The Annals of Statistics
1996, Vol. 24, No. 5, 2058–2080

ON LATIN HYPERCUBE SAMPLING[1]

BY WEI-LIEM LOH

Purdue University and National University of Singapore

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 establishes sufficient conditions 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. McKay, Beckman and Conover (1979) 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:

1. π_k, $1 \le k \le d$, be random permutations of $\{1, \dots, n\}$ each uniformly distributed over all the $n!$ possible permutations;
2. $U_{i_1, \dots, i_d, j}$, $1 \le i_1, \dots, i_d \le n$, $1 \le j \le d$, be $[0, 1]$ uniform random variables;
3. the $U_{i_1, \dots, i_d, j}$'s and π_k's all be stochastically independent.

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

$$X_j(i_1, \dots, i_d) = (i_j - U_{i_1, \dots, i_d, j})/n \qquad \forall 1 \le j \le d,$$
$$X(i_1, \dots, i_d) = (X_1(i_1, \dots, i_d), \dots, X_d(i_1, \dots, i_d))'.$$

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].

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

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

Received October 1993; revised March 1995.
[1]Research supported in part by NSA Grant MDA 904-93-H-3011.
AMS 1991 *subject classifications*. Primary 62D05; secondary 62E20, 62G15.
Key words and phrases. Berry–Esseen bound, confidence regions, Latin hypercube sampling, multivariate central limit theorem, Stein's method, strong law of large numbers.

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 being 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) and Ho and Chen (1978), and a Berry–Esseen-type bound was obtained by Bolthausen (1984) for univariate linear statistics and Bolthausen and Götze (1993) for multivariate statistics.

In Section 3, we establish sufficient conditions on the rate of convergence in the strong law of large numbers for $\hat{\mu}_n$ (Proposition 1). The main result (Theorem 3) shows that $\hat{\mu}_n$ converges almost surely to μ under finiteness of second moments.

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, $\|\cdot\|$ will be the usual Euclidean metric on $\mathscr{R}^p$, Φ_p is the standard p-variate normal distribution and, given any measurable function $h: \mathscr{R}^p \to \mathscr{R}$, we write

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

Also if $x \in \mathscr{R}^p$, 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* $\Sigma_{\mathrm{lhs}} = \mathrm{Cov}(\hat{\mu}_n)$ *and* Σ_{iid} *be the covariance matrix of* $\hat{\mu}_n$ *when the* X*'s are independently and identically distributed*, *that is*

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

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

$$\Sigma_{\mathrm{lhs}} = n^{-1}\int_{[0,1]^d} f_{\mathrm{rem}}(x) f'_{\mathrm{rem}}(x)\,dx + o(n^{-1}),$$

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

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

$$f_{-k}(x_k) = \int_{[0,1]^{d-1}} [f(x) - \mu] \prod_{j \neq k} dx_j,$$
$$f_{\mathrm{rem}}(x) = f(x) - \mu - \sum_{k=1}^{d} f_{-k}(x_k). \tag{3}$$

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

COROLLARY 1. $\Sigma_{\mathrm{iid}} - \Sigma_{\mathrm{lhs}}$ *is asymptotically positive semidefinite*, *that is*,

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

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

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

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 (1993). Let

$$Ef \circ X(i_1, \dots, i_d) = \mu(i_1, \dots, i_d) \qquad \forall 1 \leq i_1, \dots, i_d \leq n,$$
$$\mu_{-k}(i_k) = (1/n^{d-1}) \sum_{j \neq k} \sum_{i_j = 1}^{n} \mu(i_1, \dots, i_d)$$

and

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

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

$$g_\delta^+(w) = \sup\{g(w+y): \|y\| \le \delta\} \qquad \forall w \in \mathscr{R}^p,$$
$$g_\delta^-(w) = -\inf\{g(w+y): \|y\| \le \delta\} \qquad \forall w \in \mathscr{R}^p,$$
$$\omega(g,\delta) = \int_{\mathscr{R}^p}[g_\delta^+(y) - g_\delta^-(y)]\Phi_p(dy).$$

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

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

REMARK. If $\gamma \ge 2\sqrt{p}$, then $\mathscr{A}$ can be taken to be the class of all indicator functions of measurable convex sets in $\mathscr{R}^p$ [see, for example, Bolthausen and Götze (1993)].

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

$$(6)\qquad \sup\left\{\left|Eg(W) - \int_{\mathscr{R}^p} g(x)\Phi_p(dx)\right| : g \in \mathscr{A}\right\} \le C_{d,p}\beta_3,$$

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

The following is an immediate corollary.

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

$$\sup\left\{\left|Eg(W) - \int_{\mathscr{R}^p} g(x)\Phi_p(dx)\right| : g \in \mathscr{A}\right\} \le c^* n^{-1/2}.$$

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

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

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

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* $w = (w_1, \dots, w_p)' \in \mathscr{R}^p$, *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*

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

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

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

$$\begin{aligned} &\sup_{g \in \mathscr{A}} \left| \int_{\mathscr{R}^p} g(w) [Q(dw) - \Phi_p(dw)] \right| \\ &\qquad \le \frac{4}{3} \sup_{h \in \mathscr{A}} \left| \int_{\mathscr{R}^p} \chi_{\varepsilon^2}(w|h) Q(dw) \right| + \frac{5 \varepsilon \gamma a}{2(1 - \varepsilon^2)}, \end{aligned}$$

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 $\{1, \dots, n\}^{d^2}$ such that:

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

(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 $\{1, \dots, n\}/\{J_{j,j,K}: 1 \le k \le k_0\}$.

Let $\pi_2^{(1)}, \dots, \pi_d^{(1)}$ be independent random permutations (each uniformly distributed on the permutations of $\{1, \dots, n\}$), which are also independent of $\{I_1, J_{j,j,k}: 1 \le i,\ k \le d,\ 2 \le j \le d\}$. Define for $2 \le i,\ j \le d$ and $1 \le k \le d$,

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

Let $1 \le i_1, \dots, i_d, j_1, \dots, j_d \le n$ and $\beta(i_1, \dots, i_d, j_1, \dots, j_d)$ be a permutation of $\{1, \dots, n\}$ leaving the numbers outside $\{i_1, \dots, i_d, j_1, \dots, i_d\}$ 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 $2 \le j \le d$,

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

Finally define

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

where:

(i) $Y^{(1)}(i_1, \dots, i_d) = Y(i_1, \dots, i_d)$ whenever $1 \le i_1, \dots, i_d \le n$.

(ii) Given $\{(I_k, \pi_2^{(1)}(I_k), \dots, \pi_d^{(1)}(I_k)): 1 \le k \le d\}$, $Y^{(2)}(I_k, \pi_2^{(1)}(I_k), \dots, \pi_d^{(1)}(I_k))$ is an independent replicate of $Y^{(1)}(I_k, \pi_2^{(1)}(I_k), \dots, \pi_d^{(1)}(I_k))$ (which is also independent of all other previously defined random quantities) for all $1 \le k \le d$ and

$$Y^{(2)}(i_1, \dots, i_d) = Y^{(1)}(i_1, \dots, i_d),$$

if $(i_1, \dots, k_d) \notin \{(I_k, \pi_2^{(1)}(I_k), \dots, \pi_d^{(1)}(I_k)): 1 \le k \le d\}$.

(iii) Given $(I_1, \pi_2^{(2)}(I_1), \dots, \pi_d^{(2)}(I_1))$, $Y^{(3)}(I_1, \pi_2^{(2)}(I_1), \dots, \pi_d^{(2)}(I_1))$ is an independent replicate of $Y^{(2)}(I_1, \pi_2^{(2)}(I_1), \dots, \pi_d^{(2)}(I_1))$ (which is also independent of all other previously defined random quantities) and

$$Y^{(3)}(i_1, \dots, i_d) = Y^{(2)}(i_1, \dots, i_d),$$

if $(i_1, \dots, i_d) \ne (I_1, \pi_2^{(2)}(I_1), \dots, \pi_d^{(2)}(I_1))$.

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,

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

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. Writing $W = (W_1, \dots, W_p)'$, we have for $\varepsilon > 0$,

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

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

$$
E\sum_{i=1}^{p} nY_i^{(3)}(I_1, J_{2,2,1}, \dots, J_{d,d,1}) \frac{\partial}{\partial w_i} \psi_{\varepsilon^2}(W^{(2)}) = 0
$$

and, hence,

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

where

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

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

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

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

$$(13)\qquad \sup\{|R_1 + R_2|: h \in \mathscr{A}\} \le c\beta_3\big(1 + \varepsilon^{-1} C_{d,p}\,\beta_3\big).$$

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

$$(14)\qquad \sup_{g\in\mathscr{A}}\left| Eg(W) - \int_{\mathscr{R}^p} g(x)\Phi_p(dx)\right| \le c\beta_3\big(1 + \varepsilon^{-1}C_{d,p}\,\beta_3\big) + c\varepsilon.$$

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

3. Convergence rates in the strong law. Let $\{f \circ X(\pi_1(k), \dots, \pi_d(k)): 1 \le k \le n\}$ and $\hat{\mu}_n$ be as in (1). We shall study the rate of convergence in the strong law of large numbers for $\hat{\mu}_n$.

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

$$\sum_{n=1}^{\infty} n^{\alpha q - 2} P\big(\|n\hat{\mu}_n\| \ge \varepsilon n^{\alpha}\big) < \infty \qquad \forall \varepsilon > 0,$$

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

The following theorem is the main result of this section. It follows directly from Proposition 1 and the Borel–Cantelli lemma.

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

PROOF OF PROPOSITION 1. Without loss of generality, we assume throughout this proof that $E(f\circ X) = 0$ if the expectation exists. Suppose that $E\|f\circ X\|^q < \infty$. For $\varepsilon > 0$, we define as in Erdös (1949) and Katz (1963),

$$a_i = P\big(\|f\circ X\| > \varepsilon 2^{i\alpha}\big) \qquad \forall i \ge 0,$$

$$f^{+}(x) = \begin{cases} f(x), & \text{if } \|f(x)\| \le \varepsilon n^{\theta\alpha},\\ 0, & \text{otherwise,}\end{cases}$$

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 $2^i \le n < 2^{i+1}$, we write

$$A_n = \big\{\|n\hat{\mu}_n\| \ge \varepsilon n^{\alpha}\big\},$$

$$A_n^{(1)} = \big\{\|f\circ X(\pi_1(k),\dots,\pi_d(k))\| \ge \varepsilon 2^{(i-2)\alpha} \text{ for at least one } k \le n\big\},$$

$$A_n^{(2)} = \big\{\|f\circ X(\pi_1(k_1),\dots,\pi_d(k_1))\| \ge \varepsilon n^{\theta\alpha},$$
$$\|f\circ X(\pi_1(k_2),\dots,\pi_d(k_2))\| \ge \varepsilon n^{\theta\alpha} \text{ for at least two } k_1, k_2 \le n\big\},$$

$$A_n^{(3)} = \left\{\left\|\sum_{k=1}^{n} f^{+}\circ X(\pi_1(k),\dots,\pi_d(k))\right\| \ge \varepsilon 2^{(i-2)\alpha}\right\}.$$

We observe that

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

Hence to prove the proposition, it suffices to show that

$$\sum_{n=1}^{\infty} n^{\alpha q-2} P\big(A_n^{(j)}\big) < \infty \qquad \forall 1 \le j \le 3.$$

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

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

Next we observe that

$$\begin{aligned} \sum_{n=1}^{\infty} n^{\alpha q-2} P\big(A_n^{(2)}\big) &\le \sum_{n=1}^{\infty} n^{\alpha q} P\big(\{\|f \circ X(\pi_1(1), \dots, \pi_d(1))\| \ge \varepsilon n^{\theta\alpha}\} \\ &\qquad\qquad \cap \{\|f \circ X(\pi_1(2), \dots, \pi_d(2))\| \ge \varepsilon n^{\theta\alpha}\}\big) \\ &\le \sum_{n=1}^{\infty} c n^{\alpha q} P^2\big(\|f \circ X\| \ge \varepsilon n^{\theta\alpha}\big) \qquad (15) \\ &\le \sum_{n=1}^{\infty} c^* n^{\alpha q - 2\theta\alpha q} \big(E\|f \circ X\|^q\big)^2 < \infty. \end{aligned}$$

Here the second inequality of (15) 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

$$\begin{aligned} &\sum_{n=1}^{\infty} n^{\alpha q-2} P\big(A_n^{(3)}\big) \\ &\quad \le c^* \sum_{n=1}^{\infty} n^{-2-\delta\alpha} E\left\{\left\| \sum_{k=1}^{n} f^+ \circ X(\pi_1(k), \dots, \pi_d(k)) \right\|^{q+\delta}\right\} \\ &\quad \le c^* \sum_{n=1}^{\infty} n^{(\theta-1)\delta\alpha - 1} E\|f \circ X\|^q < \infty. \end{aligned}$$

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

$$mq^*(2\alpha - 1) > \alpha q - 1. \qquad (16)$$

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

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

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

$$
\begin{aligned}
&\sum_{n=1}^{\infty} n^{\alpha q-2} P\left(A_n^{(3)}\right) \\
&\qquad \le \sum_{n=1}^{\infty} n^{\alpha q-2} P\left(\left\| \sum_{k=1}^{n} \tilde{f}^{+} \circ X(\pi_1(k),\dots,\pi_d(k)) \right\| \ge c^* n^{\alpha}\right) \\
&\qquad \le c^* \sum_{n=1}^{\infty} n^{\alpha q-2-2m\alpha q^*} E\left\{\left\| \sum_{k=1}^{n} \tilde{f}^{+} \circ X(\pi_1(k),\dots,\pi_d(k)) \right\|^{2mq^*}\right\} \\
(17)\qquad &\qquad \le c^* \sum_{n=1}^{\infty} n^{\alpha q-2-2m\alpha q^*} \left\{ \sum_{i=2mq^*-q^*+1}^{2mq^*-2} n^{2mq^*/(2mq^*-i)} \right. \\
&\qquad\qquad \left. + \sum_{i=0}^{(2mq^*-q^*)\wedge(2mq^*-2)} n^{2mq^*/(2mq^*-i)} \left(E\|\tilde{f}^{+} \circ X\|^{2mq^*-i}\right)^{2mq^*/(2mq^*-i)} \right\}.
\end{aligned}
$$

Now we observe from (16) that

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

and

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

The finiteness of $\sum_{n=1}^{\infty} n^{\alpha q-2} P(A_n^{(3)})$ follows from (17), (18) and (19). This proves Proposition 1. □

4. Nonparametric confidence regions. Owen (1988, 1990) 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 $\{f \circ X(\pi_1(k),\dots,\pi_d(k))\colon 1 \le k \le n\}$ be as in (1), $\mu = E(f \circ X)$ and $\mathscr{S} = \{w = (w_1,\dots,w_n)'\colon \sum_{k=1}^{n} w_k \le 1,\ w_k \ge 0\ \forall k\}$. Define for $0 < r < 1$,

$$
(20)\qquad \Theta_{n,r} = \left\{ \sum_{k=1}^{n} w_k f \circ X(\pi_1(k),\dots,\pi_d(k))\colon w \in \mathscr{S},\ \prod_{k=1}^{n} n w_k \ge r \right\},
$$

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

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

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

REMARK. A very nice discussion of the motivation (in terms of empirical likelihood ratio) for the preceding 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 nonrepresentative ones, that is, those that are not "evenly distributed" over $[0,1]^d$. Since empirical likelihiood 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 Theorem 4, which give the asymptotic validity as well as the convergence rates for the procedure.

THEOREM 4. *Let* $\{f \circ X(\pi_1(k), \ldots, \pi_d(k))\colon 1 \le k \le n\}$ *be as in* (1) *with* $\mu = E(f \circ X)$. *Also let* $\Theta_{n,r}$ *be as in* (20) *for some* $0 < r < 1$ *and let* $f_{\mathrm{rem}}(x)$ *be as in* (3) *such that* $\int_{[0,1]^d} f_{\mathrm{rem}}(x) f'_{\mathrm{rem}}(x)\,dx$ *is nonsingular. Then* $\Theta_{n,r}$ *is a convex set.*

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

$$\begin{aligned}\lim_{n\to\infty} P(\mu \in \Theta_{n,r}) &= P\big(Z' M^{1/2} (M+N)^{-1} M^{1/2} Z \le -2\log r\big)\\ &\ge P\big(\chi^2_{(p)} \le -2\log r\big),\end{aligned} \tag{22}$$

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

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

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

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) *Let* ε *be a constant satisfying* $0 < \varepsilon < 1/2$. *Then if* $E\|f \circ X\|^{10} < \infty$, *we have*

$$\big|P(\mu \in \Theta_{n,r}) - P\big(Z' \Sigma_{\mathrm{lhs}}^{1/2} \Sigma_{\mathrm{iid}}^{-1} \Sigma_{\mathrm{lhs}}^{1/2} Z \le -2\log r\big)\big| \le c^* n^{\varepsilon - 1/2},$$

where Σ_{lhs} *and* Σ_{iid} *are as defined in Theorem* 1.

REMARK. The limit in (22) 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.

REMARK. The moment conditions in Theorem 4 are probably excessive. However, under these conditions, the proof of Theorem 4 can be carried out using essentially only results from previous sections.

PROOF OF THEOREM 4. The convexity of $\Theta_{n,r}$ follows from Jensen's inequality and the observation that $(\Pi_{k=1}^{n} nw_k)^{1/n}$ is concave (strictly concave if $n \geq 2$) in $w \in \mathscr{S}$ [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 (22) it suffices to show

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

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

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

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

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

almost surely for sufficiently large n. This implies that μ is an interior point of the convex hull of $\{f\circ X(\pi_1(k),\dots,\pi_d(k))\colon 1 \leq k \leq n\}$ and $R_n(\mu) > c^*$ almost surely for large n. Using Lagrange multipliers, the solution of (21) is found to be

$$nw_k = (1+\gamma_k)^{-1}, \tag{23}$$

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

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

For simplicity we write

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

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

$$\begin{aligned}\zeta &= \eta - S^{-1}(\hat{\mu}_n - \mu)\\ &= n^{-1}S^{-1}\sum_{k=1}^{n}\{f\circ X(\pi_1(k),\dots,\pi_d(k)) - \mu\}\gamma_k^2/(1+\gamma_k).\end{aligned} \tag{24}$$

Next we note from (23) that

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

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 Corollary 2 and Theorem 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 (25) that

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

This proves (a).

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

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

it follows from (25) and Lemma 7 (see Appendix) that

$$
(26)\quad \begin{aligned}
&\left|P(R_n(\mu)\geq r)-P\left(Z'\Sigma_{\mathrm{lhs}}^{1/2}\Sigma_{\mathrm{iid}}^{-1}\Sigma_{\mathrm{lhs}}^{1/2}Z\leq -2\log r\right)\right| \\
&\quad\leq \max_{\delta=1,-1}\Big\{\Big|P\left\{n(\hat{\mu}_n-\mu)'(n\Sigma_{\mathrm{iid}})^{-1}(\hat{\mu}_n-\mu)\leq 2\delta n^{\varepsilon-1/2}-2\log r\right\} \\
&\qquad\qquad -P\left(Z'\Sigma_{\mathrm{lhs}}^{1/2}\Sigma_{\mathrm{iid}}^{-1}\Sigma_{\mathrm{lhs}}^{1/2}Z\leq 2\delta n^{\varepsilon-1/2}-2\log r\right)\Big| \\
&\qquad +\Big|P\left\{n(\hat{\mu}_n-\mu)'S^{-1}(\hat{\mu}_n-\mu)\leq 2\delta n^{\varepsilon-1/2}-2\log r\right\} \\
&\qquad\qquad -P\left\{n(\hat{\mu}_n-\mu)'(n\Sigma_{\mathrm{iid}})^{-1}(\hat{\mu}_n-\mu)\leq 2\delta n^{\varepsilon-1/2}-2\log r\right\}\Big|\Big\} \\
&\quad +c^*n^{\varepsilon-1/2}.
\end{aligned}
$$

Since $\{x\in\mathscr{R}^p\colon x'\Sigma_{\mathrm{lhs}}^{1/2}\Sigma_{\mathrm{iid}}^{-1}\Sigma_{\mathrm{lhs}}^{1/2}x\leq 2\delta n^{\varepsilon-1/2}-2\log r\}$ is a convex set in $\mathscr{R}^p$, we conclude from Corollary 2 that the right-hand side of (26) is bounded by $c^*n^{\varepsilon-1/2}$. This proves (b). □

APPENDIX

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

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

whenever $x \in \Pi_{i=1}^{r}[j_i/n, (j_i+1)/n)$ *for some* $0 \le j_1, \dots, j_r \le n-1$. *Then* $\|h - \rho_n(\cdot; h)\|_q \to 0$ *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: $[0,1]^r \to \mathscr{R}^p$, then we write

$$\rho_n(x;h) = \big(\rho_n(x;h_1), \dots, \rho_n(x;h_p)\big)', \qquad \forall x \in [0,1]^r. \tag{27}$$

PROOF OF THEOREM 1. We observe that

$$\begin{aligned} n\Sigma_{\mathrm{lhs}} = n^{-1} \sum_{i=1}^{n} & E[f \circ X(i, \pi_2(i), \dots, \pi_d(i)) - \mu] \\ & \times [f \circ X(i, \pi_2(i), \dots, \pi_d(i)) - \mu]' \\ + n^{-1} \sum_{i \ne j} & E[f \circ X(i, \pi_2(i), \dots, \pi_d(i)) - \mu] \\ & \times [f \circ X(j, \pi_2(j), \dots, \pi_d(j)) - \mu]'. \end{aligned} \tag{28}$$

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

$$\delta_n(s,t) = \begin{cases} 1, & \text{if } \lfloor ns \rfloor = \lfloor nt \rfloor, \\ 0, & \text{otherwise,} \end{cases} \tag{29}$$

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

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

Thus it follows from Lemma 3 (with $q = 2$) and (27) that

$$
\begin{aligned}
&\lim_{n\to\infty} n^{-1} \sum_{i\neq j} E[f\circ X(i,\pi_2(i),\dots,\pi_d(i)) - \mu] \\
&\qquad\qquad \times [f\circ X(j,\pi_2(j),\dots,\pi_d(j)) - \mu]' \\
&\quad = \lim_{n\to\infty} -n \sum_{i=1}^{n}\sum_{k=1}^{d} \int_{(i-1)/n}^{i/n} f_{-k}(x_k)\,dx_k \int_{(i-1)/n}^{i/n} f'_{-k}(y_k)\,dy_k \\
&\quad = \lim_{n\to\infty} -n^{-1} \sum_{i=1}^{n}\sum_{k=1}^{d} \rho_n((i-1)/n; f_{-k})\rho'_n((i-1)/n; f_{-k}) \\
&\quad = -\sum_{k=1}^{d} \int_0^1 f_{-k}(x_k) f'_{-k}(x_k)\,dx_k,
\end{aligned} \tag{30}
$$

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

$$
\begin{aligned}
&E(f\circ X - \mu)(f\circ X-\mu)' \\
&\quad = \int_{[0,1]^d} f_{\mathrm{rem}}(x) f'_{\mathrm{rem}}(x)\,dx + \sum_{k=1}^{d}\int_0^1 f_{-k}(x_k) f'_{-k}(x_k)\,dx_k.
\end{aligned} \tag{31}
$$

The theorem now follows from (28), (30) and (31). □

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

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

$$
\tilde{Y}(i_1,\dots,i_d) = \begin{cases} Y(i_1,\dots,i_d), & \text{if } \|Y(i_1,\dots,i_d)\| \leq 1/(4d), \\ 0, & \text{otherwise,} \end{cases} \tag{32}
$$

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

$$
\begin{aligned}
P(W \neq \tilde{W}) &\leq P\biggl(\sum_{i=1}^{n} I\{\|Y(i,\pi_2(i),\dots,\pi_d(i))\| > 1/(4d)\} \geq 1\biggr) \\
&\leq n^{1-d} \sum_{1\leq i_1,\dots,i_d\leq n} P(\|Y(i_1,\dots,i_d)\| > 1/(4d)) \\
&\leq (4d)^3\beta_3,
\end{aligned} \tag{33}
$$

and

$$
\begin{aligned}
\|\tilde{\mu}\| \leq n^{1-d} &\sum_{1\leq i_1,\dots,i_d\leq n} E\|Y(i_1,\dots,i_d)\| I\{\|Y(i_1,\dots,i_d)\| > 1/(4d)\} \\
&\leq (4d)^2\beta_3.
\end{aligned} \tag{34}
$$

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

$$\begin{aligned}
&\delta_{i,j} - \tilde{\Sigma}_{i,j} - \tilde{\mu}_i \tilde{\mu}_j \\
&\quad = E\Biggl\{ \sum_{a=1}^{n} Y_i(a, \pi_2(a), \dots, \pi_d(a)) Y_j(a, \pi_2(a), \dots, \pi_d(a)) \\
&\qquad\qquad \times I\{\|Y(a, \pi_2(a), \dots, \pi_d(a))\| > 1/(4d)\} \\
&\qquad\qquad - \sum_{a \ne b} Y_i(a, \pi_2(a), \dots, \pi_d(a)) \\
&\qquad\qquad\qquad \times I\{\|Y(a, \pi_2(a), \dots, \pi_d(a))\| > 1/(4d)\} \\
&\qquad\qquad\qquad \times Y_i(b, \pi_2(b), \dots, \pi_d(b)) \\
&\qquad\qquad\qquad \times I\{\|Y(b, \pi_2(b), \dots, \pi_d(b))\| > 1/(4d)\} \\
&\qquad\qquad + 2 \sum_{a \ne b} Y_i(a, \pi_2(a), \dots, \pi_d(a)) \\
&\qquad\qquad\qquad \times I\{\|Y(a, \pi_2(a), \dots, \pi_d(a))\| > 1/(4d)\} \\
&\qquad\qquad\qquad\qquad \times Y_i(b, \pi_2(b), \dots, \pi_d(b)) \Biggr\} \\
&\quad = \Delta_1 - \Delta_2 + 2\Delta_3, \quad \text{say.}
\end{aligned} \tag{35}$$

We note that

$$\begin{aligned}
1|\Delta_1| &\le n^{1-d} \sum_{1 \le i_1, \dots, i_d \le n} E\|Y(i_1, \dots, i_d)\|^2 I\{\|Y(i_1, \dots, i_d)\| > 1/(4d)\} \\
&\le 4\,d\beta_3,
\end{aligned} \tag{36}$$

and in a similar way,

$$|\Delta_2| \le \left(\frac{n}{n-1}\right)^{d-1} \left[(4d)^2 \beta_3\right]^2 \le \left(\frac{n}{n-1}\right)^{d-1} (4d)^4 \varepsilon_0 \beta_3. \tag{37}$$

Also we have from (5),

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

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

$$|\Delta_3| \le cn^{-1}\beta_3. \tag{38}$$

Since $\beta_3 \le \varepsilon_0$, it follows from (34)–(38) 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),

$$E\tilde{Y}(i_1, \dots, i_d) = \tilde{\mu}(i_1, \dots, i_d) \qquad \forall 1 \le i_1, \dots, i_d \le n,$$
$$\tilde{\mu}_{-k}(i_k) = (1/n^{d-1}) \sum_{j \ne k} \sum_{i_j = 1}^{n} \tilde{\mu}(i_1, \dots, i_d)$$

and

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

Now it follows from (32) that for sufficiently small $\varepsilon_0 > 0$,

$$\|\tilde{Y}^*(i_1, \dots, i_d)\| \le 1 \qquad \forall 1 \le i_1, \dots, i_d \le n, \tag{39}$$

and

$$(1/n^{d-1}) \sum_{1 \le i_1, \dots, i_d \le n} E\|\tilde{Y}^*(i_1, \dots, i_d)\|^3 \le c\beta_3. \tag{40}$$

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

$$\begin{aligned}
&\sup_{g \in \mathscr{A}} |E[g(W) - g(Z)]| \\
&\quad = \sup_{g \in \mathscr{A}} \Big| E\big[g(W) - g(Z) | W = \tilde{W}\big] P(W = \tilde{W}) \\
&\qquad\qquad + E\big[g(W) - g(Z) | W \ne \tilde{W}\big] P(W \ne \tilde{W}) \Big| \\
&\quad \le \sup_{g \in \mathscr{A}} \Big| E\big[g(\tilde{W}) - g(Z)\big]\Big| + 4P(W \ne \tilde{W}) \\
&\quad \le \sup_{g \in \mathscr{A}} \Big| E\big[g(\tilde{W}) - g(\tilde{Z})\big]\Big| + \Big| E\big[g(Z) - g(\tilde{Z})\big]\Big| + 4P(W \ne \tilde{W}),
\end{aligned}$$

where $\tilde{Z}$ denotes the p-variate normal random vector having the same mean and covariance matrix as $\tilde{W}$. Using (33), the Taylor expansion for the density of $\tilde{Z}$ and the fact that $\mathscr{A}$ is closed under affine transformations, we have

$$\sup_{g \in \mathscr{A}} |E[g(W) - g(Z)]| \le \sup_{g \in \mathscr{A}} \Big| E\Big[g\big(\tilde{\Sigma}^{-1/2}(\tilde{W} - E\tilde{W})\big) - g(Z)\Big]\Big| + c\beta_3. \tag{41}$$

Thus it follows from (39), (40) and (41) that to prove Theorem 2 it suffices to prove (6) under the assumption that $\|Y(i_1, \dots, i_d)\| \le 1$ for all $1 \le i_1, \dots, i_d \le n$. □

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

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

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

$$R_1 = \sum_{i,j=1}^{p} \Big\{ EnY_i^{(3)}(I_1, J_{2,2,1}, \dots, J_{d,d,1}) W_j^{(3)}$$

$$-EnY_i^{(3)}(I_1, J_{2,2,1}, \dots, J_{d,d,1}) EW_j^{(2)} - \delta_{i,j}\Big\} E \frac{\partial^2}{\partial w_i\, \partial w_j} \psi_{\varepsilon^2}(W^{(1)}) \tag{42}$$

$$= \sum_{i,j=1}^{p} \Big\{ EnY_i^{(3)}(I_1, J_{2,2,1}, \dots, J_{d,d,1}) W_j^{(3)} - \delta_{i,j}\Big\} E \frac{\partial^2}{\partial w_i\, \partial w_j} \psi_{\varepsilon^2}(W^{(1)}).$$

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

$$EnY_i^{(3)}(I_1, J_{2,2,1}, \dots, J_{d,d,1}) W_j^{(3)} = EW_i^{(3)} W_j^{(3)} = \delta_{i,j}. \tag{43}$$

It now follows from (42) and (43) that $R_1 = 0$.

For simplicity of notation, we write

$$\Omega_1 = \big\{ I_k, L_{j,k}: 2 \leq j \leq d,\ 1 \leq k \leq d \big\},$$

$$\Omega_2 = \big\{ J_{i,j,k}: 2 \leq i \leq d,\ 1 \leq j, k \leq d \big\}$$

and $\Omega_2 = \cdots = \Omega_d$. Let $\mathscr{C}$ denote the sigma-field generated by $\Omega_1 \cup \Omega_2 \cup \{Y^{(j)}(i, \pi_2^{(j)}(i), \dots, \pi_d^{(j)}(i)): i \in \Omega_1,\ 1 \leq j \leq 3\}$. Then

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

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

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

where $R_{s,t,\mathscr{C}}$ is a constant and $|\Omega_1|$ denotes the number of distinct elements in Ω_1. We note from (11) that $\Sigma_{s,t,\mathscr{C}}$ approximate arbitrarily closely to the identity matrix uniformly over $0 \le s, t \le 1$ and $\mathscr{C}$ by choosing ε_0 sufficiently small [see Bolthausen (1984), page 385, for a similar argument]. Define for all $i_k \notin \Omega_k$, $1 \le k \le d$,

$$E\big[V_{s,t,\mathscr{C}}(i_1, \dots, i_d) | \mathscr{C}\big] = \mu_{s,t,\mathscr{C}}(i_1, \dots, i_d),$$

$$\mu_{s,t,\mathscr{C},-k}(i_k) = \big(n - |\Omega_1|\big)^{1-d} \sum_{j \ne k} \sum_{i_j \notin \Omega_j} \mu_{s,t,\mathscr{C}}(i_1, \dots, i_d)$$

and

$$\begin{aligned} V^*_{s,t,\mathscr{C}}(i_1, \dots, i_d) = \Sigma_{s,t,\mathscr{C}}^{-1/2} \Bigg[& V_{s,t,\mathscr{C}}(i_1, \dots, i_d) - \sum_{k=1}^{d} \mu_{s,t,\mathscr{C},-k}(i_k) \\ & + (d-1) \mu_{s,t,\mathscr{C}} \Bigg]. \end{aligned}$$

Next we observe that

$$(44) \qquad \big(n - |\Omega_1|\big)^{1-d} \sum_{k=1}^{d} \sum_{i_k \notin \Omega_k} E\big(\|V^*_{s,t,\mathscr{C}}(i_1, \dots, i_d)\|^3 | \mathscr{C}\big) \le c\beta_3$$

and

$$(45) \qquad \begin{aligned} E\Big| nY_i^{(3)}(I_1, J_{2,2,1}, \dots, J_{d,d,1})\big(W_j^{(3)} - W_j^{(2)}\big)\Big| & \\ \times \big(|W_k^{(2)} - W_k^{(1)}| + |W_k^{(3)} - W_k^{(2)}|\big) \le c\beta_3. & \end{aligned}$$

Now it follows from (10), (44), (45), the induction hypothesis and the fact that $\mathscr{A}$ is closed under affine transformations that

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

Finally from (7), (8) and (45), we get $R_4 \le c\beta_3$ [see Bolthausen and Götze (1993), page 1703, for a similar argument]. This proves the lemma. □

LEMMA 6. *With the notation and assumptions of Proposition* 1, *we have*

$$E\left\{ \left\| \sum_{k=1}^{n} \tilde{f}^+ \circ X(\pi_1(k), \dots, \pi_d(k)) \right\|^{2mq^*} \right\} \le c^{**} n \lambda^{2mq^*}(n),$$

where c^{**} is a generic constant that depends only on d, p and $2mq^*$ and

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

PROOF. We observe that

$$\begin{aligned} &E\left\{\left\|\sum_{k=1}^{n} \tilde{f}^+ \circ X(\pi_1(k),\dots,\pi_d(k))\right\|^{2mq^*}\right\} \\ &\quad = E\left\{\sum_{i=1}^{p}\sum_{1\le k_1,k_2\le n}\big(\tilde{f}^+\circ X\big)_i(\pi_1(k_1),\dots,\pi_d(k_1)) \right. \\ &\qquad\qquad \left. \times\big(\tilde{f}^+\circ X\big)_i(\pi_1(k_2),\dots,\pi_d(k_2))\right\}^{mq^*}. \end{aligned} \tag{46}$$

On simplification, it can be seen that the right-hand side of (46) can be expressed as a finite sum (which depends only on p and mq^*) of terms each of the form

$$c^{**}\sum_{1\le k_1<\cdots<k_l\le n} E\prod_{t=1}^{l}\prod_{a_t=1}^{q_t}\big(\tilde{f}^+\circ X\big)_{i_{t,a_t}}(\pi_1(k_t),\dots,\pi_d(k_t)) \tag{47}$$

for some l such that $q_t \ge 1$, $1 \le i_{t,a_t} \le p$ for all $1 \le t \le l$ and $2mq^* = \sum_{t=1}^{l} q_t$. Now if $q_t \ge 2$ for all $1 \le t \le l$, we observe that the absolute value of (47) is bounded by

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

where $\delta_n(x_k^{(r)}, x_k^{(s)})$ is as in (29). Next we suppose that there exists a $q_t = 1$ for some $1 \le t \le l$. Without loss of generality, assume that $q_1 = \cdots = q_b = 1$ and $q_t \ge 2$ whenever $t > b \ge 1$. Then the absolute value of (47) is bounded by

$$\begin{aligned} &c^{**}n^l\left|\int_{[0,1]^{ld}}\left\{\prod_{t=1}^{b}\tilde{f}^+_{i_{t,1}}(x^{(t)})\right\}\left\{\prod_{t=b+1}^{l}\prod_{a_t=1}^{q_t}\tilde{f}^+_{i_{t,a_t}}(x^{(t)})\right\}\right| \\ &\qquad\times\left\{\prod_{1\le r<s\le l}\prod_{k=1}^{d}\big(1-\delta_n\big(x_k^{(r)},x_k^{(s)}\big)\big)\right\}dx^{(1)}\cdots dx^{(l)}. \end{aligned} \tag{48}$$

Since $E(\tilde{f}^+ \circ X) = 0$, by expanding the third product in (48), we observe that (48) can be rewritten as a finite sum (which depends only on d and l) of terms each of which is bounded by

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

for some l^*, q_t^*, where $q_t^* \ge 2$ for all $1 \le t \le l^*$ and $\sum_{t=1}^{l^*} q_t^* = 2mq^*$. This proves Lemma 6. □

LEMMA 7. *With the notation and assumptions of Theorem* 4(b),

$$P(n\zeta' S\zeta > n^{-1/2}) \le c^* n^{-1/2}, \tag{49}$$

$$P\left\{\left|\sum_{k=1}^{n} [2\log(1+\gamma_k) - 2\gamma_k + \gamma_k^2]\right| > n^{\varepsilon - 1/2}\right\} \le c^* n^{-1/2} \tag{50}$$

and

$$\begin{aligned} &\big|P\{n(\hat{\mu}_n - \mu)' S^{-1}(\hat{\mu}_n - \mu) > 2\delta n^{\varepsilon - 1/2} - 2\log r\} \\ &\quad - P\{n(\hat{\mu}_n - \mu)'(n\Sigma_{\text{iid}})^{-1}(\hat{\mu}_n - \mu) > 2\delta n^{\varepsilon - 1/2} - 2\log r\}\big| \\ &\quad \le c^* n^{\varepsilon - 1/2}. \end{aligned} \tag{51}$$

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

$$\begin{aligned} E\{\|\eta\|^{2j}\} &\le c^* n^{-j} \qquad \forall 1 \le j \le 3, \\ P\Big(\max_{1 \le k \le n} |\gamma_k| > 1/4\Big) &\le c^* n^{-1/2}. \end{aligned} \tag{52}$$

Since $E\|f \circ X\|^{10} < \infty$, it follows from Proposition 1 that

$$P\Big(\max_{1 \le i, j \le p} |(S - n\Sigma_{\text{iid}})_{i,j}| > n^{(\varepsilon/2) - (1/2)}\Big) \le c^* n^{-1/2}, \tag{53}$$

$$P\big(n\|\hat{\mu}_n - \mu\|^2 > n^{\varepsilon/2}\big) \le c^* n^{-1/2} \tag{54}$$

and

$$\begin{aligned} P\bigg(\bigg| n^{-1} \sum_{k=1}^{n} \|f \circ X(\pi_1(k), \dots, \pi_d(k)) - \mu\|^3 - E\|f \circ X - \mu\|^3\bigg| > 1\bigg) \\ \le c^* n^{-1/2}. \end{aligned} \tag{55}$$

By conditioning on the occurrence or nonoccurrence of $\{\max_{1\le k\le n}|\gamma_k| > 1/4\}$, it follows form (53) and the definition of ζ in (24) that for sufficiently large n,

$$
\begin{aligned}
&P(n\zeta' S\zeta > n^{-1/2})\\
&\quad\le P(\|\zeta\|^2 > c^* n^{-3/2}) + c^* n^{-1/2}\\
&\quad\le P\left(\left\{n^{-1}\sum_{k=1}^{n}\|f\circ X(\pi_1(k),\dots,\pi_d(k)) - \mu\|\gamma_k^2\right\}^2 > c^* n^{-3/2}\right)\\
&\qquad + c^* n^{-1/2}\\
&\quad\le P\left(n^{-1}\|\eta\|^2\sum_{k=1}^{n}\|f\circ X(\pi_1(k),\dots,\pi_d(k)) - \mu\|^3 > c^* n^{-3/4}\right)\\
&\qquad + c^* n^{-1/2}.
\end{aligned}
\tag{56}
$$

Thus it follows from (52), (55) and Markov's inequality that the right-hand side of (56) is bounded by $c^* n^{-1/2}$. This proves (49).

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

$$
\begin{aligned}
&P\left\{\left|\sum_{k=1}^{n}\left[2\log(1+\gamma_k) - 2\gamma_k + \gamma_k^2\right]\right| > n^{\varepsilon-1/2}\right\}\\
&\quad\le P\left(\sum_{k=1}^{n}|\gamma_k|^3 > c^* n^{\varepsilon-1/2}\right) + c^* n^{-1/2}\\
&\quad\le P\left(\|\eta\|^3\sum_{k=1}^{n}\|f\circ X(\pi_1(k),\dots,\pi_d(k))\|^3/n > c^* n^{\varepsilon-3/2}\right) + c^* n^{-1/2}\\
&\quad\le P\left(\|\zeta\|^3 + \|\hat\mu_n - \mu\|^3 > c^* n^{\varepsilon-3/2}\right) + c^* n^{-1/2}.
\end{aligned}
\tag{57}
$$

From (54) and (56), we observe that the right-hand side of (57) is bounded by $c^* n^{-1/2}$. This proves (50).

Finally observing that matrix inversion is a continuous operation for sufficiently large n, it follows from (53), (54) and (56) that

$$
\begin{aligned}
&P\{n(\hat\mu_n - \mu)' S^{-1}(\hat\mu_n - \mu) > 2\delta n^{\varepsilon-1/2} - 2\log r\}\\
&\quad -P\{n(\hat\mu_n - \mu)'(n\Sigma_{\text{iid}})^{-1}(\hat\mu_n - \mu) > 2\delta n^{\varepsilon-1/2} - 2\log r\}\\
&\qquad = P\{n(\hat\mu_n - \mu)'(n\Sigma_{\text{iid}})^{-1}(\hat\mu_n - \mu) > 2\delta n^{\varepsilon-1/2} - 2\log r\\
&\qquad\qquad\qquad -n(\hat\mu_n - \mu)'[S^{-1} - (n\Sigma_{\text{iid}})^{-1}](\hat\mu_n - \mu)\}\\
&\qquad\quad -P\{n(\hat\mu_n - \mu)'(n\Sigma_{\text{iid}})^{-1}(\hat\mu_n - \mu) > 2\delta n^{\varepsilon-1/2} - 2\log r\}\\
&\qquad \le c^* n^{-1/2} + P\{n(\hat\mu_n - \mu)'(n\Sigma_{\text{iid}})^{-1}(\hat\mu_n - \mu) \le 2\delta n^{\varepsilon-1/2} - 2\log r\}\\
&\qquad\quad -P\{n(\hat\mu_n - \mu)'(n\Sigma_{\text{iid}})^{-1}(\hat\mu_n - \mu)\\
&\qquad\qquad \le -c^* n^{\varepsilon-1/2} + 2\delta n^{\varepsilon-1/2} - 2\log r\}\\
&\qquad \le c^* n^{\varepsilon-1/2}.
\end{aligned}
$$

The last inequality follows from Corollary 2 and the observation that $\{x \in \mathscr{R}^p: x'\Sigma_{\text{lhs}}^{1/2}\Sigma_{\text{iid}}^{-1}\Sigma_{\text{lhs}}^{1/2}x \le 2\delta n^{\varepsilon-1/2} - 2\log r\}$ is a convex set in $\mathscr{R}^p$. Similarly we have

$$P\big\{n(\hat{\mu}_n - \mu)'S^{-1}(\hat{\mu}_n - \mu) > 2\delta n^{\varepsilon-1/2} - 2\log r\big\}$$
$$- P\big\{n(\hat{\mu}_n - \mu)'(n\Sigma_{\text{iid}})^{-1}(\hat{\mu}_n - \mu) > 2\delta n^{\varepsilon-1/2} - 2\log r\big\} \ge -c^* n^{\varepsilon-1/2}$$

and this proves (51). □

Acknowledgments. I would like to thank Professor Friedrich Götze for sending me preprints of his recent work and a referee for pointing out an oversight in an earlier proof of Proposition 1.

REFERENCES

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

BOLTHAUSEN, E. and GÖTZE, F. (1993). The rate of convergence for multivariate sampling statistics. *Ann. Statist.* **21** 1692–1710.

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

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

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.

HOEFFDING, W. (1951). A combinatorial central limit theorem. *Ann. Math. Statist.* **22** 558–566.

KATZ, M. (1963). The probability in the tail of a distribution. *Ann. Math. Statist.* **34** 312–318.

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

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.

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

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

OWEN, A. B. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18** 90–120.

OWEN, A. B. (1992). A central limit theorem for Latin hypercube sampling. *J. Roy. Statist. Soc. Ser. B* **54** 541–551.

ROYDEN, H. L. (1988). *Real Analysis*, 3rd ed. Macmillan, New York.

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.

STEIN, C. M. (1986). *Approximate Computation of Expectations*. IMS, Hayward, CA.

STEIN, M. L. (1987). Large sample properties of simulations using Latin hypercube sampling. *Technometrics* **29** 143–151.

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

DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
KENT RIDGE
SINGAPORE 119260
REPUBLIC OF SINGAPORE
E-MAIL: matlohwl@math.nus.sg