ON LATIN HYPERCUBE SAMPLING

by

Wei-Liem Loh
Purdue University

Technical Report 93-52#

Department of Statistics
Purdue University

October 1993

*

ON LATIN HYPERCUBE SAMPLING

by

Wei-Liem Loh
Purdue University

Abstract

This paper contains a collection of results on latin hypercube sampling. The first result is a Berry-Esseen type bound for the multivariate central limit theorem of the sample mean $\hat{\mu}_n$ based on a latin hypercube sample. The second gives a necessary and sufficient condition on the convergence rate in the strong law for $\hat{\mu}_n$. Finally motivated by the concept of empirical likelihood, a way of constructing nonparametric confidence regions based on latin hypercube samples is proposed for vector means.

ON LATIN HYPERCUBE SAMPLING[1]

BY WEI-LIEM LOH

Purdue University

This paper contains a collection of results on latin hypercube sampling. The first result is a Berry-Esseen type bound for the multivariate central limit theorem of the sample mean $\hat{\mu}_n$ based on a latin hypercube sample. The second gives a necessary and sufficient condition on the convergence rate in the strong law for $\hat{\mu}_n$. Finally motivated by the concept of empirical likelihood, a way of constructing nonparametric confidence regions based on latin hypercube samples is proposed for vector means.

1 Introduction

In 1979 McKay, Beckman and Conover proposed latin hypercube sampling as an attractive alternative to simple random sampling in computer experiments. The main feature of latin hypercube sampling is that, in contrast to simple random sampling, it simultaneously stratifies on all input dimensions. More precisely, for positive integers d and n, let

(i) π_k, $1 \leq k \leq d$, be random permutations of $\{1, \ldots, n\}$ each uniformly distributed over all the $n!$ possible permutations,

(ii) $U_{i_1,\ldots,i_d,j}$, $1 \leq i_1, \ldots, i_d \leq n$, $1 \leq j \leq d$, be $[0,1]$ uniform random variables and

(iii) that the $U_{i_1,\ldots,i_d,j}$'s and π_k's are all stochastically independent.

A latin hypercube sample of size n (taken from the d-dimensional hypercube $[0,1]^d$) is defined to be $\{X(\pi_1(i), \pi_2(i), \ldots, \pi_d(i)) : 1 \leq i \leq n\}$ where for all $1 \leq i_1, \ldots, i_d \leq n$,

$$\begin{aligned} X_j(i_1, \ldots, i_d) &= (i_j - U_{i_1,\ldots,i_d,j})/n, \quad \forall 1 \leq j \leq d, \\ X(i_1, \ldots, i_d) &= (X_1(i_1, \ldots, i_d), \ldots, X_d(i_1, \ldots, i_d))'. \end{aligned}$$

We remark that no generality is lost in this paper by restricting sampling to the unit hypercube as long as the sampling distribution of interest is a product measure (see for example Owen (1992) page 543).

---

[1]Research supported in part by NSA Grant MDA 904-93-H-3011.

AMS 1980 *subject classifications.* Primary 62D05; secondary 62E20, 62G15.

Key words and phrases. Latin hypercube sampling, Berry-Esseen bound, Stein's method, strong law of large numbers, confidence regions.

In many computer experiments, we are interested in estimating $\mu = E(f \circ X)$ where f is a measurable function from $\mathcal{R}^d$ to $\mathcal{R}^p$ and X is uniformly distributed on $[0,1]^d$. Let

$$\hat{\mu}_n = n^{-1} \sum_{k=1}^{n} f \circ X(\pi_1(k), \pi_2(k), \ldots, \pi_d(k)). \tag{1}$$

Then $\hat{\mu}_n$ is an unbiased estimator for μ. McKay, Beckman and Conover showed that in a great number of instances with $p = 1$, the variance of $\hat{\mu}_n$ is substantially smaller than that of the estimator based on simple random sampling. Stein (1987) further proved that the asymptotic variance of $\hat{\mu}_n$ is less than the asymptotic variance of an analogous estimator based on an independently and identically distributed sample. Recently Owen (1992) showed that the multivariate central limit theorem holds for $\hat{\mu}_n$ when f is a bounded function.

This paper contains a number of results, which we think are of interest in their own rights, all with the underlying theme on the construction of asymptotically valid confidence regions for μ using latin hypercube samples. In particular, Section 2 first shows that the result of Stein (1987), mentioned in the previous paragraph, generalizes naturally and easily to the multivariate setting (see Theorem 1). Also a Berry-Esseen type bound (Theorem 2) is obtained for the multivariate central limit theorem for $\hat{\mu}_n$ under the finiteness of third moments. This gives a "rate" to the asymptotic justification for the use of the contours of constant probability density of a multivariate normal distribution as confidence regions for μ. We remark that in the special case of $d = 2$, this reduces to the classical combinatorial central limit theorem (see for example Hoeffding (1951) and Motoo (1957)). The convergence rate of the combinatorial central limit theorem was investigated by von Bahr (1976), Ho and Chen (1978) and a Berry-Esseen type bound was obtained by Bolthausen (1984) for univariate linear statistics and Bolthausen and Götze (1989) for multivariate statistics.

In Section 3, we obtain a necessary and sufficient condition on the rate of convergence in the strong law of large numbers for $\hat{\mu}_n$. The result (Theorem 3) generalizes a well known theorem of Baum and Katz (1965) for independent and identically distributed random variables to multivariate latin hypercube samples. A special case of their result was considered earlier by Hsu and Robbins (1947) and Erdös (1949). We remark also that the result of Baum and Katz has also been extended by Lai (1977), Hipp (1979) and Peligrad (1985) to a number of other stationary sequences under a variety of mixing conditions.

Motivated by the empirical likelihood ratio confidence regions introduced by Owen (1988), (1990) for independent observations, a way of constructing nonparametric confidence regions based on latin hypercube samples is proposed for vector means in Section 4. Theorem 4 provides conditions for the asymptotic validity of the procedure as well as its convergence rates.

Finally the Appendix contains a number of somewhat technical lemmas that are needed in previous sections.

Throughout this paper, c will denote a generic constant which only depends on d and p, c^* denotes a strictly positive generic constant independent of n, $\|.\|$ be the usual Euclidean metric on $\mathcal{R}^p$, Φ_p the standard p-variate normal distribution and given any measurable function $h : \mathcal{R}^p \to \mathcal{R}$, we write

$$\|h\|_q = \begin{cases} (\int_{\mathcal{R}^p} |h(y)|^q dy)^{1/q} & \text{if } 0 < q < \infty, \\ \text{ess sup}_{y \in \mathcal{R}^p} |h(y)| & \text{if } q = \infty. \end{cases}$$

Also if $x \in \mathcal{R}^p$, then x' denotes the transpose of x and if A is some event, then $I\{A\}$ is its indicator function.

2 Rate of convergence to normality

We shall first show that the result of Stein (1987) mentioned in the Introduction generalizes naturally and easily to many dimensions.

Theorem 1 *Suppose* $E\|f \circ X\|^2 < \infty$. *Let* $\Sigma_{lhs} = Cov(\hat{\mu}_n)$ *and* Σ_{iid} *be the covariance matrix of* $\hat{\mu}_n$ *when the* X*'s are independently and identically distributed, that is*

$$\Sigma_{iid} = n^{-1} E(f \circ X - \mu)(f \circ X - \mu)'.$$

Then as $n \to \infty$, *we have*

$$\begin{aligned} \Sigma_{lhs} &= n^{-1} \int_{[0,1]^d} f_{rem}(x) f'_{rem}(x) dx + o(n^{-1}), \\ \Sigma_{iid} &= n^{-1} \int_{[0,1]^d} f_{rem}(x) f'_{rem}(x) dx \\ &\quad + n^{-1} \sum_{k=1}^{d} \int_0^1 f_{-k}(x_k) f'_{-k}(x_k) dx_k, \end{aligned} \tag{2}$$

where for all $x = (x_1, \ldots, x_d)' \in [0,1]^d$,

$$f_{-k}(x_k) = \int_{[0,1]^{d-1}} [f(x) - \mu] \prod_{j \neq k} dx_j,$$

$$f_{rem}(x) = f(x) - \mu - \sum_{k=1}^{d} f_{-k}(x_k). \tag{3}$$

The proof of the above theorem is deferred to the Appendix and the following corollary is an immediate consequence of Theorem 1.

Corollary 1 $\Sigma_{iid} - \Sigma_{lhs}$ *is asymptotically positive semidefinite, i.e.*

$$\lim_{n\to\infty} n\xi'(\Sigma_{iid} - \Sigma_{lhs})\xi \geq \sum_{k=1}^{d} \int_0^1 \xi' f_{-k}(x_k) f'_{-k}(x_k) \xi dx_k \geq 0, \quad \forall \xi \in \mathcal{R}^p.$$

Suppose that $\int_{[0,1]^d} f_{rem}(x) f'_{rem}(x) dx$ is nonsingular. Then it follows from Theorem 1 that for sufficiently large n, $\Sigma_{lhs}^{-1/2}$ exists and define

$$W = \Sigma_{lhs}^{-1/2}(\hat{\mu}_n - \mu). \tag{4}$$

The rest of this section is devoted to establishing a Berry-Esseen type bound for the rate of convergence of W to the standard p-variate normal distribution Φ_p. To do so, we shall make extensive use of the multivariate normal version of Stein's method [see Stein (1972), (1986)] as given in Götze (1991) and Bolthausen and Götze (1989). Let

$$Ef \circ X(i_1, \ldots, i_d) = \mu(i_1, \ldots, i_d), \quad \forall 1 \leq i_1, \ldots, i_d \leq n,$$

$$\mu_{-k}(i_k) = (1/n^{d-1}) \sum_{j\neq k} \sum_{i_j=1}^{n} \mu(i_1, \ldots, i_d),$$

and

$$Y(i_1, \ldots, i_d) = n^{-1}\Sigma_{lhs}^{-1/2}[f \circ X(i_1, \ldots, i_d) - \sum_{k=1}^{d} \mu_{-k}(i_k) + (d-1)\mu]. \tag{5}$$

Then we have $W = \sum_{i=1}^{n} Y(\pi_1(i), \pi_2(i), \ldots, \pi_d(i))$. Next let $\mathcal{A}$ be a class of measurable functions from $\mathcal{R}^p \to \mathcal{R}$ such that $\|g\|_\infty \leq 1$ for all $g \in \mathcal{A}$. Also for $g \in \mathcal{A}$ and $\delta > 0$, define

$$g_\delta^+(w) = \sup\{g(w+y) : \|y\| \leq \delta\}, \quad \forall w \in \mathcal{R}^p,$$

$$g_\delta^-(w) = -\inf\{g(w+y) : \|y\| \leq \delta\}, \quad \forall w \in \mathcal{R}^p,$$

$$\omega(g, \delta) = \int_{\mathcal{R}^p} [g_\delta^+(y) - g_\delta^-(y)] d\Phi_p(y).$$

We further assume that $\mathcal{A}$ is closed under supremum and affine transformations, i.e. $g \in \mathcal{A}$ implies $g_\delta^+ \in \mathcal{A}$, $g_\delta^- \in \mathcal{A}$ and $g \circ T \in \mathcal{A}$ whenever $T : \mathcal{R}^p \to \mathcal{R}^p$ is affine and that there exists a constant $\gamma > 0$ such that

$$\sup\{\omega(g,\delta) : g \in \mathcal{A}\} \leq \gamma\delta, \qquad \forall \delta > 0.$$

REMARK. $\mathcal{A}$ includes the class of all indicator functions of measurable convex sets in $\mathcal{R}^p$ if $\gamma \geq 2\sqrt{p}$ (see for example Bolthausen and Götze (1989)).

Theorem 2 *Suppose $\int_{[0,1]^d} f_{rem}(x) f'_{rem}(x)dx$ is nonsingular. Then there exists a positive constant $C_{d,p}$ which depends only on d and p such that for sufficiently large n,*

$$\sup\{|Eg(W) - \int_{\mathcal{R}^p} g(x)d\Phi_p(x)| : g \in \mathcal{A}\} \leq C_{d,p}\beta_3, \tag{6}$$

where $\beta_3 = (1/n^{d-1}) \sum_{1 \leq i_1,\ldots,i_d \leq n} E\|Y(i_1,\ldots,i_d)\|^3$.

The following is an immediate corollary.

Corollary 2 *Suppose $E\|f \circ X\|^3 < \infty$ and $\int_{[0,1]^d} f_{rem}(x) f'_{rem}(x)dx$ is nonsingular. Then*

$$\sup\{|Eg(W) - \int_{\mathcal{R}^p} g(x)d\Phi_p(x)| : g \in \mathcal{A}\} \leq c^* n^{-1/2}.$$

In order to prove Theorem 2, we first need some preliminary results. For $h \in \mathcal{A}$ and $0 \leq t < 1$, define

$$\chi_t(w|h) = \int_{\mathcal{R}^p} \{h(y) - h(t^{1/2}y + (1-t)^{1/2}w)\}\Phi_p(dy), \tag{7}$$

$$\psi_t(w) = \frac{1}{2}\int_t^1 \chi_s(w|h)\frac{ds}{1-s}. \tag{8}$$

Then $-\chi_0(w|h) = h(w) - \Phi_p(h)$ and $\chi_t(w|h)$ is a smooth approximation of $\chi_0(w|h)$ for small t. (Here $\Phi_p(h) = Eh(Z)$ where Z is a random vector having distribution Φ_p.) The following two lemmas are due to Götze (1991) and we refer the reader to his paper for the proofs.

Lemma 1 *For $0 < \varepsilon < 1$ and $w = (w_1,\ldots,w_p)' \in \mathcal{R}^p$, we have*

$$\sum_{i=1}^p \frac{\partial^2}{\partial w_i^2}\psi_{\varepsilon^2}(w) - \sum_{i=1}^p w_i \frac{\partial}{\partial w_i}\psi_{\varepsilon^2}(w) = -\chi_{\varepsilon^2}(w|h), \tag{9}$$

and there exists a positive constant c_p, depending only on p, such that

$$\sup_{1\le i,j,k\le p} \left| \int_{\mathcal{R}^p} \frac{\partial^3}{\partial w_i \partial w_j \partial w_k} \psi_{\varepsilon^2}(w) Q(dw) \right|$$
$$\le c_p \varepsilon^{-1} \sup\{ \left| \int_{\mathcal{R}^p} h(tw+y) Q(dw) \right| : 0 \le t \le 1, y \in \mathcal{R}^p \}, \tag{10}$$

for all finite signed measures Q on $\mathcal{R}^p$ satisfying $Q(\mathcal{R}^p)=0$.

Lemma 2 *Let Q be a probability distribution on $\mathcal{R}^p$ and $\varepsilon > 0$. Then*

$$\sup_{g\in\mathcal{A}} \left| \int_{\mathcal{R}^p} g(w)[dQ(w) - d\Phi_p(w)] \right| \le \frac{4}{3} \sup_{h\in\mathcal{A}} \left| \int_{\mathcal{R}^p} \chi_{\varepsilon^2}(w|h) dQ(w) \right| + \frac{5\varepsilon\gamma a}{2(1-\varepsilon^2)},$$

where a^2 is the 7/8-quantile of the chi-square distribution with p degrees of freedom.

Proof of Theorem 2. Let n_0 and ε_0 be arbitrary but fixed positive constants. We observe that the theorem is true if $n \le n_0$ or $\beta_3 > \varepsilon_0$. Hence without loss of generality, we shall assume that $n > n_0$ and $\beta_3 \le \varepsilon_0$ for positive constants $n_0 > d^2$ and ε_0 to be suitably chosen later. Next we consider the following combinatorial construction inspired by that given in Bolthausen (1984). Let $(I_i, J_{j,j,k} : 1 \le i,k \le d, 2 \le j \le d)$ be a random element in $\{1,\ldots,n\}^{d^2}$ such that

(i) $(I_i, J_{j,j,1} : 1 \le i \le d, 2 \le j \le d)$ is uniformly distributed on $\{1,\ldots,n\}^{2d-1}$,

(ii) given $(I_i, J_{j,j,k} : 1 \le i \le d, 2 \le j \le d, 1 \le k \le k_0 < d)$, we set $J_{j,j,k_0+1} = J_{j,j,k}$ for all $2 \le j \le d$ if $I_{k_0+1} = I_k$ for some $1 \le k \le k_0$; otherwise J_{j,j,k_0+1} is independently uniformly distributed on $\{1,\ldots,n\}/\{J_{j,j,k} : 1 \le k \le k_0\}$.

Let $\pi_2^{(1)},\ldots,\pi_d^{(1)}$ be independent random permutations (each uniformly distributed on the permutations of $\{1,\ldots,n\}$) which are also independent of $\{I_i, J_{j,j,k} : 1 \le i,k \le d, 2 \le j \le d\}$. Define for $2 \le i,j \le d$ and $1 \le k \le d$,

$$\begin{aligned} L_{j,k} &= [\pi_j^{(1)}]^{-1}(J_{j,j,k}), \\ J_{i,j,k} &= \pi_i^{(1)}(L_{j,k}), \\ J_{i,1,k} &= \pi_i^{(1)}(I_k). \end{aligned}$$

Let $1 \le i_1,\ldots,i_d,j_1,\ldots,j_d \le n$ and $\beta(i_1,\ldots,i_d,j_1,\ldots,j_d)$ be a permutation of $\{1,\ldots,n\}$ leaving the numbers outside $\{i_1,\ldots,i_d,j_1,\ldots,j_d\}$ unchanged such that for each $1 \le k \le d$, $i_k \mapsto j_k$ if $i_k \ne i_j$ for all $1 \le j < k$. Also let $\tau(i,j)$ represent the permutation of $\{1,\dots,n\}$ which transposes i and j leaving other numbers fixed. Now define for $2 \le j \le d$,

$$\begin{aligned}
\pi_j^{(2)} &= \pi_j^{(1)} \circ \beta(I_j, I_1, \dots, I_{j-1}, I_{j+1}, \dots, I_d, L_{j,1}, L_{j,j}, L_{j,2}, \dots, L_{j,d}), \\
\pi_j^{(3)} &= \pi_j^{(2)} \circ \tau(I_1, I_j).
\end{aligned}$$

Finally define

$$W^{(j)} = \sum_{i=1}^{n} Y(i, \pi_2^{(j)}(i), \dots, \pi_d^{(j)}(i)), \quad \forall 1 \le j \le 3.$$

We observe from Lemma 4 (see Appendix) that by choosing ε_0 sufficiently small, we can without loss of generality assume that with probability 1,

$$\|Y(i_1, \dots, i_d)\| \le 1, \quad \forall 1 \le i_1, \dots, i_d \le n. \tag{11}$$

We shall now use an induction argument to prove the theorem by assuming that (6) holds for all values of n less than the current value now being considered. Then writing $W = (W_1, \dots, W_p)'$, we have for $\varepsilon > 0$,

$$\begin{aligned}
& E[\sum_{i=1}^{p} W_i \frac{\partial}{\partial w_i} \psi_{\varepsilon^2}(W)] \\
= \; & E[\sum_{i=1}^{p} W_i^{(3)} \frac{\partial}{\partial w_i} \psi_{\varepsilon^2}(W^{(3)})] \\
= \; & E\{\sum_{i=1}^{p} nY_i(I_1, \pi_2^{(3)}(I_1), \dots, \pi_d^{(3)}(I_1)) \frac{\partial}{\partial w_i} \psi_{\varepsilon^2}(W^{(3)})\} \\
= \; & E \sum_{i=1}^{p} nY_i(I_1, J_{2,2,1}, \dots, J_{d,d,1}) \{ \frac{\partial}{\partial w_i} \psi_{\varepsilon^2}(W^{(2)}) \\
& + \sum_{j=1}^{p} (W^{(3)} - W^{(2)})_j \int_0^1 \frac{\partial^2}{\partial w_i \partial w_j} \psi_{\varepsilon^2}(W^{(2)} + t(W^{(3)} - W^{(2)})) dt \}.
\end{aligned}$$

By construction, $\{I_1, J_{2,2,1}, \dots, J_{d,d,1}\}$ and $\{\pi_2^{(2)}, \dots, \pi_d^{(2)}\}$ are independent. Consequently we have

$$E \sum_{i=1}^{p} nY_i(I_1, J_{2,2,1}, \dots, J_{d,d,1}) \frac{\partial}{\partial w_i} \psi_{\varepsilon^2}(W^{(2)}) = 0,$$

and hence

$$E[\sum_{i=1}^{p} W_i \frac{\partial}{\partial w_i}\psi_{\varepsilon^2}(W) - \sum_{i=1}^{p} \frac{\partial^2}{\partial w_i^2}\psi_{\varepsilon^2}(W)] = R_1 + R_2, \tag{12}$$

where

$$\begin{aligned} R_1 &= E\sum_{i,j=1}^{p}\{nY_i(I_1, J_{2,2,1}, \ldots, J_{d,d,1})(W^{(3)} - W^{(2)})_j \\ &\quad -\delta_{i,j}\}\frac{\partial^2}{\partial w_i \partial w_j}\psi_{\varepsilon^2}(W^{(1)}), \end{aligned}$$

$\delta_{i,j}$ being the Kronecker delta, and

$$\begin{aligned} R_2 &= E\sum_{i,j=1}^{p} nY_i(I_1, J_{2,2,1}, \ldots, J_{d,d,1})(W^{(3)} - W^{(2)})_j \\ &\quad \times \int_0^1 \frac{\partial^2}{\partial w_i \partial w_j}\{\psi_{\varepsilon^2}(W^{(2)} + t(W^{(3)} - W^{(2)})) - \psi_{\varepsilon^2}(W^{(1)})\}dt. \end{aligned}$$

We observe from Lemma 5 (see Appendix) that by choosing ε_0 sufficiently small, we have

$$\sup\{|R_1 + R_2| : h \in \mathcal{A}\} \leq c\beta_3(1 + \varepsilon^{-1}C_{d,p}\beta_3). \tag{13}$$

Now it follows from Lemma 2, (9), (12) and (13) that

$$\sup_{g\in\mathcal{A}} |Eg(W) - \int_{\mathcal{R}^p} g(x)d\Phi_p(x)| \leq c\beta_3(1 + \varepsilon^{-1}C_{d,p}\beta_3) + c\varepsilon. \tag{14}$$

Choosing $\varepsilon = 2c\beta_3$ and $C_{d,p} \geq 2c(2c+1)$ in (14) proves Theorem 2. □

3 Convergence rates in the strong law

Let $\{f \circ X(\pi_1(k), \ldots, \pi_d(k)) : 1 \leq k \leq n\}$ and $\hat{\mu}_n$ be as in (1). In this section, we shall study the rate of convergence in the strong law of large numbers for $\hat{\mu}_n$.

Theorem 3 *Let $\alpha > 1/2$, $\alpha q > 1$ and assume that $E(f \circ X) = 0$ if $\alpha \leq 1$. Then a necessary and sufficient condition for*

$$\sum_{n=1}^{\infty} n^{\alpha q - 2} P(\sup_{1\leq j\leq n} \|j\hat{\mu}_j\| \geq \varepsilon n^{\alpha}) < \infty \quad \forall \varepsilon > 0, \tag{15}$$

is $E\|f \circ X\|^q < \infty$.

The following corollary, which follows directly from the above theorem and the Borel-Cantelli lemma, is needed in the next section.

Corollary 3 *Suppose* $E\|f \circ X\|^2 < \infty$. *Then* $\|\hat{\mu}_n - \mu\| \to 0$ *almost surely as* $n \to \infty$.

PROOF OF THEOREM 3. Without loss of generality, we assume throughout this proof that $E(f \circ X) = 0$ if the expectation exists.

SUFFICIENCY. We first suppose that $E\|f \circ X\|^q < \infty$. For $\varepsilon > 0$, we define as in Erdös (1949) and Katz (1963),

$$\begin{aligned} a_i &= P(\|f \circ X\| > \varepsilon 2^{i\alpha}) \quad \forall i \geq 0, \\ f^+(x) &= \begin{cases} f(x) & \text{if } \|f(x)\| \leq \varepsilon n^{\theta\alpha}, \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

and $\tilde{f}^+ = f^+ - E(f^+ \circ X)$, where θ satisfies $(\alpha q+1)/(2\alpha q) < \theta < 1$, $\theta\alpha q > 1$, and $2\theta\alpha > 1$. For $2^i \leq n < 2^{i+1}$, we write

$$\begin{aligned} A_n &= \{\sup_{1\leq j\leq n} \|j\hat{\mu}_j\| \geq \varepsilon n^\alpha\}, \\ A_n^{(1)} &= \{\|f \circ X(\pi_1(k), \ldots, \pi_d(k))\| \geq \varepsilon 2^{(i-2)\alpha} \text{ for at least one } k \leq n\}, \\ A_n^{(2)} &= \{\|f \circ X(\pi_1(k_1), \ldots, \pi_d(k_1))\| \geq \varepsilon n^{\theta\alpha}, \\ &\quad\ \|f \circ X(\pi_1(k_2), \ldots, \pi_d(k_2))\| \geq \varepsilon n^{\theta\alpha} \text{ for at least two } k_1, k_2 \leq n\}, \\ A_n^{(3)} &= \{\sup_{1\leq j\leq n} \|\sum_{k=1}^{j} f^+ \circ X(\pi_1(k), \ldots, \pi_d(k))\| \geq \varepsilon 2^{(i-2)\alpha}\}. \end{aligned}$$

We observe that

$$A_n \subseteq A_n^{(1)} \cup A_n^{(2)} \cup A_n^{(3)}.$$

Hence to prove (15), it suffices to show that

$$\sum_{n=1}^{\infty} n^{\alpha q-2} P(A_n^{(j)}) < \infty, \quad \forall 1 \leq j \leq 3.$$

Since $E\|f \circ X\|^q < \infty$ is equivalent to $\sum_{i=0}^{\infty} 2^{i\alpha q} a_i < \infty$, we note that

$$\sum_{n=1}^{\infty} n^{\alpha q-2} P(A_n^{(1)}) \leq \sum_{i=0}^{\infty} \sum_{n=2^i}^{2^{i+1}} n^{\alpha q-2} 2^{i+1} a_{i-2} < \infty.$$

Next we observe that

$$\begin{aligned}\sum_{n=1}^{\infty} n^{\alpha q-2} P(A_n^{(2)}) &\leq \sum_{n=1}^{\infty} n^{\alpha q} P(\{\|f\circ X(\pi_1(1),\ldots,\pi_d(1))\| \geq \varepsilon n^{\theta\alpha}\} \cap \\ &\qquad \{\|f\circ X(\pi_1(2),\ldots,\pi_d(2))\| \geq \varepsilon n^{\theta\alpha}\}) \\ &\leq \sum_{n=1}^{\infty} c n^{\alpha q} P^2(\|f\circ X\| \geq \varepsilon n^{\theta\alpha}) \\ &\leq \sum_{n=1}^{\infty} c^* n^{\alpha q - 2\theta\alpha q}(E\|f\circ X\|^q)^2 < \infty. \qquad (16)\end{aligned}$$

Here the second inequality of (16) follows from an argument similar to that given in the proof of Theorem 1 and the third inequality uses Markov's inequality.

To show $\sum_{n=1}^{\infty} n^{\alpha q-2} P(A_n^{(3)}) < \infty$, we first consider the case $0 < q < 1$. Then for $q < q+\delta < 1$, we have

$$\begin{aligned}&\sum_{n=1}^{\infty} n^{\alpha q-2} P(A_n^{(3)}) \\ \leq\ & c^* \sum_{n=1}^{\infty} n^{-2-\delta\alpha} E\{\sup_{1\leq j\leq n} \|\sum_{k=1}^{j} f^+\circ X(\pi_1(k),\ldots,\pi_d(k))\|^{q+\delta}\} \\ \leq\ & c^* \sum_{n=1}^{\infty} n^{(\theta-1)\delta\alpha-1} E\|f\circ X\|^q < \infty.\end{aligned}$$

Next suppose that $q \geq 1$. Let $q^* = \lceil q \rceil$, the smallest integer greater than or equal to q, and m be a positive integer satisfying

$$mq^*(2\alpha - 1) > \alpha q - 1. \qquad (17)$$

Since $E(f\circ X) = 0$, we observe that for $q > 1$,

$$\begin{aligned}n^{1-\alpha}\|Ef^+\circ X\| &\leq n^{1-\alpha} E\|f\circ X\| I\{\|f\circ X\| > \varepsilon n^{\theta\alpha}\} \\ &\leq n^{1-\alpha}\{E\|f\circ X\|^q\}^{1/q}\{P(\|f\circ X\| > \varepsilon n^{\theta\alpha})\}^{1-1/q} \\ &\leq c^* n^{1-\theta\alpha q-(1-\theta)\alpha}\end{aligned}$$

which tends to 0 as $n \to \infty$. Hence from Markov's inequality and Lemma 6 (see Appendix), we obtain

$$\sum_{n=1}^{\infty} n^{\alpha q-2} P(A_n^{(3)})$$

$$
\begin{aligned}
&\le \sum_{n=1}^{\infty} n^{\alpha q-2} P(\sup_{1\le j\le n} \|\sum_{k=1}^{j} \tilde{f}^{+}\circ X(\pi_1(k),\ldots,\pi_d(k))\| \ge c^* n^{\alpha})\\
&\le c^* \sum_{n=1}^{\infty} n^{\alpha q-2-2m\alpha q^*} E\{\sup_{1\le j\le n} \|\sum_{k=1}^{j} \tilde{f}^{+}\circ X(\pi_1(k),\ldots,\pi_d(k))\|^{2mq^*}\}\\
&\le c^* \sum_{n=1}^{\infty} n^{\alpha q-2-2m\alpha q^*} (\log_2 n)^{2mq^*} \{ \sum_{i=2mq^*-q^*+1}^{2mq^*-2} n^{2mq^*/(2mq^*-i)}\\
&\quad + \sum_{i=0}^{(2mq^*-q^*)\wedge(2mq^*-2)} n^{2mq^*/(2mq^*-i)} (E\|\tilde{f}^{+}\circ X\|^{2mq^*-i})^{2mq^*/(2mq^*-i)}\}. \qquad (18)
\end{aligned}
$$

Now we observe from (17) that

$$
\begin{aligned}
&\sum_{n=1}^{\infty} n^{\alpha q-2-2m\alpha q^*} (\log_2 n)^{2mq^*} \sum_{i=0}^{(2mq^*-q^*)\wedge(2mq^*-2)} n^{2mq^*/(2mq^*-i)}\\
&\qquad \times (E\|\tilde{f}^{+}\circ X\|^{2mq^*-i})^{2mq^*/(2mq^*-i)}\\
&\le c^* \sum_{n=1}^{\infty} n^{-1-\alpha(1-\theta)(2mq^*-q)} (\log_2 n)^{2mq^*} < \infty, \qquad (19)
\end{aligned}
$$

and

$$
\sum_{n=1}^{\infty} n^{\alpha q-2-2m\alpha q^*} (\log_2 n)^{2mq^*} \sum_{i=2mq^*-q^*+1}^{2mq^*-2} n^{2mq^*/(2mq^*-i)} < \infty. \qquad (20)
$$

(15) now follows from (18), (19) and (20).

NECESSITY. Suppose now that (15) holds. We observe as in Lai (1977) and Peligrad (1985) that for all $\varepsilon > 0$,

$$
\begin{aligned}
&\sum_{n=1}^{\infty} n^{\alpha q-2} P(\max_{1\le j\le n} \|f\circ X(\pi_1(j),\ldots,\pi_d(j))\| \ge 2\varepsilon n^{\alpha})\\
&\le \sum_{n=1}^{\infty} n^{\alpha q-2} P(\max_{1\le j\le n} \|j\hat{\mu}_j\| \ge \varepsilon n^{\alpha}) < \infty, \qquad (21)
\end{aligned}
$$

and hence

$$
\begin{aligned}
&n^{\alpha q-1} P(\max_{1\le j\le n} \|f\circ X(\pi_1(j),\ldots,\pi_d(j))\| \ge \varepsilon n^{\alpha})\\
&\le c^* \sum_{k=n}^{2n} k^{\alpha q-2} P(\max_{1\le j\le k} \|f\circ X(\pi_1(j),\ldots,\pi_d(j))\| \ge \varepsilon (k/2)^{\alpha}) \to 0, \qquad (22)
\end{aligned}
$$

as $n \to \infty$. Next we shall prove that

$$nP(\|f \circ X\| \geq \varepsilon n^{\alpha}) \to 0 \quad \text{as } n \to \infty. \tag{23}$$

If $\alpha q \geq 2$, (23) follows from (22). Now suppose $1 < \alpha q < 2$. Define $n^* = \lfloor n^{2-\alpha q} \rfloor$. We note from the definition of latin hypercube sampling that for $1 \leq k \leq n^*$,

$$\begin{aligned}
& P(\{\|f \circ X(\pi_1(k), \ldots, \pi_d(k))\| \geq \varepsilon n^{\alpha}\} \cap \\
& \qquad \{\max_{1 \leq j < k} \|f \circ X(\pi_1(j), \ldots, \pi_d(j))\| \geq \varepsilon n^{\alpha}\}) \\
= \; & \{\prod_{i=1}^{k-1} (\frac{n}{n-i})^d\} \int_{[0,1]^{kd}} I\{\|f(x^{(k)})\| \geq \varepsilon n^{\alpha}\} I\{\max_{1 \leq j < k} \|f(x^{(j)})\| \geq \varepsilon n^{\alpha}\} \\
& \times \prod_{1 \leq a < b \leq k} \prod_{l=1}^{d} (1 - \delta_n(x_l^{(a)}, x_l^{(b)})) dx^{(1)} \cdots dx^{(k)} \\
\leq \; & c^* P(\|f \circ X\| \geq \varepsilon n^{\alpha}) P(\max_{1 \leq j \leq n^*} \|f \circ X(\pi_1(j), \ldots, \pi_d(j))\| \geq \varepsilon n^{\alpha}). \qquad (24)
\end{aligned}$$

Since

$$P(\max_{1 \leq j \leq n^*} \|f \circ X(\pi_1(j), \ldots, \pi_d(j))\| \geq \varepsilon n^{\alpha}) \to 0 \quad \text{as } n \to \infty,$$

we observe from (24) that

$$\begin{aligned}
& P(\max_{1 \leq j \leq n^*} \|f \circ X(\pi_1(j), \ldots, \pi_d(j))\| \geq \varepsilon n^{\alpha}) \\
= \; & \sum_{k=1}^{n^*} \{P(\|f \circ X\| \geq \varepsilon n^{\alpha}) - P(\{\|f \circ X(\pi_1(k), \ldots, \pi_d(k))\| \geq \varepsilon n^{\alpha}\} \cap \\
& \qquad \{\max_{1 \leq j < k} \|f \circ X(\pi_1(j), \ldots, \pi_d(j))\| \geq \varepsilon n^{\alpha}\})\} \\
\geq \; & c^* n^* P(\|f \circ X\| \geq \varepsilon n^{\alpha}). \qquad (25)
\end{aligned}$$

Now (23) follows from (22), (25) and the definition of n^*. Finally as in (25), we observe from (23) that

$$\begin{aligned}
& P(\max_{1 \leq j \leq n} \|f \circ X(\pi_1(j), \ldots, \pi_d(j))\| \geq \varepsilon n^{\alpha}) \\
\geq \; & \sum_{k=1}^{n} P(\|f \circ X\| \geq \varepsilon n^{\alpha})[1 - c(k-1)P(\|f \circ X\| \geq \varepsilon n^{\alpha})] \\
\geq \; & c^* n P(\|f \circ X\| \geq \varepsilon n^{\alpha}). \qquad (26)
\end{aligned}$$

Consequently from (21) and (26), we obtain

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{\alpha q-1} P(\|f \circ X\| \geq \varepsilon n^{\alpha}) \\
\leq\ & c^* \sum_{n=1}^{\infty} n^{\alpha q-2} P(\max_{1\leq j\leq n} \|f \circ X(\pi_1(j), \ldots, \pi_d(j))\| \geq \varepsilon n^{\alpha}) < \infty,
\end{aligned}$$

which implies that $E\|f \circ X\|^q < \infty$. This proves Theorem 3. □

4 Nonparametric confidence regions

Owen (1988), (1990) in a series of papers introduced a method of constructing asymptotically valid nonparametric confidence regions for vector-valued statistical functionals using independent and identically distributed observations. In this section we shall show that this method can be readily adaptable to latin hypercube sampling as well.

Let $\{f \circ X(\pi_1(k), \ldots, \pi_d(k)) : 1 \leq k \leq n\}$ be as in (1), $\mu = E(f \circ X)$ and $\mathcal{S} = \{w = (w_1, \ldots, w_n)' : \sum_{k=1}^{n} w_k \leq 1, w_k \geq 0\ \forall k\}$. Define for $0 < r < 1$,

$$(27)\quad \Theta_{n,r} = \{\sum_{k=1}^{n} w_k f \circ X(\pi_1(k), \ldots, \pi_d(k)) : w \in \mathcal{S}, \prod_{k=1}^{n} n w_k \geq r\},$$

$$(28)\quad R_n(\mu) = \sup \prod_{k=1}^{n} n w_k,$$

where the supremum is over $w \in \mathcal{S}$ satisfying

$$\sum_{k=1}^{n} w_k[f \circ X(\pi_1(k), \ldots, \pi_d(k)) - \mu] = 0.$$

REMARK. A very nice discussion of the motivation (in terms of empirical likelihood ratio) for the above construction is given by Owen (1990) in the context of independent observations. In the case of latin hypercube sampling, due to the inherent dependence among the observations, the motivation is less clear. The main motivation here is that this formulation is mathematically tractable. Also intuitively, we can think of latin hypercube samples as a subset of the set of all possible samples of that size by leaving out the non-representative ones, that is, those that are not "evenly distributed" over $[0, 1]^d$. Since empirical likelihood ratio confidence regions work remarkably well for simple random samples, it is plausible that they also perform well for a smaller more representative subset (for instance, latin hypercube samples). However this is just a heuristic, the final justification rests with the results of the theorem below which give an asymptotic validity as well as convergence rates to the procedure.

Theorem 4 *Let* $\{f \circ X(\pi_1(k), \ldots, \pi_d(k)) : 1 \leq k \leq n\}$ *be as in (1) with* $\mu = E(f \circ X)$*. Also let* $\Theta_{n,r}$ *be as in (27) for some* $0 < r < 1$*, and* $f_{rem}(x)$ *be as in (3) such that* $\int_{[0,1]^d} f_{rem}(x) f'_{rem}(x) dx$ *is nonsingular. Then* $\Theta_{n,r}$ *is a convex set.*

(a) If $E\|f \circ X\|^4 < \infty$*, we have*

$$(29) \quad \lim_{n\to\infty} P(\mu \in \Theta_{n,r}) = P(Z'M^{1/2}(M+N)^{-1}M^{1/2}Z \leq -2\log r) \geq P(\chi^2_{(p)} \leq -2\log r),$$

where Z *denotes the random vector having distribution* Φ_p*,*

$$M = \int_{[0,1]^d} f_{rem}(x) f'_{rem}(x) dx,$$

$$N = \sum_{k=1}^{d} \int_0^1 f_{-k}(x_k) f'_{-k}(x_k) dx_k,$$

with $f_{-k}(x_k)$ *as in (3) and* $\chi^2_{(p)}$ *denotes the random variable having the chi-square distribution with p degrees of freedom..*

(b) If $E\|f \circ X\|^6 < \infty$*, we have*

$$(30) \quad |P(\mu \in \Theta_{n,r}) - P(Z'\Sigma_{lhs}^{1/2}\Sigma_{iid}^{-1}\Sigma_{lhs}^{1/2} Z \leq -2\log r)| \leq c^* n^{\varepsilon - 1/2}, \quad \forall 0 < \varepsilon < 1/2,$$

where Σ_{lhs} *and* Σ_{iid} *are as in (2).*

Remark. (29) provides a way of calibrating r to ensure that $\Theta_{n,r}$ is an asymptotically valid confidence region for μ having the desired degree of confidence.

Proof of Theorem 4. The convexity of $\Theta_{n,r}$ follows from Jensen's inequality and the observation that $(\prod_{k=1}^n nw_k)^{1/n}$ is concave (strictly concave if $n \geq 2$) in $w \in \mathcal{S}$ (see for example Marshall and Olkin (1979) page 79).

(a) Since $\mu \in \Theta_{n,r}$ is equivalent to $R_n(\mu) \geq r$, to prove (29) it suffices to show

$$\lim_{n\to\infty} P(R_n(\mu) \geq r) = P(Z'M^{1/2}(M+N)^{-1}M^{1/2}Z \leq -2\log r).$$

Define $\Xi = \{\xi \in \mathcal{R}^p : \|\xi\| = 1\}$. From Lemma 2 of Owen (1990), we have $\inf_{\xi\in\Xi} P((f\circ X - \mu)'\xi > 0) > 0$. Hence it follows from Corollary 3 and the Glivenko-Cantelli theorem that

$$\sup_{\xi\in\Xi} |P((f\circ X-\mu)'\xi > 0) - n^{-1}\sum_{k=1}^{n} I\{[f\circ X(\pi_1(k),\ldots,\pi_d(k)) - \mu]'\xi > 0\}| \to 0$$

almost surely as $n \to \infty$ and thus

$$\inf_{\xi\in\Xi} n^{-1}\sum_{k=1}^{n} I\{[f\circ X(\pi_1(k),\ldots,\pi_d(k)) - \mu]'\xi > 0\} > c^*$$

almost surely for sufficiently large n. This implies that μ is an interior point of the convex hull of $\{f\circ X(\pi_1(k),\ldots,\pi_d(k)) : 1 \leq k \leq n\}$ and $R_n(\mu) > c^*$ almost surely for large n. Using Lagrange multipliers, the solution of (28) is found to be

$$nw_k = (1+\gamma_k)^{-1}, \tag{31}$$

where $\gamma_k = \eta'(f\circ X(\pi_1(k),\ldots,\pi_d(k)) - \mu)$ for all $1 \leq k \leq n$ and $\eta \in \mathcal{R}^p$ satisfies

$$\sum_{k=1}^{n} (f\circ X(\pi_1(k),\ldots,\pi_d(k)) - \mu)/(1+\gamma_k) = 0.$$

For simplicity we write

$$S = n^{-1}\sum_{k=1}^{n}[f\circ X(\pi_1(k),\ldots,\pi_d(k)) - \mu][f\circ X(\pi_1(k),\ldots,\pi_d(k)) - \mu]'.$$

Then it follows from Corollary 3 that S^{-1} exists almost surely for large n and we define

$$\begin{aligned} \zeta &= \eta - S^{-1}(\hat{\mu}_n - \mu) \\ &= n^{-1}S^{-1}\sum_{k=1}^{n}\{f\circ X(\pi_1(k),\ldots,\pi_d(k)) - \mu\}\gamma_k^2/(1+\gamma_k). \end{aligned} \tag{32}$$

Next we note from (31) that

$$
\begin{aligned}
-2\log R_n(\mu) &= 2\sum_{k=1}^{n}\log(1+\gamma_k) \\
&= \sum_{k=1}^{n}\{2\gamma_k-\gamma_k^2+[2\log(1+\gamma_k)-2\gamma_k+\gamma_k^2]\} \\
&= n(\hat{\mu}_n-\mu)'S^{-1}(\hat{\mu}_n-\mu)-n\zeta'S\zeta \\
&\quad +\sum_{k=1}^{n}[2\log(1+\gamma_k)-2\gamma_k+\gamma_k^2]. \qquad (33)
\end{aligned}
$$

Now we observe as in Owen (1990) pages 101-102 that $n\zeta'S\zeta \to 0$ and $\sum_{k=1}^{n}[2\log(1+\gamma_k)-2\gamma_k+\gamma_k^2]\to 0$ in probability as $n\to\infty$. Since $E\|f\circ X\|^4<\infty$, it follows from Corollaries 2 and 3 that $n^{1/2}(\hat{\mu}_n-\mu)$ converges in distribution to $M^{1/2}Z$ and S converges almost surely to $M+N$ as $n\to\infty$. Thus we conclude from (33) that

$$\lim_{n\to\infty} P(R_n(\mu)\geq r)=P(Z'M^{1/2}(M+N)^{-1}M^{1/2}Z\leq -2\log r).$$

This proves (a).

(b) Clearly the arguments in part (a) apply equally well here. Conditioning on the occurrence or non-occurrence of the event

$$\{n\zeta'S\zeta\leq cn^{-1/2}\}\cap\{|\sum_{k=1}^{n}2\log(1+\gamma_k)-2\gamma_k+\gamma_k^2|\leq cn^{\varepsilon-1/2}\},$$

it follows from (33) and Lemma 7 (see Appendix) that

$$
\begin{aligned}
&|P(R_n(\mu)\geq r)-P(Z'\Sigma_{lhs}^{1/2}\Sigma_{iid}^{-1}\Sigma_{lhs}^{1/2}Z\leq -2\log r)| \\
\leq\ & |P\{n(\hat{\mu}_n-\mu)'(n\Sigma_{iid})^{-1}(\hat{\mu}_n-\mu)\leq -2\log r\} \\
&-P(Z'\Sigma_{lhs}^{1/2}\Sigma_{iid}^{-1}\Sigma_{lhs}^{1/2}Z\leq -2\log r)| \\
&+|P\{n(\hat{\mu}_n-\mu)'S^{-1}(\hat{\mu}_n-\mu)\leq -2\log r\} \\
&-P\{n(\hat{\mu}_n-\mu)'(n\Sigma_{iid})^{-1}(\hat{\mu}_n-\mu)\leq -2\log r\}|+c^*n^{\varepsilon-1/2}. \qquad (34)
\end{aligned}
$$

Since $\{x\in\mathcal{R}^p : x'\Sigma_{lhs}^{1/2}\Sigma_{iid}^{-1}\Sigma_{lhs}^{1/2}x\leq -2\log r\}$ is a convex set in $\mathcal{R}^p$, we conclude from Corollary 2 that the right hand side of (34) is bounded by $c^*n^{\varepsilon-1/2}$. This proves (30). □

5 Acknowledgments

I would like to thank Professor Friedrich Götze for sending me preprints of his recent work.

6 Appendix

Lemma 3 *Suppose* $h : [0,1]^r \to \mathcal{R}$ *is a measurable function such that* $\|h\|_q < \infty$ *for some* $1 \le q \le \infty$ *and*

$$\rho_n(x;h) = n^r \int_{j_1/n}^{(j_1+1)/n} \cdots \int_{j_r}^{(j_r+1)/n} h(y)dy,$$

whenever $x \in \prod_{i=1}^r [j_i/n, (j_i+1)/n)$ *for some* $0 \le j_1, \ldots, j_r \le n-1$. *Then* $\|h - \rho_n(.;h)\|_q \to 0$ *as* $n \to \infty$.

PROOF. We refer the reader to Royden (1988) page 129 for a proof when $r = 1$. The proof of the lemma for $r > 1$ is similar and is omitted. □

REMARK. If $h : [0,1]^r \to \mathcal{R}^p$, then we write

$$\rho_n(x;h) = (\rho_n(x;h_1), \ldots, \rho_n(x;h_p))', \quad \forall x \in [0,1]^r.$$

PROOF OF THEOREM 1. We observe that

$$\begin{aligned} n\Sigma_{lhs} &= n^{-1} \sum_{i=1}^n E[f \circ X(i, \pi_2(i), \ldots, \pi_d(i)) - \mu] \\ &\quad \times [f \circ X(i, \pi_2(i), \ldots, \pi_d(i)) - \mu]' \\ &\quad + n^{-1} \sum_{i \ne j} E[f \circ X(i, \pi_2(i), \ldots, \pi_d(i)) - \mu] \\ &\quad \times [f \circ X(j, \pi_2(j), \ldots, \pi_d(j)) - \mu]'. \end{aligned} \tag{35}$$

Define for $0 \le s, t < 1$,

$$\delta_n(s,t) = \begin{cases} 1 & \text{if } \lfloor ns \rfloor = \lfloor nt \rfloor, \\ 0 & \text{otherwise,} \end{cases} \tag{36}$$

where $\lfloor t \rfloor$ denotes the greatest integer less than or equal to t. We further observe that

$$[n(n-1)]^{-1} \sum_{i \ne j} E[f \circ X(i, \pi_2(2), \ldots, \pi_d(i)) - \mu]$$

$$\begin{aligned}
&\times[f \circ X(j, \pi_2(j), \ldots, \pi_d(j)) - \mu]' \\
= \quad & n^d(n-1)^{-d} \int_{[0,1]^{2d}} [f(x) - \mu][f(y) - \mu]' \prod_{k=1}^{d} [1 - \delta_n(x_k, y_k)] dx dy \\
= \quad & -n^d(n-1)^{-d} \sum_{k=1}^{d} \int_0^1 \int_0^1 f_{-k}(x_k) f'_{-k}(y_k) \delta_n(x_k, y_k) dx_k dy_k + R, \quad \text{say.}
\end{aligned}$$

Thus it follows from Lemma 3 (with $q = 2$) that

$$\begin{aligned}
&\lim_{n\to\infty} n^{-1} \sum_{i \neq j} E[f \circ X(i, \pi_2(i), \ldots, \pi_d(i)) - \mu] \\
&\times[f \circ X(j, \pi_2(j), \ldots, \pi_d(j)) - \mu]' \\
= \quad & \lim_{n\to\infty} -n \sum_{i=1}^{n} \sum_{k=1}^{d} \int_{(i-1)/n}^{i/n} f_{-k}(x_k) dx_k \int_{(i-1)/n}^{i/n} f'_{-k}(y_k) dy_k \\
= \quad & \lim_{n\to\infty} -n^{-1} \sum_{i=1}^{n} \sum_{k=1}^{d} \rho_n((i-1)/n; f_{-k}) \rho'_n((i-1)/n; f_{-k})
\end{aligned}$$

$$= -\sum_{k=1}^{d} \int_0^1 f_{-k}(x_k) f'_{-k}(x_k) dx_k, \tag{37}$$

and it can similarly be shown that $|R| \leq cn^{-2} E\|f \circ X\|^2$. Also we observe from (3) that

$$E(f \circ X - \mu)(f \circ X - \mu)'$$

$$= \int_{[0,1]^d} f_{rem}(x) f'_{rem}(x) dx + \sum_{k=1}^{d} \int_0^1 f_{-k}(x_k) f'_{-k}(x_k) dx_k. \tag{38}$$

The theorem now follows from (35), (37) and (38). □

Lemma 4 *With the notations of Theorem 2, to prove (6) it suffices without loss of generality to assume that* $\|Y(i_1, \ldots, i_d)\| \leq 1$ *for all* $1 \leq i_1, \ldots, i_d \leq n$.

PROOF. The following proof is heavily motivated by the truncation-type argument of Bolthausen (1984) page 382. Define

$$\tilde{Y}(i_1, \ldots, i_d) = \begin{cases} Y(i_1, \ldots, i_d) & \text{if } \|Y(i_1, \ldots, i_d)\| \leq 1/(4d), \\ 0 & \text{otherwise,} \end{cases} \tag{39}$$

$\tilde{\mu} = E\tilde{W}$ and $\tilde{W} = \sum_{i=1}^{n} \tilde{Y}(\pi_1(i), \ldots, \pi_d(i))$. We observe by Markov's inequality that

$$\begin{aligned} P(W \neq \tilde{W}) &\leq P(\sum_{i=1}^{n} I\{\|Y(i, \pi_2(i), \ldots, \pi_d(i))\| > 1/(4d)\} \geq 1) \\ &\leq n^{1-d} \sum_{1 \leq i_1, \ldots, i_d \leq n} P(\|Y(i_1, \ldots, i_d)\| > 1/(4d)) \\ &\leq (4d)^3 \beta_3, \end{aligned} \tag{40}$$

and

$$\begin{aligned} \|\tilde{\mu}\| &\leq n^{1-d} \sum_{1 \leq i_1, \ldots, i_d \leq n} E\|Y(i_1, \ldots, i_d)\| I\{\|Y(i_1, \ldots, i_d)\| > 1/(4d)\} \\ &\leq (4d)^2 \beta_3. \end{aligned} \tag{41}$$

Writing $\tilde{\Sigma} = \mathrm{Cov}(\tilde{W})$, we further observe that for $1 \leq i, j \leq p$, $\mathrm{Cov}(W)_{i,j} = \delta_{i,j}$ and

$$\begin{aligned} &\delta_{i,j} - \tilde{\Sigma}_{i,j} - \tilde{\mu}_i \tilde{\mu}_j \\ = \; & E\{\sum_{a=1}^{n} Y_i(a, \pi_2(a), \ldots, \pi_d(a)) Y_j(a, \pi_2(a), \ldots, \pi_d(a)) \\ & \times I\{\|Y(a, \pi_2(a), \ldots, \pi_d(a))\| > 1/(4d)\} \\ & - \sum_{a \neq b} Y_i(a, \pi_2(a), \ldots, \pi_d(a)) I\{\|Y(a, \pi_2(a), \ldots, \pi_d(a))\| > 1/(4d)\} \\ & \times Y_i(b, \pi_2(b), \ldots, \pi_d(b)) I\{\|Y(b, \pi_2(b), \ldots, \pi_d(b))\| > 1/(4d)\} \\ & + 2 \sum_{a \neq b} Y_i(a, \pi_2(a), \ldots, \pi_d(a)) I\{\|Y(a, \pi_2(a), \ldots, \pi_d(a))\| > 1/(4d)\} \\ & \times Y_i(b, \pi_2(b), \ldots, \pi_d(b))\} \\ = \; & \Delta_1 - \Delta_2 + 2\Delta_3, \quad \text{say.} \end{aligned} \tag{42}$$

We note that

$$\begin{aligned} |\Delta_1| &\leq n^{1-d} \sum_{1 \leq i_1, \ldots, i_d \leq n} E\|Y(i_1, \ldots, i_d)\|^2 I\{\|Y(i_1, \ldots, i_d)\| > 1/(4d)\} \\ &\leq 4d\beta_3, \end{aligned} \tag{43}$$

and in a similar way,

$$|\Delta_2| \leq (\frac{n}{n-1})^{d-1} [(4d)^2 \beta_3]^2 \leq (\frac{n}{n-1})^{d-1} (4d)^4 \varepsilon_0 \beta_3. \tag{44}$$

Also we have from (5),

$$\begin{aligned}|\Delta_3| &\leq |[n(n-1)]^{1-d}\sum_{k=1}^{d}\sum_{i_k\neq j_k} EY_i(i_1,\ldots,i_d)\\ &\quad\times I\{\|Y(i_1,\ldots,i_d)\| > (1/(4d)\}\mu_j(j_1,\ldots,j_d)|\\ &= |[n(n-1)]^{1-d}\sum_{1\leq i_1,\ldots,i_d\leq n} EY_i(i_1,\ldots,i_d)\\ &\quad\times I\{\|Y(i_1,\ldots,i_d)\| > (1/(4d)\}\sum_{\nu=0}^{d-2}\sum{}^{(\nu)}(-1)^{d-\nu}\mu_j(j_1,\ldots,j_d)|\end{aligned}$$

where given $i_1,\ldots,i_d$, $\sum^{(\nu)}$ denotes the sum over $j_1,\ldots,j_d$ with exactly $d-\nu$ of the j's satisfying $j_{k_1}=i_{k_1},\ldots,j_{k_{d-\nu}}=i_{k_{d-\nu}}$ for some $1\leq k_1<\cdots<k_{d-\nu}\leq d$. Consequently it follows from Hölder's and Markov's inequalities that

(45) $$|\Delta_3|\leq cn^{-1}\beta_3.$$

Since $\beta_3\leq\varepsilon_0$, it follows from (41), (42), (43), (44) and (45) that $\tilde{\Sigma}$ tends to the identity matrix as $\varepsilon_0\to 0$. Thus by choosing $\varepsilon_0>0$ sufficiently small, $\tilde{\Sigma}^{-1}$ exists. Next define as in (5),

$$\begin{aligned}E\tilde{Y}(i_1,\ldots,i_d) &= \tilde{\mu}(i_1,\ldots,i_d),\quad \forall 1\leq i_1,\ldots,i_d\leq n,\\ \tilde{\mu}_{-k}(i_k) &= (1/n^{d-1})\sum_{j\neq k}\sum_{i_j=1}^{n}\tilde{\mu}(i_1,\ldots,i_d),\end{aligned}$$

and

$$\tilde{Y}^*(i_1,\ldots,i_d)=\tilde{\Sigma}^{-1/2}[\tilde{Y}(i_1,\ldots,i_d)-\sum_{k=1}^{d}\tilde{\mu}_{-k}(i_k)+(d-1)\tilde{\mu}].$$

Now it follows from (39) that for sufficiently small $\varepsilon_0>0$,

(46) $$\|\tilde{Y}^*(i_1,\ldots,i_d)\|\leq 1,\quad \forall 1\leq i_1,\ldots,i_d\leq n$$

and

(47) $$(1/n^{d-1})\sum_{1\leq i_1,\ldots,i_d\leq n} E\|\tilde{Y}^*(i_1,\ldots,i_d)\|^3\leq c\beta_3.$$

Let Z denote the random vector having probability distribution Φ_p. Then

$$\sup_{g\in\mathcal{A}}|E[g(W)-g(Z)]|$$

$$\begin{aligned} &= \sup_{g\in\mathcal{A}} |E[g(W)-g(Z)|W=\tilde{W}]P(W=\tilde{W}) \\ &\quad +E[g(W)-g(Z)|W\neq\tilde{W}]P(W\neq\tilde{W})| \\ &\leq \sup_{g\in\mathcal{A}} |E[g(\tilde{W})-g(Z)]| + 4P(W\neq\tilde{W}) \\ &\leq \sup_{g\in\mathcal{A}} |E[g(\tilde{W})-g(\tilde{Z})]| + |E[g(Z)-g(\tilde{Z})]| + 4P(W\neq\tilde{W}), \end{aligned}$$

where $\tilde{Z}$ denotes the p-variate normal random vector having the same mean and covariance matrix as $\tilde{W}$. Using (40), the Taylor expansion for the density of $\tilde{Z}$ and the fact that $\mathcal{A}$ is closed under affine transformations, we have

$$\sup_{g\in\mathcal{A}} |E[g(W)-g(Z)]| \leq \sup_{g\in\mathcal{A}} |E[g(\tilde{\Sigma}^{-1/2}(\tilde{W}-E\tilde{W}))-g(Z)]| + c\beta_3. \tag{48}$$

Thus it follows from (46), (47) and (48) that to prove Theorem 2 it suffices to prove (6) under the assumption that $\|Y(i_1,\ldots,i_d)\| \leq 1$ for all $1 \leq i_1,\ldots,i_d \leq n$. □

Lemma 5 *With the notations and assumptions of Theorem 2, we have* $R_1 = 0$ *and by choosing* ε_0 *sufficiently small, we have*

$$\sup\{|R_2| : h \in \mathcal{A}\} \leq c\beta_3(1+\varepsilon^{-1}C_{d,p}\beta_3).$$

PROOF. From the combinatorial construction of $\pi^{(1)}$, $\pi^{(2)}$ and $\pi^{(3)}$, we observe that

$$\begin{aligned} R_1 &= \sum_{i,j=1}^{p} \{EnY_i(I_1,J_{2,2,1},\ldots,J_{d,d,1})W_j^{(3)} \\ &\quad -EnY_i(I_1,J_{2,2,1},\ldots,J_{d,d,1})EW_j^{(2)} - \delta_{i,j}\}E\frac{\partial^2}{\partial w_i \partial w_j}\psi_{\varepsilon^2}(W^{(1)}) \\ &= \sum_{i,j=1}^{p} \{EnY_i(I_1,J_{2,2,1},\ldots,J_{d,d,1})W_j^{(3)} - \delta_{i,j}\}E\frac{\partial^2}{\partial w_i \partial w_j}\psi_{\varepsilon^2}(W^{(1)}). \qquad (49) \end{aligned}$$

We further observe that for all $1 \leq i,j \leq p$,

$$EnY_i(I_1,J_{2,2,1},\ldots,J_{d,d,1})W_j^{(3)} = EW_iW_j = \delta_{i,j}. \tag{50}$$

It now follows from (49) and (50) that $R_1 = 0$.

For simplicity of notation, we write

$$\begin{aligned} \Omega_1 &= \{I_k, L_{j,k} : 2 \le j \le d, 1 \le k \le d\}, \\ \Omega_2 &= \{J_{i,j,k} : 2 \le i \le d, 1 \le j, k \le d\}, \end{aligned}$$

and $\Omega_2 = \cdots = \Omega_d$. Let $\mathcal{C}$ denote the sigma-field generated by $\Omega_1 \cup \Omega_2$. Then

$$\begin{aligned} |R_2| \le& \sum_{i,j,k=1}^{p} \int_0^1 \int_0^1 E|nY_i(I_1, J_{2,2,1}, \ldots, J_{d,d,1})(W^{(3)} - W^{(2)})_j \\ &\times (W^{(2)} - W^{(1)} + t(W^{(3)} - W^{(2)}))_k E[\frac{\partial^3}{\partial w_i \partial w_j \partial w_k}(\psi_{\varepsilon^2}(W^{(1)} \\ &+ s(W^{(2)} - W^{(1)}) + st(W^{(3)} - W^{(2)})) - \psi_{\varepsilon^2}(\tilde{Z}_{s,t}))|\mathcal{C}]|dsdt \\ &+ \sum_{i,j,k=1}^{p} \int_0^1 \int_0^1 E|nY_i(I_1, J_{2,2,1}, \ldots, J_{d,d,1})(W^{(3)} - W^{(2)})_j \\ &\times (W^{(2)} - W^{(1)} + t(W^{(3)} - W^{(2)}))_k \\ &\times E[\frac{\partial^3}{\partial w_i \partial w_j \partial w_k} \psi_{\varepsilon^2}(\tilde{Z}_{s,t})|\mathcal{C}]|dsdt \\ =& R_3 + R_4, \quad \text{say}, \end{aligned}$$

where given $\mathcal{C}$, $\tilde{Z}_{s,t}$ has the p-variate normal distribution with the same mean and covariance matrix $\Sigma_{s,t,\mathcal{C}}$ as

$$\begin{aligned} & W^{(1)} + s(W^{(2)} - W^{(1)}) + st(W^{(3)} - W^{(2)}) \\ =& \sum_{i \notin \Omega_1} [Y(i, \pi_2^{(1)}(i), \ldots, \pi_d^{(1)}(i)) + (n - |\Omega_1|)^{-1} R_{s,t,\mathcal{C}}] \\ =& \sum_{i \notin \Omega_1} V_{s,t,\mathcal{C}}(i, \pi_2^{(1)}(i), \ldots, \pi_d^{(1)}(i)), \quad \text{say}, \end{aligned}$$

where $|\Omega_1|$ denotes the number of distinct elements in Ω_1. We note from (11) that the existence of $\Sigma_{s,t,\mathcal{C}}^{-1}$, uniformly over $0 \le s, t \le 1$ and $\mathcal{C}$, is ensured by choosing ε_0 sufficiently small. Define for all $i_k \notin \Omega_k$, $1 \le k \le d$,

$$\begin{aligned} E[V_{s,t,\mathcal{C}}(i_1, \ldots, i_d)|\mathcal{C}] &= \mu_{s,t,\mathcal{C}}(i_1, \ldots, i_d), \\ \mu_{s,t,\mathcal{C},-k}(i_k) &= (n - |\Omega_1|)^{1-d} \sum_{j \ne k} \sum_{i_j \notin \Omega_j} \mu_{s,t,\mathcal{C}}(i_1, \ldots, i_d), \end{aligned}$$

and

$$V^*_{s,t,\mathcal{C}}(i_1,\ldots,i_d) = \Sigma^{-1/2}_{s,t,\mathcal{C}}[V_{s,t,\mathcal{C}}(i_1,\ldots,i_d) - \sum_{k=1}^{d} \mu_{s,t,\mathcal{C},-k}(i_k) + (d-1)\mu_{s,t,\mathcal{C}}].$$

Next we observe that

$$(n - |\Omega_1|)^{1-d} \sum_{k=1}^{d} \sum_{i_k \notin \Omega_k} E(\|V^*_{s,t,\mathcal{C}}(i_1,\ldots,i_d)\|^3 |\mathcal{C}) \le c\beta_3, \tag{51}$$

and

$$\begin{aligned} &E|nY_i(I_1, J_{2,2,1},\ldots,J_{d,d,1})(W_j^{(3)} - W_j^{(2)})| \\ &\times(|W_k^{(2)} - W_k^{(1)}| + |W_k^{(3)} - W_k^{(2)}|) = c\beta_3. \end{aligned} \tag{52}$$

Now it follows from (10), (51), the induction hypothesis and the fact that $\mathcal{A}$ is closed under affine transformations that

$$\sup\{R_3 : h \in \mathcal{A}\} \le c\varepsilon^{-1} C_{d,p}\beta_3^2.$$

Finally from (7), (8) and (52), we get $R_4 \le c\beta_3$. This proves the lemma. □

Lemma 6 *With the notations and assumptions of Theorem 3, we have*

$$\begin{aligned} &E\{\sup_{1\le j\le n} \|\sum_{k=1}^{j} \tilde{f}^+ \circ X(\pi_1(k),\ldots,\pi_d(k))\|^{2mq^*}\} \\ \le\ & c^{**} n\{\sum_{k=0}^{\lfloor \log_2 n \rfloor} \lambda(\lceil n/2^{k+1} \rceil)\}^{2mq^*}, \end{aligned} \tag{53}$$

where c^{**} is a generic constant that depends only on d, p and $2mq^*$ and that for $1 \le j \le n$,

$$\lambda(j) = \sum_{i=0}^{2mq^*-2} j^{i/[2mq^*(2mq^*-i)]} (E\|\tilde{f}^+ \circ X\|^{2mq^*-i})^{1/(2mq^*-i)}.$$

PROOF. We shall first show that

$$E\|\sum_{k=1}^{j} \tilde{f}^+ \circ X(\pi_1(k),\ldots,\pi_d(k))\|^{2mq^*} \le c^{**} j \lambda^{2mq^*}(j), \quad \forall 1 \le j \le n. \tag{54}$$

We observe that for $1 \leq j \leq n$,

$$
\begin{aligned}
& E\|\sum_{k=1}^{j} \tilde{f}^{+} \circ X(\pi_1(k),\dots,\pi_d(k))\|^{2mq^*} \\
= \; & E\{\sum_{i=1}^{p} \sum_{1\leq k_1,k_2\leq j} (\tilde{f}^{+} \circ X)_i(\pi_1(k_1),\dots,\pi_d(k_1)) \\
& \times(\tilde{f}^{+} \circ X)_i(\pi_1(k_2),\dots,\pi_d(k_2))\}^{mq^*}. \qquad (55)
\end{aligned}
$$

On simplification, it can be seen that the right hand side of (55) can be expressed as a finite sum of terms each of the form

$$
c^{**} \sum_{1\leq k_1<\cdots<k_l\leq j} E\prod_{t=1}^{l}\prod_{a_t=1}^{q_t}(\tilde{f}^{+} \circ X)_{i_{t,a_t}}(\pi_1(k_t),\dots,\pi_d(k_t)), \qquad (56)
$$

for some l such that $q_t \geq 1, 1 \leq i_{t,a_t} \leq p$ for all $1 \leq t \leq l$ and $2mq^* = \sum_{t=1}^{l} q_t$. Now if $q_t \geq 2$ for all $1 \leq t \leq l$, we observe that the absolute value of (56) is bounded by

$$
\begin{aligned}
& c^{**}j^l|\int_{[0,1]^{ld}}\prod_{t=1}^{l}\prod_{a_t=1}^{q_t}\tilde{f}^{+}_{i_{t,a_t}}(x^{(t)}) \\
& \times \prod_{1\leq r<s\leq l}\prod_{k=1}^{d}(1-\delta_n(x_k^{(r)},x_k^{(s)}))dx^{(1)}\dots dx^{(l)}| \\
& \leq c^{**}j^l\prod_{t=1}^{l}E\|\tilde{f}^{+}\circ X\|^{q_t} \leq c^{**}j\lambda^{2mq^*}(j),
\end{aligned}
$$

where $\delta_n(x_k^{(r)},x_k^{(s)})$ is as in (36). Next we suppose that there exists a $q_t = 1$ for some $1 \leq t \leq l$. Without loss of generality, assume that $q_1 = \cdots = q_b = 1$ and $q_j \geq 2$ whenever $t > b \geq 1$. Then the absolute value of (56) is bounded by

$$
\begin{aligned}
& c^{**}j^l|\int_{[0,1]^{ld}}\{\prod_{t=1}^{b}\tilde{f}^{+}_{i_{t,1}}(x^{(t)})\}\{\prod_{t=b+1}^{l}\prod_{a_t=1}^{q_t}\tilde{f}^{+}_{i_{t,a_t}}(x^{(t)})\} \\
& \times\{\prod_{1\leq r<s\leq l}\prod_{k=1}^{d}(1-\delta_n(x_k^{(r)},x_k^{(s)}))\}dx^{(1)}\dots dx^{(l)}|. \qquad (57)
\end{aligned}
$$

Since $E(\tilde{f}^{+} \circ X) = 0$, by expanding the third product in (57), we observe that (57) can be rewritten as a finite sum of terms each of which is bounded by

$$c^{**} j^{l^*} \prod_{t=1}^{l^*} E\|\tilde{f}^{+} \circ X\|^{q_t^*} \leq c^{**} j \lambda^{2mq^*}(j),$$

for some l^*, q_t^* where $q_t^* \geq 2$ for all $1 \leq t \leq l^*$ and $\sum_{t=1}^{l^*} q_t^* = 2mq^*$. This proves (54). Finally (53) follows from Theorem 3.3 of Móricz, Serfling and Stout (1982) and (54). □

Lemma 7 *With the notations and assumptions of Theorem 4(b),*

$$\begin{aligned} P(n\zeta' S \zeta > cn^{-1/2}) &\leq c^* n^{-1/2}, \\ P\{|\sum_{k=1}^{n} [2\log(1+\gamma_k) - 2\gamma_k + \gamma_k^2]| > cn^{\varepsilon - 1/2}\} &\leq c^* n^{-1/2}, \end{aligned}$$

and

$$\begin{aligned} &|P\{n(\hat{\mu}_n - \mu)' S^{-1} (\hat{\mu}_n - \mu) \leq -2\log r\} \\ &-P\{n(\hat{\mu}_n - \mu)'(n\Sigma_{iid})^{-1}(\hat{\mu}_n - \mu) \leq -2\log r\}| \leq c^* n^{\varepsilon - 1/2}. \end{aligned}$$

PROOF. We first observe from the definition of latin hypercube sampling and also as in Owen (1990) page 103 that

$$\begin{aligned} E\|\eta\|^{2j} &\leq c^* n^{-j}, \quad \forall 1 \leq j \leq 3, \\ P(\max_{1 \leq k \leq n} |\gamma_k| > 1/4) &\leq c^* n^{-1/2}. \end{aligned} \tag{58}$$

By conditioning on the occurrence or non-occurrence of $\{\max_{1 \leq k \leq n} |\gamma_k| > 1/4\}$, it follows from the definition of ζ in (32) that

$$\begin{aligned} &P(n\zeta' S \zeta > cn^{-1/2}) \\ \leq\; &P(\{n^{-1} \sum_{k=1}^{n} \|f \circ X(\pi_1(k), \ldots, \pi_d(k)) - \mu\| \gamma_k^2\}^2 > c^* n^{-3/2}) \\ &+2P(\max_{1 \leq k \leq n} |\gamma_k| > 1/4) \\ \leq\; &P(n^{-1} \|\eta\|^2 \sum_{k=1}^{n} \|f \circ X(\pi_1(k), \ldots, \pi_d(k)) - \mu\|^3 > c^* n^{-3/4}) \\ &+c^* n^{-1/2}. \end{aligned}$$

Since $E\|f\circ X\|^6 < \infty$, it follows from Corollary 3 that $n^{-1}\sum_{k=1}^{n}\|f\circ X(\pi_1(k),\ldots,\pi_d(k))-\mu\|^3$ converges almost surely to $E\|f\circ X-\mu\|^3 > 0$. Thus it follows from (58) and Markov's inequality that $P(n\zeta' S\zeta > cn^{-1/2}) \le c^* n^{-1/2}$.

We observe from Theorem 3 and a lemma from Baum and Katz (1965) page 113 that for $0<\varepsilon^*<\varepsilon$,

$$P(\max_{1\le i,j\le p}|(S-n\Sigma_{iid})_{i,j}| > cn^{\varepsilon^*-1/2}) \le c^* n^{-1/2}, \tag{59}$$

$$P(n\|\hat{\mu}_n-\mu\|^2 > cn^{\varepsilon-\varepsilon^*}) \le c^* n^{-1/2}. \tag{60}$$

Also using Markov's inequality, we have

$$P(\max_{1\le k\le n}\|f\circ X(\pi_1(k),\ldots,\pi_d(k))-\mu\| > n^{3/8}) \le c^* n^{-1/2},$$

and hence from Theorem 3 we obtain

$$P(\{\max_{1\le k\le n}\|f\circ X(\pi_1(k),\ldots,\pi_d(k))-\mu\|\}\|\hat{\mu}_n-\mu\| > c) \le c^* n^{-1/2}. \tag{61}$$

Now by conditioning on the occurrence or non-occurrence of the event $\{\max_{1\le k\le n}|\gamma_k| > 1/4\}$, we have

$$\begin{aligned}
& P\{\sum_{k=1}^{n}[2\log(1+\gamma_k)-2\gamma_k+\gamma_k^2] > cn^{\varepsilon-1/2}\} \\
\le\ & P(\sum_{k=1}^{n}\|\gamma_k\|^3 > c^* n^{\varepsilon-1/2}) + c^* n^{-1/2} \\
\le\ & P(\|\eta\|^3 > c^* n^{\varepsilon-3/2}) + c^* n^{-1/2} \\
\le\ & P(\|\hat{\mu}_n-\mu\|^3 > c^* n^{\varepsilon-3/2}) + c^* n^{-1/2} \le c^* n^{-1/2}.
\end{aligned}$$

Here the second last inequality uses (2.12) of Owen (1990) page 101, (59) and (61).

Finally by expanding S^{-1} as a Taylor series up to linear terms, it follows from (59) and (60) that

$$\begin{aligned}
&|P\{n(\hat{\mu}_n-\mu)'S^{-1}(\hat{\mu}_n-\mu) \le -2\log r\} \\
&-P\{n(\hat{\mu}_n-\mu)'(n\Sigma_{iid})^{-1}(\hat{\mu}_n-\mu) \le -2\log r\}| \le c^* n^{\varepsilon-1/2}.
\end{aligned}$$

This proves the lemma. □

References

[1] BAUM, L. E. and KATZ, M. (1965). Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120** 108-123.

[2] BOLTHAUSEN, E. (1984). An estimate of the remainder in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **66** 379-386.

[3] BOLTHAUSEN, E. and GÖTZE, F. (1989). The rate of convergence for multivariate sampling statistics. Preprint.

[4] ERDÖS, P. (1949). On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20** 286-291.

[5] GÖTZE, F. (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** 724-739.

[6] HIPP, C. (1979). Convergence rates of the strong law for stationary mixing sequences. *Z. Wahrsch. Verw. Gebiete* **49** 49-62.

[7] HO, S. T. and CHEN, L. H. Y. (1978). An L_p bound for the remainder in a combinatorial central limit theorem. *Ann. Probab.* **6** 231-249.

[8] HOEFFDING, W. (1951). A combinatorial central limit theorem. *Ann. Math. Statist.* **22** 558-566.

[9] HSU, P. L. and ROBBINS, H. (1947). Complete convergence and the large of large numbers. *Proc. Nat. Acad. Sci. U.S.A.* **33** 25-31.

[10] KATZ, M. The probability in the tail of a distribution. *Ann. Math. Statist.* **34** 312-318.

[11] LAI, T. L. (1977). Convergence rates and r-quick versions of the strong law for stationary mixing sequences. *Ann. Probab.* **5** 693-706.

[12] MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

[13] MCKAY, M. D., BECKMAN, R. J. and CONOVER, W. J. (1979). A comparison of three methods for selecting values of output variables in the analysis of output from a computer code. *Technometrics* **21** 239-245.

[14] MÓRICZ, F. A., SERFLING, R. J. and STOUT, W. F. (1982). Moment and probability bounds with quasi-superadditive structure for the maximum partial sum. *Ann. Probab.* **10** 1032-1040.

[15] MOTOO, M. (1957). On the Hoeffding's combinatorial central limit theorem. *Ann. Inst. Statist. Math.* **8** 145-154.

[16] OWEN, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75** 237-249.

[17] OWEN, A. B. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18** 90-120.

[18] OWEN, A. B. (1992). A central limit theorem for latin hypercube sampling. *J. R. Statist. Soc.* B **54** 541-551.

[19] PELIGRAD, M. (1985). Convergence rates of the strong law for stationary mixing sequences. *Z. Wahrsch. Verw. Gebiete* **70** 307-314.

[20] ROYDEN, H. L. (1988). *Real Analysis,* 3rd edition. Macmillan, New York.

[21] STEIN, C. M. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 583-602. Univ. California Press, Berkeley.

[22] STEIN, C. M. (1986). *Approximate Computation of Expectations. IMS Lecture Notes–Monograph Ser.* **7**. IMS, Hayward, California.

[23] STEIN, M. L. (1987). Large sample properties of simulations using latin hypercube sampling. *Technometrics* **29** 143-151.

[24] VON BAHR, B. (1976). Remainder term estimate in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **35** 131-139.