 (N(N-1))^{-3} \\ & + \cdots \\ & + (-1)^K \sum_{z_1} \cdots \sum_{z_K} \alpha_K(z_1, \ldots, z_K) (N(N-1))^{-(K-1)}. \end{aligned} \tag{28}$$

In order to use induction on z_1, we utilize the $\binom{K-1}{1}$ monotonicities in $(z_1, z_2), \ldots, (z_1, z_K)$. If the cell mean function is monotone in each $(z_1, z_2), \ldots, (z_1, z_K)$, we call it *quadratic monotonicity* in z_1. We obtain an inequality like (28) for each pair and sum them up. By symmetry and

$$\binom{K-1}{1} \binom{K-2}{s-2} / \binom{K-1}{s-1} = \binom{s-1}{1}$$

this yields

$$\begin{aligned} 0 \;\leq\; & \sum_{i \neq 1} \sum_{z_1} \sum_{z_i} \alpha_2(z_1, z_i)^2 (N(N-1))^{-1} \binom{2-1}{1} \\ & - \sum_{1<i_1<i_2} \sum_{z_1} \sum_{z_{i_1}} \sum_{z_{i_2}} \alpha_3(z_1, z_{i_1}, z_{i_2})^2 (N(N-1))^{-2} \binom{3-1}{1} \\ & + \sum_{1<i_1<i_2<i_3} \sum_{z_1} \sum_{z_{i_1}} \sum_{z_{i_2}} \sum_{z_{i_3}} \alpha_4(z_1, z_{i_1}, z_{i_2}, z_{i_3})^2 (N(N-1))^{-3} \binom{4-1}{1} \\ & + \cdots \\ & + (-1)^K \sum_{z_1} \cdots \sum_{z_K} \alpha_K(z_1, \ldots, z_K) (N(N-1))^{-(K-1)} \binom{K-1}{1}. \end{aligned} \tag{29}$$

Therefore the reduction on z_1 is achieved. Applying this process for each axis leads us to the case $K = 3$. Then (26) is

$$-\{\sum_{z_1} \sum_{z_2} \alpha_2(z_1, z_2)^2 + \sum_{z_1} \sum_{z_3} \alpha_2(z_1, z_3)^2 + \sum_{z_2} \sum_{z_3} \alpha_2(z_2, z_3)^2\} \cdot N^{-1}(N-1)^{-1}$$

$$+\sum_{z_1}\sum_{z_2}\sum_{z_3}\alpha_3(z_1,z_2,z_3)^2 N^{-2}(N-1)^{-2}\binom{2}{1}. \tag{30}$$

(29) for $K=3$ is

$$\begin{aligned} 0 \ \geq \ & \{\sum_{z_1}\sum_{z_2}\alpha_2(z_1,z_2)^2+\sum_{z_1}\sum_{z_3}\alpha_2(z_1,z_3)^2\}\cdot N^{-1}(N-1)^{-1}(-1)^1\binom{1}{1} \\ & +\sum_{z_1}\sum_{z_2}\sum_{z_3}\alpha_3(z_1,z_2,z_3)^2 N^{-2}(N-1)^{-2}(-1)^2\binom{2}{1}. \end{aligned}$$

By $\sum_{z_2}\sum_{z_3}\alpha_2(z_2,z_3)^2\geq 0$, (29) implies that (30) is nonpositive for $K=3$. By induction we have proved the following proposition for $m=2$.

Proposition 1 *Let $m=2$. If the cell mean function is monotone in any $K-1$ out of the K axes and quadratic monotonicity holds in any $K-2$ axes, then* $\mathrm{Var}(T_{EL})\leq\mathrm{Var}(T_R)$.

For general m we consider t-th order monotonicity. We define *monotonicity of $\mu_{\boldsymbol{z}}$ in* $(z_1,\ldots,z_t)$ as $\forall v_1,v_1',v_2,v_2',\ldots,v_t,v_t'\ z_{t+1},\ldots,z_K,y_{t+1},\ldots,y_K$,

$$\begin{aligned} 0 \ \leq \ & \{\mu_{v_1\cdots v_t\boldsymbol{z}} \\ & +(-1)(\mu_{v_1'v_2v_3\cdots v_t\boldsymbol{z}}+\mu_{v_1v_2'v_3\cdots v_t\boldsymbol{z}}+\cdots+\mu_{v_1v_2v_3\cdots v_t'\boldsymbol{z}}) \\ & +(-1)^2(\mu_{v_1'v_2'v_3\cdots v_t\boldsymbol{z}}+\mu_{v_1'v_2v_3'\cdots v_t\boldsymbol{z}}+\cdots+\mu_{v_1v_2v_3\cdots v_{t-1}'v_t'\boldsymbol{z}}) \\ & +(-1)^3(\mu_{v_1'v_2'v_3'\cdots v_t\boldsymbol{z}}+\cdots+\mu_{v_1\cdots v_{t-2}'v_{t-1}'v_t'\boldsymbol{z}}) \\ & +\cdots \\ & +(-1)^t(\mu_{v_1'v_2'v_3'\cdots v_t'\boldsymbol{z}})\} \\ & \times\{\mu_{v_1\cdots v_t\boldsymbol{y}} \\ & +(-1)(\mu_{v_1'v_2v_3\cdots v_t\boldsymbol{y}}+\mu_{v_1v_2'v_3\cdots v_t\boldsymbol{y}}+\cdots+\mu_{v_1v_2v_3\cdots v_t'\boldsymbol{y}}) \\ & +(-1)^2(\mu_{v_1'v_2'v_3\cdots v_t\boldsymbol{y}}+\mu_{v_1'v_2v_3'\cdots v_t\boldsymbol{y}}+\cdots+\mu_{v_1v_2v_3\cdots v_{t-1}'v_t'\boldsymbol{y}}) \\ & +(-1)^3(\mu_{v_1'v_2'v_3'\cdots v_t\boldsymbol{y}}+\cdots+\mu_{v_1\cdots v_{t-2}'v_{t-1}'v_t'\boldsymbol{y}}) \\ & +\cdots \\ & +(-1)^t(\mu_{v_1'v_2'v_3'\cdots v_t'\boldsymbol{y}})\}, \end{aligned}$$

where $\boldsymbol{z}=(z_{t+1},\ldots,z_K),\boldsymbol{y}=(y_{t+1},\ldots,y_K)$. Then we define *$t$-th order of the monotonicity of $\mu_{\boldsymbol{z}}$ in z_1* if $\boldsymbol{z}$ is monotone in $(z_1,z_{i_1},\ldots,z_{i_t})$ for all $\binom{K-1}{t-1}$ combinations $(1,i_2,\ldots,i_t)$, $1<i_2<\cdots<i_t\leq K$. The monotonicity in other axes is similarly defined.

Theorem 3 *Suppose that for all $1\leq t\leq m$, t-th order of the monotonicities concerning $K-t$ axes hold, then* $\mathrm{Var}(T_{EL})\leq\mathrm{Var}(T_R)$.

Proof We show the nonpositiveness of the covariance. The covariance term equals $\sum_{t=1}^m c_t$ where

$$c_t=(-N)^{t-1}\sum_{s=t}^{K}\varphi_s^2(-1)^s(N-1)^{s-1}\binom{s-1}{t-1}.$$

We claim that t-th order monotonicity in any $K-t$ axes implies $c_t \le 0$. The monotonicity in $(z_1, z_2, \ldots, z_t)$ yields

$$
\begin{aligned}
0 \;\le\; & \sum_{z_1}\sum_{z_2}\cdots\sum_{z_t}\alpha_t(z_1,z_2,\ldots,z_t)^2(N(N-1))^{-1} \\
& -\sum_{t<i}\sum_{z_1}\cdots\sum_{z_t}\sum_{z_i}\alpha_{t+1}(z_1,\ldots,z_t,z_i)^2(N(N-1))^{-2} \\
& +\sum_{t<i_1<i_2}\sum_{z_1}\cdots\sum_{z_t}\sum_{z_{i_1}}\sum_{z_{i_2}}\alpha_{t+2}(z_1,\ldots,z_t,z_{i_1},z_{i_2})^2(N(N-1))^{-3} \\
& +\cdots \\
& +(-1)^K\sum_{z_1}\cdots\sum_{z_K}\alpha_K(z_1,\ldots,z_K)(N(N-1))^{-(K-1)}.
\end{aligned}
$$

By

$$
\binom{K-1}{t-1}\binom{K-t}{s-t}\Big/\binom{K-1}{s-1} = \binom{s-1}{t-1},
$$

t-th order of the monotonicity in z_1 implies

$$
\begin{aligned}
0 \;\le\; & \sum_{1<i_1<\cdots<i_{t-1}}\sum_{z_1}\sum_{z_{i_1}}\cdots\sum_{z_{i_{t-1}}}\alpha_t(z_1,z_{i_1},\ldots,z_{i_{t-1}})^2(N(N-1))^{-1}\binom{t-1}{t-1} \\
& -\sum_{1<i_1<\cdots<i_t}\sum_{z_1}\sum_{z_{i_1}}\cdots\sum_{z_{i_t}}\alpha_{t+1}(z_1,z_{i_1},\ldots,z_{i_t})^2(N(N-1))^{-2}\binom{t}{t-1} \\
& +\sum_{1<i_1<\cdots<i_{t+1}}\sum_{z_1}\sum_{z_{i_1}}\cdots\sum_{z_{i_{t+1}}}\alpha_{t+2}(z_1,z_{i_1},\ldots,z_{i_{t+1}})^2(N(N-1))^{-3}\binom{t+1}{t-1} \\
& +\cdots \\
& +(-1)^K\sum_{z_1}\cdots\sum_{z_K}\alpha_K(z_1,\ldots,z_K)(N(N-1))^{-(K-1)}\binom{K-1}{t-1}.
\end{aligned}
$$

We see that the monotonicity leads to the reduction of one argument. By induction $K-(t+1)$ t-th order monotonicities lead to the initial condition

$$
\begin{aligned}
0 \;\le\; & \sum_{i_1<\cdots<i_t}\sum_{z_{i_1}}\cdots\sum_{z_{i_{t-1}}}\alpha_t(z_{i_1},\ldots,z_{i_t})^2(N(N-1))^{-1}\binom{t-1}{t-1} \\
& -\sum_{z_1}\cdots\sum_{z_{t+1}}\alpha_{t+1}(z_1,\ldots,z_{t+1})(N(N-1))^{-2}\binom{t}{t-1},
\end{aligned}
$$

which is implied by t-th order monotonicity in z_1. This proves the theorem. Q.E.D.

If we define the t-th order monotonicity of $g(X_1,\ldots,X_K)$ in an obvious manner, then as in Corollary 1, the following corollary is an immediate consequence of Theorem 3.

Corollary 2 *If t-th order of monotonicity holds in $K-t$ arguments of $g(X_1,\ldots,X_K)$ for $1 \le t \le m$, then* $\mathrm{Var}(T_{EL}) \le \mathrm{Var}(T_R)$.

5 SOME SIMULATION RESULTS

In this section we confirm our theoretical results by simulation. We choose a rather simple situation where exact variance of ELHS(m) can be computed and compared to the simulated variance. Note that this is not a realistic situation in which simulations are actually used.

Let $K = 4, N = 5, m = 2$, and consider

$$W = \exp(X_1 + X_2 + X_3 + X_4),$$

where $X_i, i = 1, \ldots, 4, \sim U[0, 1]$ i. i. d. In this example all mixed derivatives are positive and hence monotonicities of all orders hold.

We obtained simulation variance with 1000000 replications. The sample size in SRS and ELHS equals $N^m = 5^2 = 25$. The simulated sampling distributions of SRS and ELHS are plotted in Figure 1 and Figure 2. In Figure 1 the histograms of T_R and T_{EL} are compared on a same scale. Figure 2 gives the histogram of T_{EL} in more detail. From Figure 1 we see that T_{EL} is much more concentrated around the true value μ. The numerical results of our simulation are summarized in Table 1. Observe that the true values and the simulated values in Table 1 are in close agreements and this confirms our theoretical results.

Table 1. Simulation Results (1000000 replications)

	SRS	ELHS(2)
True E[W]	8.717212	8.717212
Sample mean	8.716336	8.717124
True Variance of $\mu_{\mathbf{z}}$	1.071051	0.009239
True ∇_r	0.055047	0.055047
True Variance of T	1.126098	0.064286
Sample Variance of T	1.123594	0.064175
Minimum T	4.991916	7.613056
Maximum T	15.304256	10.060986

In this example W is the product of marginal functions and it can be expected that the higher order interaction effects are not so large. As a matter of fact, $\varphi_1^2 = 23.825679$, $\varphi_2^2 = 2.801345$, $\varphi_3^2 = 0.146388$, $\varphi_4^2 = 0.002869$. $m = 2$ seems to be a good choice in this example.

6 SOME FURTHER DISCUSSION

Method to map z_{ij} into X_{ij} is a controversial point. Although it produces bias, we can make $\nabla_r = 0$ by obtaining x_{ij} deterministically given z_{ij}. Hence one should consider the tradeoff between bias and variance. Owen (1992a) discusses midpoint rule and rectangular rule. Tang (1993) introduces Latin hypercube structure to the cells, i.e. the sample points in the given cells are generated by using a method like LHS. Under this stratification, ∇_r is reduced when W is additive.

To achieve specific objectives, some authors add restriction on the Latin hypercube design. Handcock (1991) proposes "cascading" Latin hypercube design. Sample points are obtained by using modified LHS with midpoint rule, and a few points from the

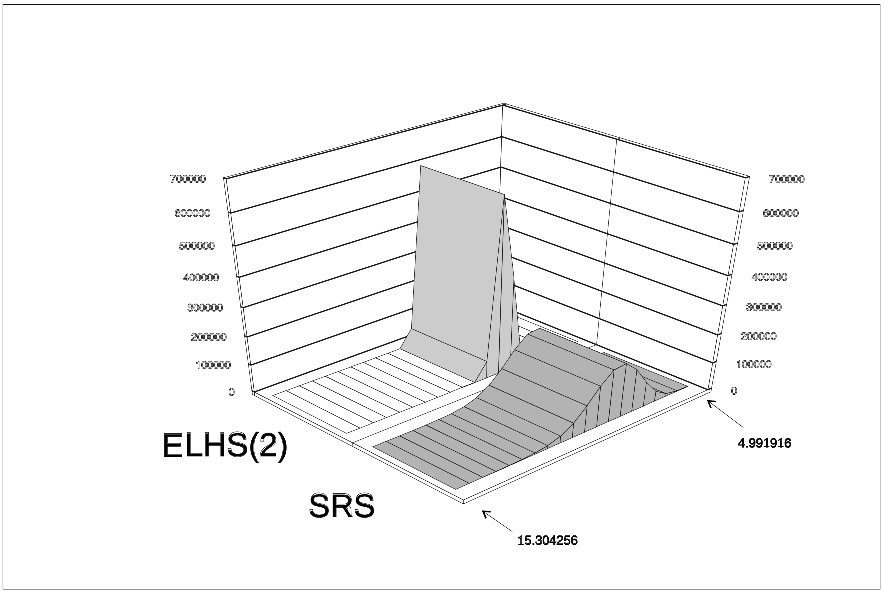

Figure 1: Comparison between T_R and T_{EL} of Section 5.

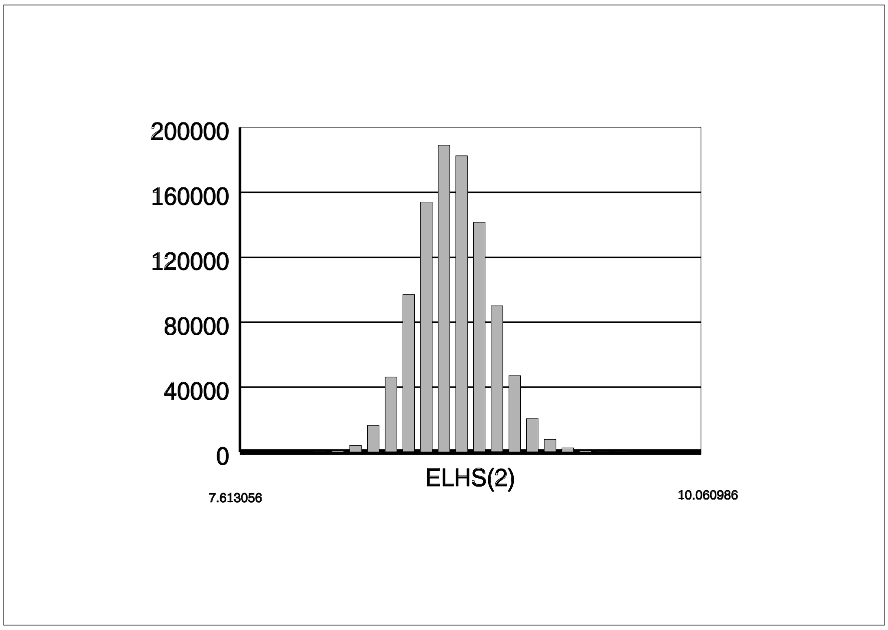

Figure 2: Histogram of T_{EL} of Section 5.

same cells are added with LHS. In the literature of experimental design, optimality of permuted generator arrays is discussed. Uniformity of the sample points is improved by restricting permutations. Tang (1994) introduces a criterion to compare design arrays. Shaw (1988) reviews criteria of uniformity.

In our assumption, each axis is independently distributed. Iman and Conover (1982) treat dependency in LHS, and Owen (1994) proposes another algorithm for controlling correlations. Stein (1987) discusses central limit theorem for LHS and Owen (1992b) gives a proof of the central limit theorem using the method of moments. It is natural to expect that the central limit theorem holds for T_{EL} under appropriate regularity conditions. See Figure 2.

References

[1] Bose, R. C. and Bush, K. A. (1952). Orthogonal Arrays of Strength Two and Three. *Annals of Mathematical Statistics*, **23**, 508–524.

[2] Galambos, J. and Simonelli, I. (1996). *Bonferroni-type Inequalities with Applications*. Springer.

[3] Handcock, M. S. (1991). On Cascading Latin Hypercube Designs. *Commun. Statist. -Theory Meth.*, **20**, 2, 417–439.

[4] 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.

[5] Lehmann, E. L. (1966). Some Concepts of Dependence. *Annals of Mathematical Statistics*, **35**, 1137–1153.

[6] McKay, M. D., Conover, W. J. and Beckman, R. 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.

[7] Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM.

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

[9] Owen, A. B. (1992b). A Central Limit Theorem for Latin Hypercube Sampling. *J. R. Statist. Soc. B*, **54**, 2, 541–551.

[10] Owen, A. B. (1994). Controlling Correlations in Latin Hypercube Samples. *Journal of the American Statistical Association*, **89**, 1517–1522.

[11] Patterson, H. D. (1954). The Errors of Lattice Sampling. *J. R. Statist. Soc. Ser.B*, **16**, 140–149.

[12] Plackett, R. L. and Burman, J. P. (1946). The Design of Optimum Multifactorial Experiments. *Biometrika*, **33**, 305–325.

[13] Rao, C. R. (1947). Factorial Experiments Derivable from Combinatorial Arrangements of Arrays. *J. R. Statist. Soc.*, Suppl., 9, 128–139.

[14] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley.

[15] Shaw, J. E. H. (1988). A Quasirandom Approach to Integration in Bayesian Statistics. *Annals of Statistics*, **16**, 2, 895–914.

[16] Stein, M. (1987). Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics*, **29**, 143–151.

[17] Tang, B. (1993). Orthogonal Array-based Latin Hypercubes. *Journal of the American Statistical Association*, **88**, 1392–1397.

[18] Tang, B. (1994). A Theorem for Selecting OA-based Latin Hypercubes Using a Distance Criterion. *Commun. Statist. -Theory Meth.*, **23**, 7, 2047–2058.