Journal of the Korean Statistical Society (1999), **28**: 2, pp 135–150

Asymptotic Comparison of Latin Hypercube Sampling and Its Stratified Version†

Jooho Lee[1]

Abstract

Latin hypercube sampling (LHS) introduced by McKay et al. (1979) is a widely used method for Monte Carlo integration. Stratified Latin hypercube sampling (SLHS) proposed by Choi and Lee (1993) improves LHS by combining it with stratified sampling. In this article it is shown that SLHS yields an asymptotically more accurate estimate than both stratified sampling and LHS.

Keywords: Monte Carlo integration; Computer simulation experiment; Latin hypercube sampling.

1. INTRODUCTION

Monte Carlo integration is often a useful alternative to quadrature in evaluating very complicated multidimensional integrals. Let $\mathbf{X} = (X_1, \ldots, X_K)$ be a random vector distributed uniformly on $[0,1]^K$. Then Monte Carlo integration may be considered as the problem of estimating the expected value of $Y = g(\mathbf{X})$ with $\bar{Y} = N^{-1} \sum_{i=1}^{N} g(\mathbf{X}_i)$ by generating random vectors $\mathbf{X}_1, \ldots, \mathbf{X}_N$ from the computer. Although random vectors are most easily generated via random sampling, stratification may be used to increase the precision of estimation. Latin hypercube sampling (LHS) introduced by McKay et al. (1979) is the first approach that utilizes the stratification idea. Later, Stein(1987) provided a theoretical support by showing that LHS yields an asymptotically more accurate estimate than random sampling. Now it is probably one of the most commonly used methods for Monte Carlo integration.

Computer simulation experiments are often used as substitutes for extremely expensive physical experiments. In designing computer simulation experiments

†This paper was supported by the research grant of Chungnam National University in 1996.
[1]Department of Statistics, Chungnam National University, Taejeon, 305-764, Korea.

it is important to choose design points evenly all over the space of input variables. Optimal design approaches such as the minimum integrated mean squared prediction error design proposed by Sacks et al. (1989) and the maximum entropy design proposed by Shewry and Wynn (1987) and Currin et al. (1991) are commonly used for computer simulation experiments. However, they require high computing cost when the number of input variables are large and the less expensive Latin hypercube design (LHD), which is a nonrandom midpoint Latin hypercube sample, may be considered as an alternative despite its lack of optimal properties.

Although a Latin hypercube stratifies each univariate margin simultaneously, it does not achieve stratification on each bivariate or higher margin. Therefore, if stratification is done on multivariate as well as univariate margins, then the resulting design scatters the design points more uniformly over the design region. Moreover, we may expect that the resulting probability sample would yield a more accurate estimate than a Latin hypercube sample. Orthogonal array-based LHS (OALHS) proposed by Owen (1992) and independently by Tang (1993), stratified LHS (SLHS) proposed by Choi and Lee (1993), and two-stage LHS (TLHS) proposed by Im et al. (1995) are all in pursuit of this higher dimensional uniformity. In this article we show that SLHS, which combines the idea of stratified sampling and LHS, yields an asymptotically more accurate estimate than both stratified sampling and LHS.

In Section 2 we describe how to construct stratified Latin hypercubes. In Section 3 SLHS is shown to give an asymptotically more accurate estimate than both stratified sampling and LHS. Brief discussion of relationship between SLHS, OALHS, and TLHS is given in Section 4.

2. CONSTRUCTION OF STRATIFIED LATIN HYPERCUBES

Let $\mathbf{X} = (X_1, \ldots, X_K)$ be the $U[0,1]^K$ random vector of input variables and $Y = g(\mathbf{X})$ be the observable output variable. Assumption of $\mathbf{X}$ being $U[0,1]^K$ imposes no restriction since any random vector $\mathbf{X}$ with a component distribution function F can be generated from $\mathbf{X} = F^{-1}(\mathbf{U})$, where $\mathbf{U} = (U_1, \ldots, U_K)$ is $U[0,1]^K$, and $g(\mathbf{X})$ is then replaceable with $g(F^{-1}(\mathbf{U}))$. Suppose that we are interested in estimating $\mu = E(Y)$ with the sample mean $\bar{Y} = N^{-1}\sum_{i=1}^{N} g(\mathbf{X}_i)$ based on N values of the input vector, $\mathbf{X}_1, \ldots, \mathbf{X}_N$. Our objective here is to choose N values so that they not only come from each of the same sized strata equally often but also form a Latin hypercube sample. Such a sample would attain stratification on all of bivariate through K-variate margins and thus is named the stratified Latin hypercube sample. Figure 1 shows three different samples of size 8 on two input variables drawn by (a) LHS, (b) stratified sampling, and (c) SLHS. The sample in (a) is evenly scattered on each univariate margin but not on bivariate margin. Conversely, the sample in (b) is evenly scattered on bivariate margin but not on each univariate margin. The sample in (c), on the other hand, is evenly scattered on both each univariate and bivariate margin.

We next describe an algorithm for generating a stratified Latin hypercube sample of size N. It is assumed that each univariate margin is equally divided into M subintervals of length $1/M$ to form $I = M^K$ equal-sized strata over the input space. Equal number, n ($n \geq 2$), of observations are taken from each stratum, so the relationship $N = nM^K$ always holds. The presented algorithm is essentially for achieving exchangeability of $\mathbf{X}_1, \ldots, \mathbf{X}_N$.

Step 1. Randomly permute $N = nM^K$ stratum indices and denote the result by $(a_{1j}, \ldots, a_{Kj})$, $j = 1, \ldots, N$, where each a_{ij} $(i = 1, \ldots, K)$ belongs to $\{1, \ldots, M\}$.

Step 2. Define index sets $R_{mk} = \{1, 2, \ldots, N/M\}$, $m = 1, \ldots, M$; $k = 1, \ldots, K$. For the jth stratum index $(j = 1, \ldots, N)$, randomly select a number p_{ij} without replacement from $R_{a_{ij},j}$, $i = 1, \ldots, K$.

Step 3. Let $X_{ij} = N^{-1}\left[\frac{N}{M}(a_{ij} - 1) + p_{ij} - 1 + Z_{ij}\right]$, $i = 1, \ldots, K$; $j = 1, \ldots, N$, where Z_{ij}'s are independent $U[0,\ 1]$.

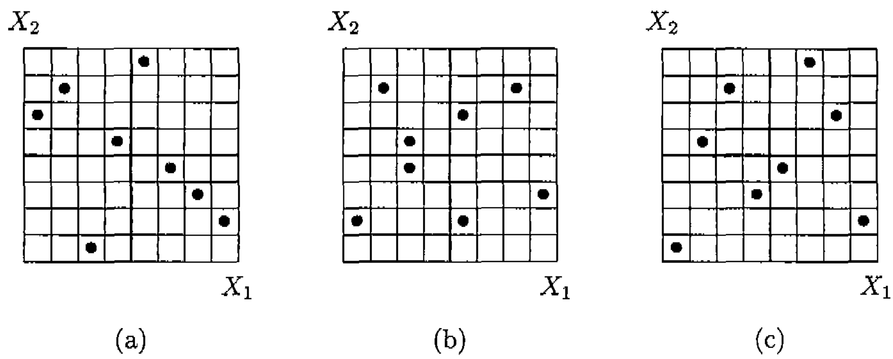

Figure 2.1: Three types of samples drawn by (a) LHS, (b) stratified sampling, and (c) SLHS

3. MAIN RESULTS

Let a cartesian product set $S_i = S_{1i} \times S_{2i} \times \cdots \times S_{Ki}$, $i = 1, \ldots I$, denote the ith stratum where each S_{ij} is one of the intervals $[0, 1/M)$, $[1/M, 2/M)$, $\ldots, [(M-1)/M, 1]$. Also let $c(i,q)$, $i = 1, \ldots, I; q = 1, \ldots, (N/M)^K$, be a hypercube denoting the qth cell in the ith stratum. Variances and covariances under random sampling, stratified sampling, LHS, and SLHS will be respectively denoted by subscripts RS, SS, LHS, and SLHS. The following theorem states that SLHS yields asymptotically smaller variance than stratified sampling.

Theorem 3.1. *If $E(Y^2) < \infty$, then as $N \to \infty$ with K fixed,*

$$N\left(Var_{SLHS}(\bar{Y}) - Var_{SS}(\bar{Y})\right) \longrightarrow -MI^{-2}\sum_{m=1}^{K}\int\left[\sum_{i=1}^{I}(g_{mi}(x) - \mu_i)I_{S_{mi}}(x)\right]^2 M dx,$$

where

$$g_{mi}(x_m) = \int_{\prod_{j\neq m} S_{ji}} g(\mathbf{x}) M^{K-1} \prod_{j\neq m} dx_j.$$

Proof: First we decompose $Var_{SLHS}(\bar{Y})$ into three parts as follows:

$$\begin{aligned}
Var_{SLHS}(\bar{Y}) &= Var_{SLHS}\left(N^{-1}\sum_{i=1}^{I}\sum_{j=1}^{n} Y_{ij}\right) \\
&= N^{-2}\sum_{i=1}^{I}\sum_{j=1}^{n} Var_{SLHS}(Y_{ij}) + N^{-2}\sum_{i=1}^{I}\sum_{j\neq l} Cov_{SLHS}(Y_{ij}, Y_{il}) \\
&\quad + N^{-2}\sum_{i\neq k}\sum_{j=1}^{n}\sum_{l=1}^{n} Cov_{SLHS}(Y_{ij}, Y_{kl}) \\
&\equiv A + B + C.
\end{aligned}$$

Let Y_{ij}, $i = 1, \ldots, I$; $j = 1, \ldots, n$, be the jth observed output drawn from the ith stratum. Then

$$\begin{aligned}
P(Y_{ij} \leq y) &= \sum_{\text{all } c(i,q)} P\left(Y_{ij} \leq y | \mathbf{X} \in c(i,q)\right) P(\mathbf{X} \in c(i,q)) \\
&= \sum_{q}\int_{c(i,q), g(\mathbf{x})\leq y} N^K d\mathbf{x} \cdot \left(\frac{N}{M}\right)^{-K} \\
&= \int_{S_i, g(\mathbf{x})\leq y} M^K d\mathbf{x},
\end{aligned}$$

so the marginal distribution of Y_{ij} drawn by SLHS is the same as that of Y_{ij} drawn by stratified sampling. Therefore

$$A = N^{-2} \cdot n \sum_{i=1}^{I} \sigma_i^2 = I^{-1} N^{-1} \sum_{i=1}^{I} \sigma_i^2 = Var_{SS}(\bar{Y}), \tag{3.1}$$

where

$$\sigma_i^2 = \int_{S_i} (g(\mathbf{x}) - \mu_i)^2 M^K d\mathbf{x} \quad \text{and} \quad \mu_i = \int_{S_i} g(\mathbf{x}) M^K d\mathbf{x}.$$

Here notice that the equality of marginal distributions of Y_{ij} drawn by SLHS and by stratified sampling implies the unbiasedness of $\bar{Y}$ based on SLHS because $\bar{Y}$ based on stratified sampling is known to be unbiased. To obtain the expression for B and C, define for $0 \le x_1, x_2 < 1$,

$$r(x_1, x_2) = \begin{cases} 1 & \text{if } [Nx_1] = [Nx_2] \\ 0 & \text{otherwise,} \end{cases}$$

where $[x]$ is the greatest integer not exceeding x. If x_1 and x_2 belong to the same stratum interval, then

$$\begin{aligned} & P(x_1 < X_{1k} \le x_1 + \Delta x_1,\ x_2 < X_{2k} \le x_2 + \Delta x_2) \\ = & \left[\frac{N}{M}\left(\frac{N}{M} - 1\right)\right]^{-1} N^2 \Delta x_1 \Delta x_2 (1 - r(x_1, x_2)) \\ = & \frac{N}{N-M} M^2 \Delta x_1 \Delta x_2 (1 - r(x_1, x_2)). \end{aligned}$$

On the other hand, if x_1 and x_2 belong to different stratum intervals, then

$$\begin{aligned} & P(x_1 < X_{1k} \le x_1 + \Delta x_1,\ x_2 < X_{2k} \le x_2 + \Delta x_2) \\ = & \left(\frac{N}{M}\right)^{-2} N^2 \Delta x_1 \Delta x_2 (1 - r(x_1, x_2)) \\ = & M^2 \Delta x_1 \Delta x_2 (1 - r(x_1, x_2)). \end{aligned}$$

Therefore the joint pdf of $\mathbf{X}_1$ and $\mathbf{X}_2$ when they have r stratum intervals in common is given by

$$p(\mathbf{x}_1, \mathbf{x}_2) = \left(\frac{N}{N-M}\right)^r M^{2K} \prod_{k=1}^{K} (1 - r(x_{k1}, x_{k2})).$$

Now, if S_i and S_k have r coordinates in common, then

$Cov_{SLHS}(Y_{i1}, Y_{k1})$

$$
\begin{aligned}
&= \int_{S_iS_k} g(\mathbf{x}_1)g(\mathbf{x}_2)\left(\frac{N}{N-M}\right)^r M^{2K}\prod_{m=1}^{K}(1-r(x_{m1},x_{m2}))d\mathbf{x}_1 d\mathbf{x}_2 - \mu_i\mu_k \\
&= \left(\frac{N}{N-M}\right)^r \Bigg[\int_{S_iS_k} g(\mathbf{x}_1)g(\mathbf{x}_2)M^{2K}d\mathbf{x}_1 d\mathbf{x}_2 \\
&\qquad - \sum_{m=1}^{K}\int_{S_iS_k} g(\mathbf{x}_1)g(\mathbf{x}_2)M^{2K}r(x_{m1},x_{m2})d\mathbf{x}_1 d\mathbf{x}_2 + O(N^{-2})\Bigg] - \mu_i\mu_k \\
&= \left(1+\frac{rM}{N}+o(N^{-1})\right)\Bigg[\int_{S_i} g(\mathbf{x}_1)M^K d\mathbf{x}_1 \cdot \int_{S_k} g(\mathbf{x}_2)M^K d\mathbf{x}_2 \\
&\qquad - \frac{M}{N}\sum_{m\in J_{ik}}\int_{S_{mi}} g_{mi}(x)g_{mk}(x)M dx + o(N^{-1})\Bigg] - \mu_i\mu_k \\
&= \frac{rM}{N}\mu_i\mu_k - \frac{M}{N}\sum_{m\in J_{ik}}\int_{S_{mi}} g_{mi}(x)g_{mk}(x)M dx + o(N^{-1}),
\end{aligned}
$$

where $O(N^{-2})$ includes all lower-order terms in the expansion of the product and J_{ik} is the set of m's for which S_{mi} equals S_{mk} . Here the third equality follows from the following argument: If S_i and S_k share the mth coordinate interval in common, then

$$
\begin{aligned}
&\int_{S_iS_k} g(\mathbf{x}_1)g(\mathbf{x}_2)M^{2K}r(x_{m1},x_{m2})d\mathbf{x}_1 d\mathbf{x}_2 \\
&\quad= \int_{S_{mi}S_{mk}} g_{mi}(x_1)g_{mk}(x_2)M^2 r(x_1,x_2)dx_1 dx_2 \\
&\quad= \sum_{l=1}^{N/M}\int_{I_{\bar{m}l}} g_{mi}(x)M dx \cdot \int_{I_{\bar{m}l}} g_{mk}(x)M dx,
\end{aligned}
$$

where $\bar{m}$ is the stratum index corresponding to S_{mi} and $I_{\bar{m}l}$ is the interval $[(\bar{m}-1)M^{-1}+(l-1)N^{-1},\ (\bar{m}-1)M^{-1}+lN^{-1})$. By the mean value theorem, there exist some $x_l',\ x_l'' \in I_{\bar{m}l}$ such that

$$
\begin{aligned}
&\frac{N}{M}\sum_{l=1}^{N/M}\int_{I_{\bar{m}l}} g_{mi}(x)M dx \cdot \int_{I_{\bar{m}l}} g_{mk}(x)M dx \\
&\quad= MN\sum_{l=1}^{N/M}\frac{1}{N}g_{mi}(x_l')\cdot\frac{1}{N}g_{mk}(x_l'') \\
&\quad= \frac{M}{N}\sum_{l=1}^{N/M} g_{mi}(x_l')g_{mk}(x_l'')
\end{aligned}
$$

$$\longrightarrow \int_{S_{mi}} g_{mi}(x) g_{mk}(x) M dx$$

as $N \to \infty$, so that

$$\int_{S_i S_k} g(\mathbf{x}_1) g(\mathbf{x}_2) M^{2K} r(x_{m1}, x_{m2}) d\mathbf{x}_1 d\mathbf{x}_2 = \frac{M}{N} \int_{S_{mi}} g_{mi}(x) g_{mk}(x) M dx + o(N^{-1}),$$

which establishes the third equality. Thus we have

$$\begin{aligned} B &= N^{-2} n(n-1) \sum_{i=1}^{I} Cov_{SLHS}(Y_{i1}, Y_{i2}) \\ &= N^{-2} n(n-1) \sum_{i=1}^{I} \left[\frac{KM}{N} \mu_i^2 - \frac{M}{N} \sum_{m=1}^{K} \int_{S_{mi}} g_{mi}{}^2(x) M dx + o(N^{-1}) \right] \\ &= I^{-1}(I^{-1} - N^{-1}) \left[\frac{KM}{N} \sum_{i=1}^{I} \mu_i^2 - \frac{M}{N} \sum_{i=1}^{I} \sum_{m=1}^{K} \int_{S_{mi}} g_{mi}{}^2(x) M dx \right] + o(N^{-1}) \\ &= -I^{-1}(I^{-1} - N^{-1}) M N^{-1} \sum_{i=1}^{I} \sum_{m=1}^{K} \int_{S_{mi}} (g_{mi}(x) - \mu_i)^2 M dx + o(N^{-1}) \quad (3.2) \end{aligned}$$

and

$$\begin{aligned} C &= N^{-2} n^2 \sum_{i \neq k} Cov_{SLHS}(Y_{i1}, Y_{k1}) \\ &= I^{-2} \sum_{i \neq k} \left[\frac{M}{N} \cdot n(J_{ik}) \cdot \mu_i \mu_k - \frac{M}{N} \sum_{m \in J_{ik}} \int_{S_{mi}} g_{mi}(x) g_{mk}(x) M dx + o(N^{-1}) \right] \\ &= -M I^{-2} N^{-1} \sum_{i \neq k} \sum_{m \in J_{ik}} \int_{S_{mi}} (g_{mi}(x) - \mu_i)(g_{mk}(x) - \mu_k) M dx + o(N^{-1}) \quad (3.3) \end{aligned}$$

where $n(E)$ denotes the number of elements in set E. Therefore from (3.1)-(3.3) the variance of $\bar{Y}$ drawn by SLHS is

$$\begin{aligned} Var_{SLHS}(\bar{Y}) &= Var_{SS}(\bar{Y}) + M I^{-1} N^{-2} \sum_{i=1}^{I} \sum_{m=1}^{K} \int_{S_{mi}} (g_{mi}(x) - \mu_i)^2 M dx \\ &\quad - M I^{-2} N^{-1} \sum_{i=1}^{I} \sum_{k=1}^{K} \sum_{m \in J_{ik}} \int_{S_{mi}} (g_{mi}(x) - \mu_i) \\ &\quad \times (g_{mk}(x) - \mu_k) M dx + o(N^{-1}). \end{aligned}$$

Now, since $E(Y^2) < \infty$ implies $\int_{S_{mi}} (g_{mi}(x) - \mu_i)^2 M dx < \infty$, the second term on the right hand side of the above identity is $o(N^{-1})$. Furthermore,

$$\sum_{i=1}^{I} \sum_{k=1}^{I} \sum_{m \in J_{ik}} \int_{S_{mi}} (g_{mi}(x) - \mu_i)(g_{mk}(x) - \mu_k) M dx$$

$$= \sum_{i=1}^{I}\sum_{k=1}^{I}\sum_{m=1}^{K}\int_{S_{mi}S_{mk}}(g_{mi}(x)-\mu_i)(g_{mk}(x)-\mu_k)M\,dx$$
$$= \sum_{m=1}^{K}\int\left[\sum_{i=1}^{I}(g_{mi}(x)-\mu_i)I_{S_{mi}}(x)\right]^2 M\,dx,$$

so we have

$$Var_{SLHS}(\bar{Y}) = Var_{SS}(\bar{Y}) - MI^{-2}N^{-1}\sum_{m=1}^{K}\int\left[\sum_{i=1}^{I}(g_{mi}(x)-\mu_i)I_{S_{mi}}(x)\right]^2 M\,dx + o(N^{-1}). \tag{3.4}$$

Hence as $N \to \infty$,

$$N\left(Var_{SLHS}(\bar{Y}) - Var_{SS}(\bar{Y})\right) \longrightarrow -MI^{-2}\sum_{m=1}^{K}\int\left[\sum_{i=1}^{I}(g_{mi}(x)-\mu_i)\cdot I_{S_{mi}}(x)\right]^2 M\,dx,$$

which completes the proof. □

The next theorem states that SLHS yields asymptotically smaller variance than LHS.

Theorem 3.2. *If $E(Y^2) < \infty$, then as $N \to \infty$ with K fixed*

$$\begin{aligned} &N\left(Var_{SLHS}(\bar{Y}) - Var_{LHS}(\bar{Y})\right) \\ &\longrightarrow -I^{-1}\sum_{i=1}^{I}(\mu_i-\mu)^2 - \sum_{m=1}^{K}\Big[\sum_{t=1}^{M}\int_{D_t}(\bar{g}_{mt}(x)-\bar{\mu}_{mt})^2 dx \\ &\qquad -\sum_{t=1}^{M}\int_{D_t}(g_m(x)-\mu)^2 dx\Big] \\ &\le -I^{-1}\sum_{i=1}^{I}(\mu_i-\mu)^2, \end{aligned}$$

where

$$D_t = [(t-1)M^{-1}, tM^{-1}], \quad g_m(x_m) = \int g(y)\prod_{j\neq m} dx_j,$$
$$\bar{g}_{mt}(x_m) = MI^{-1}\sum_{i\in T_{mt}} g_{mi}(x_m), \quad \bar{\mu}_{mt} = MI^{-1}\sum_{i\in T_{mt}}\mu_i,$$

and T_{mt} is the set of sample index i for which S_{mi} equals D_t.

Proof: Rewriting the integral on the right hand side of (3.4) as

$$\begin{aligned}
&\int \left[\sum_{i=1}^{I}(g_{mi}(x)-\mu_i)I_{S_{mi}}(x)\right]^2 M dx \\
&= \int \left[\sum_{t=1}^{M}\sum_{i\in T_{mt}}(g_{mi}(x)-\mu_i)I_{D_t}(x)\right]^2 M dx \\
&= \sum_{t=1}^{M}\sum_{i\in T_{mt}}\sum_{k\in T_{mt}}\int_{D_t}(g_{mi}(x)-\mu_i)(g_{mk}(x)-\mu_k)M dx \\
&= \sum_{t=1}^{M}\int_{D_t}\left[\sum_{i\in T_{mt}}(g_{mi}(x)-\mu_i)\right]^2 M dx,
\end{aligned}$$

$Var_{SLHS}(\bar{Y})$ can be written as

$$\begin{aligned}
&Var_{SLHS}(\bar{Y}) \\
&= Var_{SS}(\bar{Y}) - N^{-1}\sum_{m=1}^{K}\sum_{t=1}^{M}\int_{D_t}\left[MI^{-1}\sum_{i\in T_{mt}}(g_{mi}(x)-\mu_i)\right]^2 dx + o(N^{-1}) \\
&= Var_{SS}(\bar{Y}) - N^{-1}\sum_{m=1}^{K}\sum_{t=1}^{M}\int_{D_t}(\bar{g}_{mt}(x)-\bar{\mu}_{mt})^2 dx + o(N^{-1}),
\end{aligned}$$

On the other hand, from Theorem 3.1 of Stein(1987) $Var_{LHS}(\bar{Y})$ can be expressed as

$$\begin{aligned}
Var_{LHS}(\bar{Y}) &= Var_{RS}(\bar{Y}) - N^{-1}\left[\sum_{m=1}^{K}\int g_m^2(x)dx - K\mu^2\right] + o(N^{-1}) \\
&= Var_{RS}(\bar{Y}) - N^{-1}\sum_{m=1}^{K}\sum_{t=1}^{M}\int_{D_t}(g_m(x)-\mu)^2 dx + o(N^{-1}).
\end{aligned}$$

Noting that

$$\sum_{t=1}^{M}\bar{\mu}_{mt} = MI^{-1}\sum_{t=1}^{M}\sum_{i\in T_{mt}}\mu_i = MI^{-1}\sum_{i=1}^{I}\mu_i = M\mu$$

and

$$\sum_{t=1}^{M}\bar{g}_{mt}(x_m) = MI^{-1}\sum_{t=1}^{M}\sum_{i\in T_{mt}} g_{mi}(x_m)$$

$$
\begin{aligned}
&= MI^{-1}\sum_{i=1}^{I}\int_{\prod_{j\neq m}S_{ji}} g(\mathbf{x})M^{K-1}\prod_{j\neq m}dx_j \\
&= \sum_{i=1}^{I}\int_{\prod_{j\neq m}S_{ji}} g(\mathbf{x})\prod_{j\neq m}dx_j \\
&= M\int g(\mathbf{x})\prod_{j\neq m}dx_j \\
&= Mg_m(x_m),
\end{aligned}
$$

we have

$$
\begin{aligned}
\sum_{t=1}^{M}\int_{D_t}(g_m(x)-\mu)^2dx &= \int\sum_{t=1}^{M}\left[M^{-1}\sum_{u=1}^{M}(\bar{g}_{mu}(x)-\bar{\mu}_{mu})\right]^2 I_{D_t}dx \\
&\leq \int\sum_{t=1}^{M}(\bar{g}_{mt}(x)-\bar{\mu}_{mt})^2 I_{D_t}(x)dx \\
&= \sum_{t=1}^{M}\int_{D_t}(\bar{g}_{mt}(x)-\bar{\mu}_{mt})^2dx,
\end{aligned}
$$

where the inequality follows since both $\bar{g}_{mt}(x_m)$ and $\bar{\mu}_{mt}$ equal zero outside D_t by definition. Using the well-known fact

$$
Var_{SS}(\bar{Y}) = Var_{RS}(\bar{Y}) - N^{-1}I^{-1}\sum_{i=1}^{I}(\mu_i-\mu)^2,
$$

we finally have

$$
\begin{aligned}
&N\left(Var_{SLHS}(\bar{Y}) - Var_{LHS}(\bar{Y})\right) \\
&\longrightarrow\ -I^{-1}\sum_{i=1}^{I}(\mu_i-\mu)^2 - \sum_{m=1}^{K}\left[\sum_{t=1}^{M}\int_{D_t}(\bar{g}_{mt}(x)-\bar{\mu}_{mt})^2dx \right.\\
&\qquad\qquad \left. -\sum_{t=1}^{M}\int_{D_t}(g_m(x)-\mu)^2dx\right] \\
&\leq\ -I^{-1}\sum_{i=1}^{I}(\mu_i-\mu)^2. \qquad \square
\end{aligned}
$$

We may express $Var_{SLHS}(\bar{Y})$ in a more interpretable form as in the following corollary.

Corollary 3.1. *Under the same conditions as in Theorem 3.2,*

$$Var_{SLHS}(\bar{Y}) = N^{-1} \int \xi^2(\mathbf{x})d\mathbf{x} + o(N^{-1}),$$

where

$$\xi(\mathbf{x}) = g(\mathbf{x}) - \sum_{i=1}^{I} \mu_i I_{S_i}(\mathbf{x}) - MI^{-1} \sum_{i=1}^{I} \sum_{m=1}^{K} (g_{mi}(x_m) - \mu_i) I_{S_{mi}}(x_m).$$

Proof: We first expand $\xi^2(\mathbf{x})$ in the following way:

$$\begin{aligned}
\xi^2(\mathbf{x}) &= \sum_{i=1}^{I} \left[(g(\mathbf{x}) - \mu_i) I_{S_i}(\mathbf{x}) - MI^{-1} \sum_{m=1}^{K} (g_{mi}(x_m) - \mu_i) I_{S_{mi}}(x_m) \right]^2 \\
&\quad + \sum_{i \neq k} \left[(g(\mathbf{x}) - \mu_i) I_{S_i}(\mathbf{x}) - MI^{-1} \sum_{m=1}^{K} (g_{mi}(x_m) - \mu_i) I_{S_{mi}}(x_m) \right] \\
&\qquad \times \left[(g(\mathbf{x}) - \mu_k) I_{S_k}(\mathbf{x}) - MI^{-1} \sum_{l=1}^{K} (g_{lk}(x_l) - \mu_k) I_{S_{lk}}(x_l) \right] \\
&= \sum_{i=1}^{I} (g(\mathbf{x}) - \mu_i)^2 I_{S_i}(\mathbf{x}) + M^2 I^{-2} \sum_{i=1}^{I} \sum_{m=1}^{K} (g_{mi}(x_m) - \mu_i)^2 I_{S_{mi}}(x_m) \\
&\quad + M^2 I^{-2} \sum_{i=1}^{I} \sum_{m \neq l} (g_{mi}(x_m) - \mu_i)(g_{li}(x_l) - \mu_i) I_{S_{mi}}(x_m) I_{S_{li}}(x_l)
\end{aligned}$$

$$\begin{aligned}
&- 2MI^{-1} \sum_{i=1}^{I} \sum_{m=1}^{K} (g(\mathbf{x}) - \mu_i)(g_{mi}(x_m) - \mu_i) I_{S_i}(\mathbf{x}) I_{S_{mi}}(x_m) \\
&+ \sum_{i \neq k} (g(\mathbf{x}) - \mu_i)(g(\mathbf{x}) - \mu_k) I_{S_i}(\mathbf{x}) I_{S_k}(\mathbf{x}) \\
&- 2MI^{-1} \sum_{i \neq k} \sum_{m=1}^{K} (g_{mi}(x_m) - \mu_i)(g(\mathbf{x}) - \mu_k) I_{S_{mi}}(x_m) I_{S_k}(\mathbf{x}) \\
&+ M^2 I^{-2} \sum_{i \neq k} \sum_{m=1}^{K} \sum_{l=1}^{K} (g_{mi}(x_m) - \mu_i)(g_{lk}(x_l) - \mu_k) I_{S_{mi}}(x_m) I_{S_{lk}}(x_l).
\end{aligned}$$

Since the third and fifth terms on the right hand side of the last equality vanish upon integration with respect to $\mathbf{x}$, we have

$$\int \xi^2(\mathbf{x})d\mathbf{x}$$

$$
\begin{aligned}
&= I^{-1}\sum_{i=1}^{I}\int_{S_i}(g(\mathbf{x})-\mu_i)^2 M^K\prod_{j=1}^{K}dx_j + MI^{-2}\sum_{i=1}^{I}\sum_{m=1}^{K}\int_{S_{mi}}(g_{mi}(x)-\mu_i)^2 M dx \\
&\quad - 2MI^{-2}\sum_{i=1}^{I}\sum_{m=1}^{K}\int_{S_{mi}}(g_{mi}(x)-\mu_i)^2 M dx \\
&\quad - 2MI^{-2}\sum_{i\neq k}\sum_{m=1}^{K}\int_{S_{mi}S_{mk}}(g_{mi}(x)-\mu_i)(g_{mk}(x)-\mu_k) M dx \\
&\quad + MI^{-2}\sum_{i\neq k}\sum_{m=1}^{K}\int_{S_{mi}S_{mk}}(g_{mi}(x)-\mu_i)(g_{mk}(x)-\mu_k) M dx \\
&= I^{-1}\sum_{i=1}^{I}\sigma_i^2 - MI^{-2}\sum_{i=1}^{I}\sum_{m=1}^{K}\int_{S_{mi}}(g_{mi}(x)-\mu_i)^2 M dx \\
&\quad - MI^{-2}\sum_{i\neq k}\sum_{m\in J_{ik}}\int_{S_{mi}}(g_{mi}(x)-\mu_i)(g_{mk}(x)-\mu_k) M dx \\
&= I^{-1}\sum_{i=1}^{I}\sigma_i^2 - MI^{-2}\sum_{i=1}^{I}\sum_{k=1}^{I}\sum_{m\in J_{ik}}\int_{S_{mi}}(g_{mi}(x)-\mu_i)(g_{mk}(x)-\mu_k) M dx \\
&= N\cdot Var_{SLHS}(\bar{Y}) + o(1)
\end{aligned}
$$

and the proof is complete. □

From the defining equation for $\xi(\mathbf{x})$ it can be seen that $\xi(\mathbf{x})$ is the sum of differences between $g(\mathbf{x})-\mu_i$ and its additive fit on S_i, $MI^{-1}\sum_{m=1}^{K}(g_{mi}(x_m)-\mu_i)$. In fact, if we define

$$\tilde{h}(\mathbf{x}) \equiv MI^{-1}\sum_{i=1}^{I}\sum_{m=1}^{K}(g_{mi}(x_m)-\mu_i)I_{S_{mi}}(x_m),$$

then we may see that $\tilde{h}(\mathbf{x})$ is the sum of the best additive fits to $g(\mathbf{x})-\mu_i$ on S_i, $i=1,\ldots,I$. This result is formally stated in the following theorem.

Theorem 3.3. *Let the sum of arbitrary additive fits of $g(\mathbf{x})-\mu_i$ on S_i , $i=1,\ldots,I$, be*

$$h(\mathbf{x}) \equiv \sum_{i=1}^{I}\sum_{k=1}^{K}h_{ki}(x_k)I_{S_{ki}}(x_k)$$

for any set of univariate functions h_{ki} , $k=1,\ldots,K$; $i=1,\ldots,I$. Then

$$\int \xi^2(\mathbf{x})d\mathbf{x} \leq \int\left[g(\mathbf{x})-\sum_{i=1}^{I}\mu_i I_{S_i}(\mathbf{x})-h(\mathbf{x})\right]^2 d\mathbf{x}.$$

Proof: For simplicity define

$$\tilde{h}_i(\mathbf{x}) = MI^{-1}\sum_{m=1}^{K}(g_{mi}(x_m) - \mu_i)I_{S_{mi}}(x_m).$$

Then

$$\begin{aligned}
&\int\left[g(\mathbf{x}) - \sum_{i=1}^{I}\mu_i I_{S_i}(\mathbf{x}) - h(\mathbf{x})\right]^2 d\mathbf{x}\\
&= \int\left[\sum_{i=1}^{I}\left\{(g(\mathbf{x}) - \mu_i)I_{S_i}(\mathbf{x}) - \tilde{h}_i(\mathbf{x}) + \tilde{h}_i(\mathbf{x}) - \sum_{k=1}^{K}h_{ki}(x_k)I_{S_{ki}}(x_k)\right\}\right]^2 d\mathbf{x}\\
&= \int\left[\sum_{i=1}^{I}\left\{(g(\mathbf{x}) - \mu_i)I_{S_i}(\mathbf{x}) - \tilde{h}_i(\mathbf{x})\right\}\right]^2 d\mathbf{x}\\
&\quad + \int\left[\sum_{i=1}^{I}\left\{\tilde{h}_i(\mathbf{x}) - \sum_{k=1}^{K}h_{ki}(x_k)I_{S_{ki}}(x_k)\right\}\right]^2 d\mathbf{x}\\
&\quad + \sum_{i=1}^{I}\sum_{j=1}^{I}\int\left[(g(\mathbf{x}) - \mu_i)I_{S_i}(\mathbf{x}) - \tilde{h}_i(\mathbf{x})\right]\left[\tilde{h}_i(\mathbf{x}) - \sum_{k=1}^{K}h_{kj}(x_k)I_{S_{kj}}(x_k)\right] d\mathbf{x}.
\end{aligned}$$

To complete the proof, it suffices to show that the cross product term on the right hand side of the last equality vanishes. The cross product term may be written as

$$\begin{aligned}
&\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{k=1}^{K}\int\left[(g(\mathbf{x}) - \mu_i)I_{S_i}(\mathbf{x}) - MI^{-1}\sum_{m=1}^{K}(g_{mi}(x_m) - \mu_i)I_{S_{mi}}(x_m)\right]\\
&\qquad\times\left[MI^{-1}(g_{kj}(x_k) - \mu_j) - h_{kj}(x_k)\right] I_{S_{kj}}(x_k)d\mathbf{x}\\
&= \sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{k=1}^{K}\int\left[\left\{MI^{-1}(g_{kj}(x_k) - \mu_j) - h_{kj}(x_k)\right\} I_{S_{kj}}(x_k)\right.\\
&\qquad\left.\times\int\left\{(g(\mathbf{x}) - \mu_i)I_{S_i}(\mathbf{x}) - MI^{-1}\sum_{m=1}^{K}(g_{mi}(x_m) - \mu_i)I_{S_{mi}}(x_m)\right\}\prod_{l\neq k}dx_l\right] dx_k.
\end{aligned}$$

The inner integral on the right hand side of the above equality may be evaluated as

$$M^{-K+1}\int_{\prod_{l\neq k}S_{li}}(g(\mathbf{x}) - \mu_i)I_{S_{ki}}(x_k)M^{K-1}\prod_{l\neq k}dx_l - MI^{-1}\Big[\int(g_{ki}(x_k) - \mu_i)$$

$$
\begin{aligned}
& \times I_{S_{ki}}(x_k) \prod_{l \neq k} dx_l \Big] + \sum_{m \neq k} \int (g_{mi}(x_m) - \mu_i) I_{S_{mi}}(x_m) \prod_{l \neq k} dx_l \\
= \; & MI^{-1}(g_{ki}(x_k) - \mu_i) I_{S_{ki}}(x_k) \\
& - MI^{-1} \left[(g_{ki}(x_k) - \mu_i) I_{S_{ki}}(x_k) + \sum_{m \neq k} M^{-1} \int_{S_{mi}} (g_{mi}(x_m) - \mu_i) M dx_m \right] \\
= \; & 0,
\end{aligned}
$$

so the cross product term vanishes and the proof is complete. □

We may get some insight by comparing Theorem 3.3 with Stein(1987)'s result. Stein showed that LHS yields the best additive fit to $g(\mathbf{x}) - \mu$ among all sums of K univariate functions, whereas Theorem 3.3 says that SLHS yields the sum of best additive fits to $g(\mathbf{x}) - \mu_i$ on S_i, $i = 1, \ldots, I$. Since SLHS optimally approximates $g(\mathbf{x}) - \mu_i$ on each S_i with the sum of K univariate functions, it involves the total of KI univariate functions against the total of only K univariate functions for LHS. Therefore it may be conjectured that SLHS yields a more accurate estimate than LHS as stated in Theorem 3.2.

4. DISCUSSION

Since SLHS, OALHS, and TLHS all achieve high dimensional uniformity using some methods of restriction on LHS, it may be helpful to briefly mention their relationships.

Generally, OALHS depends on the existence of an appropriate orthogonal array (OA). An $N \times K$ matrix $\mathbf{A}$ consisting of s symbols is called an OA of strength r, size N, with K constraints and s levels if each $N \times r$ submatrix of $\mathbf{A}$ contains all possible $1 \times r$ row vectors with the same frequency. The array is denoted by $\text{OA}(N, K, s, r)$. Note that an $N \times K$ Latin hypercube is $\text{OA}(N, K, N, 1)$ from the definition. Although Owen (1992) and Tang (1993) used slightly different methods, their OA-based Latin hypercube is basically obtained by a random permutation of symbols in each column of OA. The resulting matrix possesses uniformity property in each r-variate margin inherent in a strength r OA used for construction. SLHS may be considered as a special case of OALHS in that a stratified Latin hypercube can be generated from an $\text{OA}(N, K, M, K)$. However, since both Owen and Tang only showed that OALHS yields asymptotically smaller variance than LHS when $r = 2$, the asymptotic superiority of SLHS over LHS was not established by their proofs.

It may happen that as K gets large, N becomes impractically large for the relation $N = nM^K$ to hold. If such is the case, we can choose Q ($Q < K$) more important input variables so that $N = nM^Q$ holds for a reasonable value of N. Then a partially stratified Latin hypercube sample is obtained by applying SLHS only to these Q variables with values of the remaining $K - Q$ variables chosen by LHS. This technique may also be applied to OALHS.

TLHS is a more specialized version of SLHS in the sense that some less evenly scattered samples by SLHS are eliminated from consideration. This is achieved by introducing an intermediate sampling unit, called the block, between the stratum and the cell. TLHS also yields an unbiased estimator but its superiority over SLHS was demonstrated only with some numerical examples (see Im et al. (1995)).

REFERENCES

Choi, S. and Lee, J. (1993). "Stratified Latin Hypercube Sampling and Its Application to Uncertainty Analysis of Rod internal Pressure," *Chungnam Journal of Sciences*, **20**, 9-23.

Currin, C., Mitchell, T., Morris, M., 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.

Im, M., Kwon, W., and Lee, J. (1995). "Two-Stage Latin Hypercube Sampling and Its Appliction," *The Korean Journal of Applied Statistics*, **8**, 99-108.

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

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

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

Shewry, M. and Wynn, H. P. (1987). "Maximum Entropy Sampling," *Journal of Applied Statistics*, **14**, 165-170.

Stein, M. (1987). "Large Sample Properties of Simulations Using Latin Hypercube Sampling," *Technometrics*, **29**, 143-151.

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