JOURNAL OF MULTIVARIATE ANALYSIS **56**, 42–59 (1996)
ARTICLE NO. 0003

On Edgeworth Expansion and Moving Block Bootstrap for Studentized *M*-Estimators in Multiple Linear Regression Models*

Soumendra Nath Lahiri

Iowa State University

This paper considers the multiple linear regression model $Y_i = x_i'\beta + \varepsilon_i$, $i = i, ..., n$, where x_i's are known $p \times 1$ vectors, β is a $p \times 1$ vector of parameters, and $\varepsilon_1, \varepsilon_2, ...$ are stationary, strongly mixing random variables. Let $\bar{\beta}_n$ denote an *M*-estimator of β corresponding to some score function ψ. Under some conditions on ψ, x_i's and ε_i's, a two-term Edgeworth expansion for Studentized multivariate *M*-estimator is proved. Furthermore, it is shown that the moving block bootstrap is second-order correct for some suitable bootstrap analog of Studentized $\bar{\beta}_n$. © 1996 Academic Press, Inc.

1. Introduction

Consider the linear regression model

$$Y_i = x_i'\beta + \varepsilon_i, \qquad i = 1, ..., n, \quad n \geqslant 1 \tag{1.1}$$

where $Y_1, ..., Y_n$ are the observations, x_i's are known $p \times 1$ design vectors, β is a $p \times 1$ vector of regression parameters and $\varepsilon_1, \varepsilon_2, ...$ is a stationary sequence of random variables (r.v.'s) defined on some probability space $(\Omega, \mathscr{F}, P)$. The design vectors $x_1, ..., x_n$ at the nth stage are allowed to depend on n, but we suppress that for notational simplicity.

The classical estimator of β is the least square estimator (LSE) $\hat{\beta}_n$, which is defined as a solution of the equation (in $t \in \mathbb{R}^p$)

$$\sum_{i=1}^{n} x_i(Y_i - x_i't) = 0. \tag{1.2}$$

Received February 2, 1993; revised October 1995.

AMS 1990 subject classifications; primary 62G05, secondary 62J05.

Key words and phrases: Edgeworth expansion, moving block bootstrap, *M*-estimators, multiple linear regression, stationarity, strong mixing, Studentization.

* Research partially supported by NSF Grant DMS-9107998.

0047-259X/96 $12.00

It is well known that the $\hat{\beta}_n$ is very sensitive to the outliers. A class of robust estimators of β is given by the so-called M-estimators (cf. Huber [13]). Let $\psi: \mathbb{R} \to \mathbb{R}$ be a Borel measurable function satisfying

$$E\psi(\varepsilon_1) = 0. \tag{1.3}$$

Then, an M-estimator $\bar{\beta}_n$ of β corresponding to ψ is defined as a solution to the following robustified version of Eq. (1.2):

$$\sum_{i=1}^{n} x_i \psi(Y_i - x_i' t) = 0. \tag{1.4}$$

When ε_i's are independent and identically distributed (i.i.d.), the asymptotic behaviour of Efron's [7] bootstrap method for Studentized $\bar{\beta}_n$ (and $\hat{\beta}_n$) has been investigated by several authors. See Freedman [8], Shorack [22], Bickel and Freedman [5], Qumsiyeh [21], Lahiri [16, 17], Tiro [24], and the references therein. For dependent ε_i's, however, the i.i.d. resampling scheme of Efron [7] is not appropriate (cf. Remark 2.1 of Singh [23]). A more effective resampling scheme (viz., the moving block bootstrap or the MBB, in short) for dealing with weakly dependent observations has been recently formulated by Künsch [14] and Liu and Singh [20]. For normalized statistics that are smooth functions of sample means of stationary, strongly mixing r.v.'s, the MBB is known to outperform the classical normal approximations. (cf. [12, 18]). The major objective of this paper is to establish the second-order correctness of the MBB for *Studentized* multivariate M-estimator of β under model (1.1).

For studying second-order properties of the MBB, in Section 2, we establish a two-term Edgeworth expansion for Studentized $\bar{\beta}_n$. Unlike the independent case, under dependence of ε_i's, we have to contend with a major technical difficulty, which does not show up in the i.i.d. case. Let $D_n = (\sum_{i=1}^{n} x_i x_i')^{-1/2}$ and $d_i = D_n x_i$, $1 \leqslant i \leqslant n$. Let I_r denote the identity matrix of order $r \geqslant 1$. When ε_i's are weakly dependent, the asymptotic covariance of $D_n^{-1}(\bar{\beta}_n - \beta)$ matrix is given by $\mathrm{COV}_n \equiv (E\psi'(\varepsilon_1))^{-2} \times \sum_{k=0}^{n} L_{kn} E\psi(\varepsilon_1)\, \psi(\varepsilon_{1+k})$, where $L_{0n} = I_p$ and $L_{kn} = \sum_{j=1}^{n-k} (d_j d_{j+k}' + d_{j+k} d_j')$, $1 \leqslant k \leqslant n-1$. Thus, to Studentize $(\bar{\beta}_n - \beta)$, one needs to estimate a progressively increasing number, say, l, of the lagged covariances $E\psi(\varepsilon_1) \times \psi(\varepsilon_{1+k})$, $0 \leqslant k \leqslant l$, such that $l \to \infty$ as $n \to \infty$. Since the number l of *estimated* lagged covariances tends to infinity, the well-known Edgeworth expansion techniques of Bhattacharya and Ghosh [3] and Bhattacharya [2] do not apply to this case.

In Sections 3 and 4, we develop necessary machinery for dealing with this issue. The key steps require some of the ideas developed by Bickel, Götze, and Van Zwet [6], Götze [10], and Götze and Hipp [11] (hereafter referred to as [GH]). Using the Edgeworth expansions for Studentized $\bar{\beta}_n$ and its bootstrap version, it is then shown in Theorem 2.2 that the MBB outperforms the normal approximation.

The rest of the paper is organized as follows. Section 2 gives the assumptions and the main results of the paper. Some auxiliary lemmas are proved in Section 3. Sections 4 and 5 give the proofs of the main results for the unbootstrapped and the bootstrapped M-estimators, respectively.

2. Main Results

The following notation is used throughout the paper. Let $\{\mathscr{D}_j: -\infty \leqslant j \leqslant \infty\}$ be a given sequence of sub-σ-fields of $\mathscr{F}$ with $\mathscr{D}_{-\infty} = \mathscr{D}_{\infty} = \{\phi, \Omega\}$. For any $-\infty \leqslant a \leqslant b \leqslant \infty$, let $\mathscr{D}_a^b = \sigma\langle \mathscr{D}_j: a \leqslant j \leqslant b\rangle$. For a matrix A, write A' for the transpose of A. Define $M_n = \max\{\|x_i\|: 1 \leqslant i \leqslant n\}$ and $\delta_0 = (p+11)^{-1}$. For a function $g: \mathbb{R} \to \mathbb{R}$, let g' and g'' denote the first and the second derivatives of g. Write $\|g\|_\infty$ for the supremum norm of a function g from a set X into $\mathbb{R}$. Let $\bar{\varepsilon}_i = Y_i - x_i'\bar{\beta}_n$ denote the ith residual, $1 \leqslant i \leqslant n$. Define

$$\sigma(k) = E\psi(\varepsilon_1)\,\psi(\varepsilon_{1+k}), \qquad k \geqslant 0; \qquad \tau = E\psi'(\varepsilon_1),$$

$$\hat{\sigma}_n(k) = (n-k)^{-1} \sum_{j=1}^{n-k} \psi(\bar{\varepsilon}_j)\,\psi(\bar{\varepsilon}_{j+k}), \qquad 0 \leqslant k \leqslant n-1, \quad \hat{\tau}_n = n^{-1} \sum_{j=1}^{n} \psi'(\bar{\varepsilon}_j).$$

For any set $A \in \mathbb{R}^p$, let $\partial A =$ the boundary of A, and $A^\eta = \{x \in \mathbb{R}^p: \|x-y\| < \eta$ for some $y \in A\}$, $\eta > 0$. Also, for any countable set B, write $|B|$ for the number of elements in B. Let $I(\cdot)$ denote the indicator function. Write Φ and ϕ for the distribution and the Lebesgue density of $N(0, I_p)$-distribution on $\mathbb{R}^p$, respectively. For any real number x, let $[x]$ denote the largest integer not exceeding x. In the following C, $C(\cdot)$ will denote generic constants, depending only on their arguments, if any. Whenever it is obvious, the dependence of $C(\cdot)$ on p, ρ and on the finite moments of $\psi(\varepsilon_i)$, $\psi'(\varepsilon_i)$, and $\psi''(\varepsilon_i)$ will be suppressed for notational simplicity. Unless otherwise stated, limits in order symbols are taken as $n \to \infty$.

Now we are ready to state the assumptions.

Assumptions. (A.1) (i) ψ is twice differentiable, and ψ'' satisfies a Lipschitz condition of order $\delta_1 > 0$,

(ii) ψ, ψ', ψ'' are bounded.

(A.2) (i) $E\psi(\varepsilon_1) = 0$, $\tau \equiv E\psi'(\varepsilon_1) \neq 0$,

(ii) $\sigma_\infty \equiv \sigma(0) - 2\sum_{k=1}^{\infty} |\sigma(k)| > 0$.

(A.3) There exists $\rho > 0$ such that

(i) $\sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathscr{D}^{n}_{-\infty},\ B \in \mathscr{D}^{\infty}_{n+m},\ n \geqslant 1\} \leqslant \rho^{-1} \exp(-\rho m)$ for all $m \geqslant 1$,

(ii) for all $n \geqslant 1$, and all $m \geqslant \rho^{-1}$, there exists a $\mathscr{D}^{n+m}_{n-m}$-measurable random variable $\tilde{\varepsilon}_{n,m}$ such that $E|\varepsilon_n - \tilde{\varepsilon}_{n,m}| \leqslant \rho^{-1} \exp(-\rho m)$,

(iii) for all $n, m, q \geqslant \rho^{-1}$ and $A \in \mathscr{D}^{n+q}_{n-q}$, $E|P(A \mid \mathscr{D}_j : j \neq n) - P(A \mid \mathscr{D}_j : 0 < |j-n| \leqslant q+m)| \leqslant \rho^{-1} \exp(-\rho m)$, and

(iv) for all $n \geqslant \rho^{-1}$, $m \leqslant n$ and all $t_{n-m}, ..., t_{n+m} \in \mathbb{R}$ with $|t_n| > \rho$, $E \mid E(\exp(\sqrt{-1} \sum_{j=n-m}^{n+m} t_j \psi(\varepsilon_j)) \mid \mathscr{D}_j : j \neq n)| < \exp(-\rho)$.

(A.4) $\operatorname{Max}\{\|x_i\| : 1 \leqslant i \leqslant n\} = 0(1)$ and $\liminf_{n \to \infty} n^{-1} \lambda_n \equiv \lambda > 0$, where λ_n denotes the smallest eigenvalue of $(\sum_{i=1}^{n} x_i x_i')$.

A few comments about the assumptions are in order. Assumption (A.1)(ii) has been assumed mainly to simplify the proofs of the theorems. (A.1)(ii) can be replaced by suitable moment conditions on $\psi(\varepsilon_1)$, $\psi'(\varepsilon_1)$, and $\psi''(\varepsilon_1)$, at the cost of considerably lengthier proofs. Since a robust M-estimator of β under model (1.1) *necessarily* corresponds to a *bounded* ψ (cf. [13]), this may not be as serious a restriction in applications as it appears at first sight. Similarly, Assumption (A.4) is used to simplify the proofs. For results on $\bar{\beta}_n$ that allow unbounded x_i's, see Lahiri [15]. Assumption (A.2)(ii) is needed here to ensure that $\operatorname{Cov}(\sum_{i=1}^{n} d_i \psi(\varepsilon_i))$ is nonsingular for n large. It is easy to construct examples (with suitable choice of x_i's), where (A.2)(ii) fails and $\bar{\beta}_n$ is not asymptotically normal on $\mathbb{R}^p$ with any Studentization.

Assumption (A.3) has been used crucially to establish the validity of Edgeworth expansions for Studentized $\bar{\beta}_n$ and its bootstrap version. Except for (A.3)(iv), the other conditions are comparable to the conditions introduced in the significant work of [GH]. (A.3)(iv) is stronger than the conditional Cramér conditions used by [GH]. Here the stronger version of the Cramér condition is needed to deal with the triangular array $\{d_j \psi(\varepsilon_j) : 1 \leqslant j \leqslant n;\ n \geqslant 1\}$, as compared to a fixed sequence of random vectors in [GH]. In the special case, where ε_j's are m-dependent and the conditional distribution of $\psi(\varepsilon_{m+1})$ given $\{\varepsilon_j : 1 \leqslant j \leqslant 2m+1, j \neq m+1\}$ has an absolutely continuous component with respect to the Lebesgue measure on $\mathbb{R}$, then (A.3)(i)–(iv) hold.

To define the Studentized version of $\bar{\beta}_n$, note that the asymptotic covariance matrix of $D_n^{-1}(\bar{\beta}_n - \beta)$ is given by

$$\Sigma_n \equiv \operatorname{Cov}\left(\sum_{i=1}^{n} d_i \psi(\varepsilon_i)\right) = \sum_{k=0}^{n-1} L_{kn} \sigma(k).$$

Therefore, a natural estimator of Σ_n is

$$\hat{\Sigma}_n = \sum_{k=0}^{l} L_{kn} \hat{\sigma}_n(k),$$

where $1 \leqslant l \equiv l_n \leqslant n-1$ is an integer. If $l \to \infty$ "slowly" with n, then $\|\hat{\Sigma}_n - \Sigma_n\| = o_p(1)$. By (A.2)(ii), $\hat{\Sigma}_n$ is non singular with high probability for n large, and can be inverted to define the Studentized statistic.

In contrast to the univariate case, where one uses the unique square root of $\hat{\Sigma}_n$ to Studentize $D_n^{-1}(\bar{\beta}_n - \beta)$, in the multivariate case the choice of the Studentizing matrix is not unique and requires some special treatments. Let $\mathscr{H}$ denote the set of all *pxp* positive definite (p.d.) matrices (on $\mathbb{R}$) and let $\mathscr{H}_1$ be the set of all $p \times p$ nonsingular matrices (on $\mathbb{R}$). Then $\mathscr{H}$ can be identified with an open subset $\mathscr{O}$ of $\mathbb{R}^{q_1}$, where $q_1 \equiv p(p+1)/2$. Furthermore, there exists a matrix valued function h: $\mathscr{H} \to \mathscr{H}_1$ such that for all $\Lambda \in \mathscr{H}$,

(i) $h(\Lambda)' \, h(\Lambda) = \Lambda^{-1}$, and (ii) if $\Lambda = ((\lambda_{ij}))$, and $\mathbf{\Lambda} = (\lambda_{11}, ..., \lambda_{1p}; \lambda_{22}, ..., \lambda_{2p}; ...; \lambda_{pp})'$, then the elements of $h(\Lambda)$, considered as functions of q_1-variables, are infinitely differentiable on $\mathscr{O}$. (2.1)

Indeed, there exist more than one function from $\mathscr{H}$ into $\mathscr{H}_1$ which satisfies requirements (i) and (ii) of (2.1). Roughly speaking, each such function defines a version of $\Lambda^{-1/2}$ for $\Lambda \in \mathscr{H}$. The most common choices of h come from the spectral decomposition and the Kholesky decomposition of a p.d. matrix. (cf. [19]). However, for the rest of this paper, it is *not* assumed that h is of a specific form. Fix any function h satisfying (2.1) , and define the Studentized M-estimator T_n by

$$T_n = h(\hat{\Sigma}_n) \, D_n^{-1}(\bar{\beta}_n - \beta).$$

The following result is useful for deriving Edgeworth expansions for T_n.

PROPOSITION 2.1. *Assume that* (A.1), (A.2), (A.3)(i), (ii), *and* (A.4) *hold. Then, there exists a sequence of statistics* $\{\bar{\beta}_n\}$ *such that*

$$P(\bar{\beta}_n \textit{ satisfies } (1.4) \textit{ and } \|D_n^{-1}(\bar{\beta}_n - \beta)\|^2 \leqslant C \log n) = 1 - o(n^{-1/2}). \quad (2.2)$$

Thus if (1.4) has a unique solution $\bar{\beta}_n$, then $\|D_n^{-1}(\bar{\beta}_n - \beta)\| = O_p((\log n)^{1/2})$. The next result asserts the validity of a two-term Edgeworth expansion for T_n.

THEOREM 2.1. *Assume that Assumptions* (A.1)–(A.4) *hold and that* $\{\bar{\beta}_n\}$ *is a sequence of measurable solutions of* (1.4) *satisfying* (2.2). *Suppose that*

$n^{\delta} l^{-1} = O(1)$ and $l = O(n^{(1-\kappa)/4})$ for some $\delta > 0$, and $\kappa > \max\{p+3, 5\} \cdot \delta_0$. Then, there exist polynomials $p_n(\cdot)$ on $\mathbb{R}^p$ such that

$$\sup_{B \in \mathscr{B}} |P(T_n \in B) - \int_B (1 + p_n(x))\, d\Phi(x)| = o(n^{-1/2})$$

for every class $\mathscr{B}$ of Borel subsets of $\mathbb{R}^p$ satisfying

$$\sup_{B \in \mathscr{B}} \Phi((\partial B)^{\eta}) = O(\eta) \qquad \text{as} \quad \eta \downarrow 0. \tag{2.3}$$

Here $\|p_n \phi\|_{\infty} = O(n^{-1/2})$, and the coefficients of $p_n(\cdot)$ are continuous functions of cross-product moments of $\psi(\varepsilon_i)$, $\psi'(\varepsilon_i)$, and $\psi''(\varepsilon_i)$.

An exact expression for the Fourier transform of the expansion for T_n is given by relation (4.10) in Section 4. Even under the stationarity assumption on ε_i's, the form of the expansion is utterly complicated, making the empirical Edgeworth expansion for T_n unfit for practical applications. As a result, the question of second-order correctness of the bootstrap approximation for T_n becomes more important in this case than in the i.i.d. error case.

To define the bootstrap version of T_n, first form the "observed" blocks of residuals of length l as $\xi_j = (\bar{\varepsilon}_j, ..., \bar{\varepsilon}_{j+l-1})$, $1 \leqslant j \leqslant b$, where $b = n - l + 1$ and $\bar{\varepsilon}_j = Y_j - x_j' \bar{\beta}_n$, $1 \leqslant j \leqslant n$. Next draw $\xi_1^*, ..., \xi_{k_0}^*$ randomly, with replacement from $\xi_1, ..., \xi_b$, where $k_0 = [n/l]$. Note that each ξ_k^* has l components. Denote the ith component of ξ_k^* by ξ_{ki}^*, $1 \leqslant i \leqslant l$. Also, set $\xi_{ki}^* = \varepsilon_{(l-1)k+i}^*$, $1 \leqslant i \leqslant l$, $1 \leqslant k \leqslant k_0$, and define the bootstrap pseudo-observations

$$Y_i^* = x_i' \bar{\beta}_n + \varepsilon_i^*, \qquad 1 \leqslant i \leqslant n_1,$$

where $n_1 = k_0 l$. Adapting Shorack's [22] approach, define the bootstrapped M-estimator β_n^* as a solution of the equation in $t \in \mathbb{R}^p$,

$$\sum_{i=1}^{n_1} x_i (\psi(Y_i^* - x_i' t) - \hat{\mu}_n) = 0, \tag{2.4}$$

where $\hat{\mu}_n = l^{-1} E_n\{\psi(\varepsilon_1^*) + \cdots + \psi(\varepsilon_l^*)\}$, and E_n denotes the conditional expectation under the MBB resampling scheme, given $\varepsilon_1, ..., \varepsilon_n$. Centering ψ by $\hat{\mu}_n$ makes the estimating equation (2.4) conditionally unbiased at $t = \bar{\beta}_n$ and ensures the bootstrap analog of (1.3).

The following result shows that conclusions similar to Proposition 2.1 hold for β_n^* as well.

PROPOSITION 2.2. *Assume that the conditions of Theorem* 2.1 *hold. Then, there exists a sequence of random vectors* $\{\beta_n^*\}$ *satisfying*

$$P_n(\beta_n^* \text{ satisfies } (2.4) \text{ and } \|D_n^{-1}(\beta_n^* - \bar{\beta}_n)\|^2 \leqslant C \log n) = 1 - o_P(n^{-1/2}), \quad (2.5)$$

where P_n *denotes the probability under* E_n.

To define the bootstrap version of T_n, note that by the independence of the resampled blocks, $\Sigma_n^* \equiv$ the conditional covariance matrix of $\sum_{i=1}^{n_1} d_i \psi(\varepsilon_i^*)$ is given by $\Sigma_n^* = \sum_{k=1}^{k_0} \mathrm{Cov}_n(\sum_{j=1}^{l} d_{(k-1)l+j} \psi(\varepsilon_j^*)) = \sum_{j=0}^{l-1} \sum_{k=1}^{k_0} \sum_{i=1}^{l-j} D_{ikj}^* \sigma_n^*(i, j)$, where $D_{ikj}^* = (1 - 2^{-1} I(j=0))(\tilde{D}_{ikj}^* + \tilde{D}_{ikj}^{*\prime})$, $\tilde{D}_{ikj}^* = d_{(k-1)l+i} d'_{(k-1)l+i+j}$, $\sigma_n^*(i, j) = b^{-1} \sum_{k=i}^{b+i-1} \tilde{\psi}(\bar{\varepsilon}_k)\, \tilde{\psi}(\bar{\varepsilon}_{k+j})$, and $\tilde{\psi} = \psi(\cdot) - \hat{\mu}_n$.

Since l is small compared to b and $1 \leqslant i \leqslant l$, $\sigma_n^*(i, j)$'s are uniformly (in i) close to $\sigma^*(1, j)$ for all $0 \leqslant j < l$. Hence, by the independence of $\{\xi_k^* : k = 1, ..., k_0\}$, a natural "estimator" of $\sigma^*(1, j)$'s is given by

$$\hat{\sigma}_n^*(j) = (k_0(l-j))^{-1} \sum_{k=1}^{k_0} \sum_{i=1}^{l-j} \tilde{\psi}(\bar{\varepsilon}^*_{(k-1)l+i})\, \tilde{\psi}(\bar{\varepsilon}^*_{(k-1)l+i+j}),$$

where $\bar{\varepsilon}_i^* = Y_i^* - x_i' \beta_n^*$ is the ith bootstrap residual, $1 \leqslant i \leqslant n_1$. Let

$$\hat{\Sigma}_n^* = \sum_{j=0}^{l_2} \sum_{k=1}^{k_0} \sum_{i=1}^{l-j} D_{ikj}^* \hat{\sigma}_n^*(j), \qquad \hat{\tau}_n^* = n_1^{-1} \sum_{1 \leqslant i \leqslant n_1} \psi'(\bar{\varepsilon}_i^*), \quad (2.6)$$

where $l_2 = [l^{1/2}]$. Then, the bootstrap version T_n^* of T_n is given by

$$T_n^* = h(\hat{\Sigma}_n^*)\, D_n^{-1}(\beta_n^* - \bar{\beta}_n)/\hat{\tau}_n^* \quad (2.7)$$

and we have the following result.

THEOREM 2.2. *Assume that the conditions of Theorem* 2.1 *hold. Suppose that* T_n^* *is defined by* (2.7) *for some measurable sequence* $\{\beta_n^*\}$ *satisfying* (2.5). *Then*

$$\sup_{B \in \mathscr{B}} |P_n(T_n^* \in B) - P(T_n \in B)| = o_P(n^{-1/2}) \quad (2.8)$$

for any class $\mathscr{B}$ *of Borel subsets of* $\mathbb{R}^p$ *satisfying* (2.3).

Remark 2.1. The conclusion of Theorem 2.2 continues to hold for the MBB with a different block size l_1 as long as $n^{\delta} l_1^{-1} = o(1)$ and $l_1 = O(n^{(1-\kappa)/4})$ for some $\delta > 0$ and $\kappa > \delta_0 \max\{p+3, 5\}$.

Theorem 2.2 shows that the MBB indeed provides more accurate approximation for Studentized *multivariate* M-estimator of the regression parameter vector β than the normal approximation. It may be noted that, except in some special cases (e.g., p-population sample means from p *independent* populations), the important issue of Studentization for multivariate estimators itself has received very little attention in the literature, let alone their asymptotic properties. Consequently, Theorem 2.2 should prove particularly useful for constructing second-order correct multivariate inference procedures (e.g., tests and confidence regions) for β under model (1.1).

3. LEMMAS

We need to introduce some more notation. Let $m_3 = [\log n \log\log(3+n)]$, $v_{1n} = n^{-1/2}(\log n)^{1/2}$, $v_n = n^{-1/2}$, $v_{2n} = v_n(\log n)^{-1}$, and $v_{3n} = v_n(\log n)^{-3/2}$. For any two real numbers u, v let $u \wedge v = \min\{u, v\}$ and $u \vee v = \max\{u, v\}$. Let $\mathbf{Z}_+ = \{0, 1, ...\}$. For $\alpha = (\alpha_1, ..., \alpha_q)' \in (\mathbf{Z}_+)^q$, and $x = (x_1, ..., x_q)' \in \mathbb{R}^q$, $q \geqslant 1$, define $x^\alpha = \prod_{i=1}^q x_i^{\alpha_i}$, $|\alpha| = a_1 + \cdots + \alpha_q$, and $\alpha! = \prod_{i=1}^q (\alpha_i!)$. For a function $g: \mathbb{R}^q \to \mathbb{R}$, let $D_j g$ denote the jth partial derivative of g. For $\alpha \in (\mathbf{Z}_+)^q$, set $D^\alpha g = D_1^{\alpha_1} \cdots D_q^{\alpha_q} g$. Let

$$\theta_n = D_n^{-1}(\bar{\beta}_n - \beta), \qquad \theta_{1n} = \sum_{i=1}^n d_i Z_{1i},$$

$$A_n = \sum_{i=1}^n d_i d_i' \psi'(\varepsilon_i), \qquad A = \tau I_p$$

$$\begin{gathered} Z_{1i} = \psi(\varepsilon_1), \qquad Z_{2i} = \psi'(\varepsilon_i) - \tau, \qquad Z_{3i} = \psi''(\varepsilon_i) - E\psi''(\varepsilon_1), \\ Z_{4i}(k) = \psi(\varepsilon_i)\,\psi(\varepsilon_{i+k}) - \sigma(k), \\ Z_{5i}(k) = \psi'(\varepsilon_i)\,\psi(\varepsilon_{i+k}) - E\psi'(\varepsilon_1)\,\psi(\varepsilon_{1+k}) \\ Z_{6i}(k) = \psi(\varepsilon_i)\,\psi'(\varepsilon_{i+k}) - E\psi(\varepsilon_1)\,\psi'(\varepsilon_{1+k}), \qquad i \geqslant 1, k \geqslant 0. \end{gathered} \tag{3.1}$$

Similarly, with $\xi_{kj}^* = j$th component of ξ_k^*, $1 \leqslant k \leqslant k_0$, and $\tilde{\psi}(\cdot) \equiv \psi(\cdot) - \hat{\mu}_n$, let $Z_{1k}^* = \sum_{j=1}^l d_{(k-1)l+j} \tilde{\psi}(\xi_{kj}^*)$ and $Z_{2k}^* = \sum_{j=1}^l d_{(k-1)l+j} d'_{(k-1)l+j} \psi'(\xi_{kj}^*)$, $1 \leqslant k \leqslant k_0$. Also, write $\sum_{1k}(\cdot) = \sum_{k=1}^{k_0}(\cdot)$, $f(y) = \exp(\sqrt{-1}\, y)$, $y \in \mathbb{R}$, and $\chi(U) = (-1)^{q/2} D_1 \cdots D_q Ef(t'U)|_{t=0}$ for a random vector U in $\mathbb{R}^q$.

LEMMA 3.1. *Let $W_1, ..., W_n$ be random variables defined on a probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, P)$ with $EW_i = 0$, $|W_i| \leqslant a_{1n}$. For $1 \leqslant m \leqslant n-1$, let $\tilde{\alpha}_n(m) =$ $\mathrm{Sup}\{|P(A \cap B) - P(A)\,P(B)| : A \in \tilde{\mathscr{F}}_1^q,\ B \in \tilde{\mathscr{F}}_{q+m}^n,\ 1 \leqslant q \leqslant n-m\}$, where $\tilde{\mathscr{F}}_a^b \equiv \sigma\langle W_i : a \leqslant i \leqslant b\rangle$. Then, for any integer $r \geqslant 1$ and $1 \leqslant m \leqslant C(r)\,n$,*

(a) $E(W_1 + \cdots + W_n)^{2r} \leqslant C(r) \cdot a_{1n}^{2r}[n^r m^{2r} + n^{2r}\tilde{\alpha}_n(m)]$, *and*

(b) *if* $ma_{1n}^2 \leqslant n$, *and* $\max\{EW_i^2 : 1 \leqslant 1 \leqslant n\} \leqslant \sigma_1^2$, *then*

$$E(W_1 + \cdots + W_n)^{2r} \leqslant C(r) \cdot n^{2r} a_{1n}^{2r} \tilde{\alpha}_n(m) + C(r, \sigma_1)[n^r + a_{1n}^{-2}(n/a_{1n})^{2r}].$$

Proof. For part (a), see Lemma 3.1 in Lahiri [18]. The second part can be proved using similar arguments. The details are omitted. ∎

LEMMA 3.2. *Under the conditions of Proposition* 2.1,

(a) $P(\|\sum_{i=1}^n d_i \psi(\varepsilon_i)\| > C(\log n)^{1/2}) = o(v_n)$,

(b) *for any integer* $r \geqslant 1$,

$$\operatorname{Sup}\left\{E\left\|\sum_{i \in I} d_i \psi(\varepsilon_i)\right\|^{2r} : I \subseteq \{1, ..., n\}\right\} \leqslant C(r).$$

Proof. Under Assumptions (A.1), (A.3)(i)–(iii), and (A.4), the arguments in the proofs of Lemma 3.16, 3.17, 3.20, and 3.33 of [GH] yield

$$\left|D^\alpha\left\{Ef\left(t'h(\Sigma_n)\sum_{i \in I} d_i \psi(\varepsilon_i)\right) - \exp(-\|t\|^2/2)(1 + E(\sqrt{-1}\, t'h(\Sigma_n)\,\theta_{1n})^3/3!)\right\}\right|$$
$$\leqslant C(\alpha, \delta)[n(n^\delta M_n)^4 + m_3^2 M_n^2](1 + \|t\|^{6+|\alpha|}) \exp(-C(\rho)\, \|t\|^2) \qquad (3.2)$$

uniformly in $I \subseteq \{1, ..., n\}$ with $|I| \geqslant n - m_3^2$ and in $\alpha \in (\mathbf{Z}_+)^p$ for some $0 < \delta < \delta_0/4$, provided $\|t\| < C(\alpha, \delta)\, n^{\delta/2}$. Lemma 3.2(a) can now be proved using the arguments in the proofs of Theorems 2.10 and 2.11 of [GH].

Next, by (3.13) and (3.14) of [GH], for any integer $r \geqslant 3$ and any $\mathbf{a} \in \mathbb{R}^p$ with $\|\mathbf{a}\| = 1$,

$$|\chi_{r,I}(\mathbf{a})| \equiv \left|\text{the } r\text{th cumulant of } \left(\sum_{i \in I} \mathbf{a}'\, d_i \psi(\varepsilon_i)\right)\right|$$
$$\leqslant C(r) \sum_{i=0}^{|I|-1} \sum_r^{(i)} |\chi(\psi(\varepsilon_{i_r}), ..., \psi(\varepsilon_{i_r}))| \left(\prod_{j=1}^r \|d_{i_j}\|\right)$$
$$\leqslant C(r) \sum_{0 \leqslant i \leqslant m_3} |I| \cdot i^{r-1} M_n^r + O(\exp(-C(\rho)\, m_3)),$$

where $\sum_r^{(i)}$ extends over all $i_1, ..., i_r \in I$ with maximal gap i. Part (b) follows from this. ∎

LEMMA 3.3 *Assume that the conditions of Theorem* 2.1 *hold.*

(a) *For any integer* $r \geqslant 1$, *and for* $N = 1$ *or* $[n^{1/2-\delta_0}]$,

$$E\left\|\sum_{k=0}^{l} L_{kn}(n-k)^{-1} \sum_{i=N}^{(n-k)} Z_{4i}(k)\right\|^{2r} \leqslant C(r)\, n^{-r} l^r (1 + n^{-1} l^{r+1}).$$

(b) $$P\left(\left\|\sum_{k=0}^{l}(n-k)^{-1}L_{kn}\sum_{i=1}^{n-k}(d_iZ_{5i}(k)+d_{i+k}Z_{6i}(k))\right\|>Cv_n(\log n)^{-2}\right)$$
$$=o(v_n).$$

Proof. Lemma 3.3 can be proved using Markov's inequality and Lemma 3.1. See Lahiri [15] for details. ∎

LEMMA 3.4. *Let* $I_j=\{jp+1, ..., (j+1)\,p\}$, $j=0, 1, ..., [n/p]-1$, *and for* $c>0$ *let* $A_n(c)=\{1\leqslant j<[n/p] : \inf\{\sum_{i\in I_j}(x_i't)^2 : \|t\|=1\}>cp\}$. *If Assumption* (A.4) *holds, then for all* $0<c<c_o$, $\lambda_1(c)\equiv\liminf_{n\to\infty} n^{-1}|A_n(c)|>0$, *where* $c_o=2^{-1}\cdot\min\{\lambda, \limsup_{n\to\infty} M_n\}$.

Proof. Follows from (A.4) and the inequality $\lambda_n\leqslant M_np\,|A_n(c)|+([n/p]-1-|A_n(c)|)\,cp+pM_n$. ∎

Since x_i's are assumed to depend on n (by a reindexing, if necessary), from now on, w.l.g. assume that $A_n(c)=\{0, 1, ..., |A_n(c)|-1\}$ for all $n\geqslant 1$.

LEMMA 3.5. *Suppose that the conditions of Theorem* 2.1 *hold. Let* $N\equiv N_n$ *be an integer satisfying* $n^{1/2-\delta_0}\leqslant N\leqslant |A_n(c_o)|\cdot p$. *Then, for* $t\in\mathbb{R}^p$,

$$\left|EB_{1n}(t)\,B_{2n}f\left(t'\sum_{i=1}^{N}d_iZ_{1i}\right)\right|$$
$$\leqslant\exp(-C(c_o, K, \Sigma)\cdot N/m_3)\cdot 1(\|t\|^2>\lambda pn)$$
$$+\exp(-C(c_o, K, \Sigma)\,\|t\|^2\,Nn^{-1}m_3^{-1})\cdot 1(\|t\|^2<\lambda pn)$$
$$+\exp(-C(K)\,m_3),$$

where $B_{1n}(t)=\prod_{i=N+1}^{n}\prod_{j=1}^{r_i}(1+\mathbf{a}_{ij}'Z_i))\,f(t'd_iZ_{1i})$ *and* $B_{2n}=\prod_{j=1}^{K}(1+\mathbf{a}_{Nj}'Z_{i_j})$, *for some* $\|\mathbf{a}_{ij}\|\leqslant 1$, $K\in\mathbf{Z}_+$, $1\leqslant i_1, ..., i_K\leqslant N$, *and* $\sum_{i=N+1}^{n}r_i\leqslant K$, *and* $Z_i=(Z_{1i}, Z_{2i}, Z_{3i})'$.

Proof. Define $I_1=\{1\leqslant i\leqslant N-m_3 : |i-i_k|>m_3,\ 1\leqslant k\leqslant K\}$, $I_2=\{1\leqslant i\leqslant N-m_3 : |i-i_k|>m_3+p,\ 1\leqslant k\leqslant K\}$, $l_3=[\{(N-m_3)-(2m_3+2p+1)\,K\}(7m_3+2p)^{-1}]$, $i_1''=\text{Inf}\{i: i\in I_2\}$, and $i_{r+1}''=\text{Inf}\{i\in I_2 : i>i_r''+7m_3+2p\}$, $r=1, ..., l_3$. Let $\bar{A}_k=\prod\{f(t'h(\Sigma_n)\,d_jZ_{1j}) : |j-i_k''|\leqslant m_3,\ j\in I_1\}$, $k=1, ..., l_3$. Then, using the arguments in the proof of Lemma 3.43 of [GH], one gets

$$\left|EB_{1n}(t)\,B_{2n}f\left(t'h(\Sigma_n)\sum_{j=1}^{N}d_jZ_{1j}\right)\right|$$
$$\leqslant C(K)\prod_{k=1}^{pl_3}E\,|E(\bar{A}_k\mid\mathscr{D}_j\colon j\neq i_k)|+\exp(-C(K)\,m_3). \qquad (3.3)$$

Also, in Lemma 3.2 of [GH], for any $m \geqslant 1$, $i \geqslant \rho^{-1}$, and $t_1, ..., t_{2m+1} \in \mathbb{R}$ with $|t_{m+1}| \leqslant \rho$,

$$
\begin{aligned}
E \left| E \left(f \left(\sum_{j=1}^{2m+1} t_j Z_{1,i+j} \right) \middle| \mathscr{D}_j : j \neq i+m+1 \right) \right|^2 &\\
\leqslant 1 - 4^{-r} \left(1 - E \left| E \left(f \left(2^r \sum_{j=1}^{2m+1} t_j Z_{1,i+j} \right) \middle| \mathscr{D}_j : j \neq i+m+1 \right) \right|^2 \right) &\\
\leqslant 1 - 4^{-r} (1 - \exp(-C(\rho))) \leqslant \exp(-C(\rho)\, t_{m+1}^2), & \qquad (3.4)
\end{aligned}
$$

where r is such that $2^r\, |t_{m+1}| > \rho \geqslant 2^{r-1}\, |t_{m+1}|$. Lemma 3.5 now follows from (3.3) and (3.4) for $\|t\| > \lambda pn$ and from Lemma 3.4 for $\|t\|^2 \leqslant \lambda pn$.

LEMMA 3.6. *Assume that the conditions of Theorem 2.1 hold. Then*

(a) $|\hat{\mu}_n| = O_p(n^{-1} l)$.

(b) *For any integer* $r \geqslant 1$, $\sum_{k=1}^{N} E_n \|Z_{1k}^*\|^{2r} = O_p(N(lm_3^2 M_n^2)^4)$

for $N = k_0$ *and* $N = [l^{-1} n^{(1-2\gamma)/2}]$, *where* $\gamma = \max\{1, p-1\}\, \delta_0/4$.

Proof. Part (a) is a consequence of (A.1) and (1.4). Part (b) follows from Lemma 3.1(a), Proposition 2.1, and part (a) above. ▌

LEMMA 3.7. *Assume that the conditions of Theorem 2.2 hold. Then,*

(a) $\|\sum_{1k} E_n Z_{2k}^* - A\| = O_p(v_n m_3^2)$,

(b) $n_1^{-1} \sum_{i=1}^{n_1} d_i (E_n \psi''(\varepsilon_i^*) - E\psi''(\varepsilon_1)) = O_p(v_n (\log n)^{-2})$,

(c) $\|E_n(\sum_{1k} Z_{1k}^* Z_{1k}^{*\prime}) - \Sigma_n\| = o_p(1)$.

Proof. $\|E_n \sum_{1k} Z_{2k}^* - A\|$ is bounded above by $C\, |b^{-1} \sum_{i=1}^{b} Z_{2i}| + \|\sum_{i=1}^{n_1} d_i d_i' - I_p\| \cdot |\tau| + Cb^{-1} \sum_{i=1}^{b} \|d_i\|\, \|\theta_n\| + Cn^{-1} l$, which is $O_p(n^{-1/2} m_3^2)$, by Lemma 3.1 and Proposition 2.1. This proves part (a). Proof of part (b) is similar. Part (c) can be proved using (A.3)(i), (ii), Lemmas 3.1 and 3.6. See Lahiri [15] for further details. ▌

4. PROOFS OF PROPOSITION 2.1 AND THEOREM 2.1

Let $e_1, ..., e_p$ denote the standard basis of $\mathbb{R}^p$. Write $a = [n^{(1-2\delta_0)/2}]$, and for $1 \leqslant i \leqslant n$, $0 \leqslant k \leqslant l$, define $\gamma_{ik} \equiv \gamma_{ikn} = \tau^{-1} [d_i E\psi(\varepsilon_{1+k})\, \psi'(\varepsilon_1) + d_{i+k} E\psi(\varepsilon_1)\, \psi'(\varepsilon_{1+k})]$.

Proof of Proposition 2.1. Let $\Delta = D_n^{-1}(t-\beta)$, $t \in \mathbb{R}^p$. Then, by Taylor's expansion, one can rewrite Eq. (1.4) as

$$\left[\sum_{i=1}^{n} d_i d_i' \psi'(\varepsilon_i)\right] \Delta = \sum_{i=1}^{n} d_i \psi(\varepsilon_i) + 2^{-1} \sum_{i=1}^{n} d_i (d_i' \Delta)^2 \psi''(\varepsilon_i) + R_{1n}(t), \tag{4.1}$$

where $\|R_{1n}(t)\| \leqslant C \sum_{i=1}^{n} \|d_i\|^{3+\delta_1} \|\Delta\|^{2+\delta_1}$, $t \in \mathbb{R}^p$. Note that by Lemmas 3.1 and 3.2,

$$\begin{aligned} &P\left(\left\|\sum_{i=1}^{n} d_i \psi(\varepsilon_i)\right\| > C(\log n)^{1/2}\right) = o(n^{-1/2}), \\ &P(\|A_n - A\| > Cn^{-1/4}(\log n)^{-2}) = o(n^{-1/2}), \qquad (4.2) \\ &P\left(\left|\sum_{i=1}^{n} d_{ik} d_{ij} d_{im} Z_{3i}\right| > Cn^{-5/8}\right) \leqslant Cn^{-3/4} \end{aligned}$$

for all $1 \leqslant j, k, m \leqslant p$, where d_{ij} denotes the jth component of d_i. The proof of Proposition 2.1 can now be completed as in Bhattacharya and Ghosh [3]. ∎

Proof of Theorem 2.1. *Step* (I): *Stochastic Approximation for* T_n. Using (4.1), (4.2), and Proposition 2.1, one can retrace the steps in Lahiri [16] to show that

$$\theta_n = (A^{-1} + \tau^{-2}(A - A_n))\, \theta_{1n} + (2\tau^3)^{-1} \sum_{i=1}^{n} d_i (d_i' \theta_{1n})^2 E\psi''(\varepsilon_i) + R_{2n}, \tag{4.3}$$

where $P(\|R_{2n}\| > C(\sigma_\infty)\, v_{2n}) = o(v_n)$. Next, note that by Taylor's expansion,

$$\begin{aligned} \hat{\tau}_n - \tau &= n^{-1} \sum_{i=1}^{n} Z_{2i} - (n\tau)^{-1} \sum_{i=1}^{n} d_i' \theta_{1n} \cdot E\psi''(\varepsilon_1) + R_{3n}, \\ \hat{\sigma}_n(k) - \sigma(k) &= (n-k)^{-1} \sum_{i=1}^{n-k} Z_{4i}(k) \\ &\quad - (n-k)^{-1} \sum_{i=1}^{n-k} \gamma_{ik}' \theta_{1n} + R_{4n}(k), \qquad 0 \leqslant k \leqslant l, \qquad (4.4) \\ \hat{\Sigma}_n - \Sigma_n &= \sum_{k=0}^{l} (n-k)^{-1} L_{kn} \sum_{i=1}^{n-k} [Z_{4i}(k) - \gamma_{ik}' \theta_{1n}] + R_{5n} \\ &\equiv \hat{\Sigma}_{1n} + R_{5n} \qquad \text{(say)} \end{aligned}$$

for some intermediate remainder term $R_{4n}(\cdot)$, where by (4.2), (4.3), and (A.3), $P(|R_{3n}| > cv_{2n}(\log n)^{-1}) = o(v_n)$ and $P(\|R_{5n}\| > cv_3) = o(v_n)$. See Lahiri [15] for details.

Note that for any $t \in \mathbb{R}^p$ with $\|t\| = 1$, by Cauchy–Schwartz inequality, $t'\Sigma_n t \geqslant \sigma(0) - 2\sum_{k=1}^{n-1} (\sum_{i=1}^{n} (t'd_i)^2) |\sigma(k)| \geqslant \sigma(0) - 2\sum_{k=1}^{\infty} |\sigma(k)| > 0$, so that Σ_n is p.d. for all $n \geqslant 1$. Now, using (4.3), (4.4), and Taylor's expansion, after some lengthy algebra one gets

$$
\begin{aligned}
T_n = h(\Sigma_n)\Bigg[\, & \theta_{1n} + \tau^{-1}\theta_{1n}\left(n^{-1}\sum_{i=1}^{n} Z_{2i} - (n\tau)^{-1}\sum_{i=1}^{n} (d_i'\theta_{1n})\, E\psi''(\varepsilon_1)\right) \\
& + \tau^{-1}(A - A_n)\,\theta_{1n} + (2\tau^2)^{-1}\sum_{i=1}^{n} d_i(d_i'\theta_{1n})^2\, E\psi''(\varepsilon_1)\Bigg] \\
& + \sum_{|\beta|=1} (\hat{\Sigma}_{1n})^{\beta}\, D^{\beta}h(\Sigma_n)\,\theta_{1n} + R_{6n},
\end{aligned} \tag{4.5}
$$

where by (4.2), (4.3), (4.4), and Lemma 3.3, $P(\|R_{6n}\| > Cv_{2n}) = o(v_n)$. Note that the stochastic approximation to T_n can be expressed in the form $T_{1n} \equiv h(\Sigma_n)\,\theta_{1n} + \sum_{r=1}^{p} \theta_{1n}'\Lambda_{0rn}\theta_{1n}e_r + \sum_{r=1}^{p} \tilde{Z}_{2n}'\Lambda_{1rn}\theta_{1n}e_r + \sum_{|\beta|=1} (\hat{\mathbf{\Sigma}}_{1n})^{\beta}$ $\Lambda_{\beta n}\theta_{1n}$, where $\tilde{Z}_{2n} = ((\mathbf{A} - \mathbf{A}_n)' \colon n^{-1}\sum_{i=1}^{n} Z_{2i})'$ and Λ_{0rn}, Λ_{1rn}, $\Lambda_{\beta n}'$ are nonrandom matrices satisfying $\max\{n^{1/2}\|\Lambda_{0rn}\| + \|\Lambda_{1rn}\| + \|\Lambda_{\beta n}\| \colon 1 \leqslant r \leqslant p, |\beta| = 1\} = 0(1)$.

Step (II): *Edgeworth expansion for* T_{1n}. Let $\theta_{2n} = \sum_{i=a}^{n} d_i Z_{1i}$, $A_{2n} = ((\sum_{i=a}^{n} d_{ij}d_{ik}Z_{2i}))_{p\times p}$, $\hat{Z}_{2n} = ((\mathbf{A}_{2n}' - E\mathbf{A}_{2n}') : n^{-1}\sum_{i=a}^{n} Z_{2i})'$, $\hat{\Sigma}_{2n} = \sum_{k=0}^{l} (n-k)^{-1} L_{kn} \sum_{i=a}^{n-k} [Z_{4i}(k) - \theta_{2n}'\gamma_{ik}]$, and $T_{2n} = h(\Sigma_n)\,\theta_{1n} + \sum_{r=1}^{p} \theta_{2n}'\Lambda_{0rn}\theta_{2n}e_r + \sum_{r=1}^{p} \hat{Z}_{2n}'\Lambda_{1rn}\theta_{2n}e_r + \sum_{|\beta|=1} (\hat{\mathbf{\Sigma}}_{2n})^{\beta}\,\Lambda_{\beta n}\theta_{2n}$.

Then, by (4.2), (4.10), Proposition 2.1, and Lemma 3.1,

$$
P(\|T_{1n} - T_{2n}\| > Cv_{2n}) = o(v_n). \tag{4.6}
$$

Let $\hat{Q}_n(t) = f(t'T_{2n})$, $t \in \mathbb{R}^p$. By standard arguments, it is enough to show

$$
\max_{|\alpha| \leqslant p+1} \int_{\Gamma_n} |D^{\alpha}(\hat{Q}_n(t) - \hat{\Psi}_n(t))|\, dt = o(v_n), \tag{4.7}
$$

where $\Gamma_n = \{t \in \mathbb{R}^p : \|t\| < v_{3n}^{-1}\}$ and $\hat{\Psi}_n$ is defined by (4.10) below. First, consider the integral over $\|t\| \leqslant m_3$. By Taylor's expansion, for any $t \in \mathbb{R}^p$, one gets

$$
Q_n(t) = E(1 + it'\,\Delta_n)\exp(it'h(\Sigma_n)\,\theta_{1n}) + R_{11,n}(t),
$$

where $\Delta_n \equiv T_{2n} - h(\Sigma_n)\,\theta_{1n}$. By Lemmas 3.1–3.3 (cf. Lahiri [15]),

$$\begin{aligned} E\,\|\Delta_n\|^{2r} &\leqslant C(r)[n^{-r}l^r(1+n^{-1}l^{r+1}) + M_n^{4r}n^r m_3^{2r} \\ |D^\alpha R_{11,n}(t)| &\leqslant C(1+\|t\|^2)[E\,\|\Delta_n\|^4)^{1/2} \\ &\qquad \times (1+E\,\|\theta_{1n}\|^{2\,|\alpha|})^{1/2} + E\,\|\Delta_n\|^{2+|\alpha|}] \\ &\leqslant C(1+\|t\|^2)\cdot n^{-1/2}\cdot n^{-\delta} \end{aligned} \tag{4.8}$$

for all $|\alpha| \leqslant p+1$ and some $\delta > 0$.

Next write $t_n = t'h(\Sigma_n)$, and $\hat{Z}_{2n,i}$ for the ith summand in $\tilde{Z}_{2n}$, $1 \leqslant i \leqslant n$ (so that $\hat{Z}_{2n} \equiv \sum_{i=a}^{n} \hat{Z}_{2n,i}$). Let $\bar{L}_{1n}(t) = \sum_{|\beta|=1} \sum_{k=0}^{l} (\mathbf{L}_{kn})^\beta (n-k)^{-1} \sum_{i=1}^{n-k} \gamma_{ik} t' \Lambda_{\beta n}$. For $1 \leqslant i,\ j \leqslant n,\ r \geqslant 1$, and $I \subseteq \{1, ..., n\}$, define $v_{1ij}(t) = \sum_{r=1}^{p} (t'e_r)\, d_i' \Lambda_{0rn} d_j - d_i' \bar{L}_{1n}(t)\, d_j$ and $\theta_{3n}(r, I) = \sum_{r,I}^{*} d_k Z_{1k}$, where $\sum_{r,I}^{*}$ extends over all indices $k \in \{1, ..., n\}$ such that $3(r-1)\, m_3 \leqslant |i-k| < 3rm_3$ for every $i \in I$. Now, expanding $f(t_n' \theta_3(1, I))$ with suitable choices of I and using Lemma 3.5, one can show (see Lahiri [15]) that for any $|\alpha| \leqslant p+1$ and all $\|t\| \leqslant m_3$,

$$|D^\alpha[E(1+\sqrt{-1}\; t'\Delta_n)\; f(t_n'\theta_{1n}) - \hat{\Psi}_n(t)]| \leqslant C(\alpha)\, n^{-1/2}\cdot n^{-\delta} \tag{4.9}$$

for some $\delta > 0$, where

$$\begin{aligned} &\exp(\|t\|^2/2)\;\hat{\Psi}_n(t) \\ &\quad = 1 + E(\sqrt{-1}\; t_n'\theta_{1n})^3/3! \\ &\qquad + \sqrt{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left[v_{1ij}(t)\, Z_{1i} Z_{ij} + \hat{Z}_{2n,i}' \left(\sum_{r=1}^{p} (t'e_r)\, \Lambda_{1rn} d_j \right) Z_{1j} \right] \\ &\qquad \times [1 - t_n'\theta_{3n}(1, \{i\})\, \theta_{3n}(1, \{j\})'\, t_n] \\ &\qquad - \sqrt{-1} \sum_{j=1}^{n} \sum_{k=0}^{l} \sum_{i=1}^{n} \left(\sum_{|\beta|=1} (L_{kn})^\beta (n-k)^{-1}\, t' \Lambda_{\beta n} d_j \right) \\ &\qquad \times E\{Z_{4i}(k)\, Z_{1j}(1 - t_n'\theta_{3n}(1, \{i, k\})\, t_n'\theta_{3n}(1, \{j\}))\}. \end{aligned} \tag{4.10}$$

Using (4.8) and Taylor's expansion for $m_3 < \|t\| \leqslant m_3(n/a)^{1/2}$ and using Lemma 3.5 and weak dependence of Δ_n and $\{Z_{1k} \colon 1 \leqslant k \leqslant a\}$ for $m_3(n/a)^{1/2} \leqslant \|t\| < v_{3n}^{-1}$,

$$\int_{\Gamma_{1n}} |D^\alpha \hat{Q}_n(t)|\; dt = o(v_n) \tag{4.11}$$

for all $|\alpha| \leqslant p+1$, where $\Gamma_{1n} = \{t \in \Gamma_n \colon \|t\| > m_3\}$. Combining (4.5)–(4.7), (4.9), and (4.11), one can now complete the proof of Theorem 2.1.

5. Proofs of Proposition 2.2 and Theorem 2.2

Proof of Proposition 2.2. Using Corollary 4.2 of Fuk and Nagaev [9] and Lemmas 3.6–3.7, one can show that

$$\begin{aligned} P_n\left(\left\|\sum_{1k} Z^*_{1k}\right\| > C(\log n)^{1/2}\right) &= O_p(v_{3n}), \\ P_n\left(\left\|\sum_{1k} (Z^*_{2k} - E_n Z^*_{2k})\right\| > C v_n l^{1/2} \log n\right) &= O_p(v_{3n}). \end{aligned} \tag{5.1}$$

Also, by Chebychev's inequality,

$$\begin{aligned} P_n\Bigg(\Bigg|\sum_{1k} \sum_{j=1}^{l} \left(\prod_{r=1}^{3} d_{(k-1)\,l+j,\, i_r}\right)(\psi''(\xi^*_{kj}) - E_n \psi''(\xi^*_{kj}))\Bigg| \\ > C v_n (\log n)^{-4} \text{ for some } 1 \leqslant i_1, i_2, i_3 \leqslant p\Bigg) = O_p(v_{3n}). \end{aligned} \tag{5.2}$$

The rest of the proof is similar to the proof of Proposition 2.1. ∎

Proof of Theorem 2.2. Let $\theta^*_n = D_n^{-1}(\beta^*_n - \bar{\beta}_n)$, $\theta^*_{1n} = \sum_{1k} Z^*_{1k}$, $A^*_n = \sum_{1k} Z^*_{2k}$, $\hat{A}_n = E_n A^*_n$, $\tau^*_{1n} = n_1^{-1} \sum_{i=1}^{n_1} \psi'(\varepsilon^*_i)$, and $\hat{\tau}_{1n} = E_n \tau^*_{1n}$. Also, let $\hat{\Sigma}^*_{1n} = \sum_{j=0}^{l_2} L^*_{jn}[\hat{\sigma}^*_{1n}(j) - E_n \hat{\sigma}^*_{1n}(j) - \gamma^{*\prime}_{jn} \theta^*_{1n}]$. By (5.1), (5.2), Proposition 2.2, and Corollary 4.2 of Fuk and Nagaev [9], one has (cf. (5.3)–(5.14) of [15])

$$T^*_n = T^*_{1n} + R^*_{8n}, \tag{5.3}$$

where $P_n(\|R^*_{8n}\| > C v_{2n}) = O_p(v_{3n})$ and the stochastic approximation T^*_{1n} is of the form

$$\begin{aligned} T^*_{1n} = h(\Sigma^*_n)\, \theta^*_{1n} &+ \sum_{r=1}^{p} \theta^{*\prime}_{1n} \hat{\Lambda}_{0rn} \theta^*_{1n} e_r \\ &+ \sum_{r=1}^{p} \tilde{Z}^{*\prime}_{2n} \hat{\Lambda}_{1rn} \theta^*_{1n} e_r + \sum_{|\beta|=1} (\Sigma^*_{1n})^{\beta}\, \hat{\Lambda}_{\beta n} \theta^*_{1n} \end{aligned} \tag{5.4}$$

such that $\tilde{Z}^*_{2n} = ((\hat{\mathbf{A}}_n - \mathbf{A}^*_n)' : (\tau^*_{1n} - \hat{\tau}_{1n}))'$ is of dimension $(p(p+1)/2) + 1$, and $\hat{\Lambda}_{0rn}$, $\hat{\Lambda}_{1rn}$, and $\hat{\Lambda}_{\beta n}$ are random matrices such that by Lemmas 3.1, 3.3, and 3.7,

$$\begin{aligned} \max\{\sqrt{n}\, \|\hat{\Lambda}_{0rn} - \Lambda_{0rn}\| + \|\hat{\Lambda}_{1rn} - \Lambda_{1rn}\| \\ + \|\hat{\Lambda}_{\beta n} - \Lambda_{\beta n}\| : 1 \leqslant r \leqslant p, |\beta| = 1\} = o_p(1). \end{aligned} \tag{5.5}$$

Let $a_1 = [l^{-1} n^{1/2-\gamma}]$ with $\gamma = \max\{p-1, 1\}\, \delta_0/4$, $\tau^*_{2n,k} = \sum_{j=1}^{l} \psi'(\xi^*_{kj})$, $\hat{Z}^*_{2n,k} = ((\mathbf{Z}^*_{2k} - E_n \mathbf{Z}^*_{2k})' \colon (\tau^*_{2n,k} - E_n \tau^*_{2n,k}))'$, and

$$V^*_{jk}(t) = \sum_{r=1}^{p} (t' e_r)[Z^*_{1j} \hat{\Lambda}_{0rn} Z^*_{1k} + \hat{Z}^*_{2n,j} \hat{\Lambda}_{1rn} Z^*_{1k}] - Z^*_{1j} \bar{L}^*_{1n}(t)\, Z^*_{1k}$$

$$+ k_0^{-1} \sum_{|\beta|=1} \left[\sum_{j=0}^{l_2} (\mathbf{L}^*_{jn})^{\beta} (Z_{4k}(j) - E_n Z_{4k}(j)) \right] (t' \hat{\Lambda}_{\beta n})\, Z^*_{1k}$$

for $1 \leqslant j, k \leqslant k_0$, where $\bar{L}^*_{1n}(t) = \sum_{j=0}^{l_2} \sum_{|\beta|=1} (\mathbf{L}^*_{jn})^{\beta}\, \gamma^*_{jn}(t' \Lambda_{\beta n})$, and $\mathbf{L}^*_{jn}$ is the $p(p+1)/2$-dimensional vector corresponding to L^*_{jn} of Lemma 3.7. Then, the Fourier transform Ψ^*_n of the Edgeworth expansion for T^*_n is defined by $\Psi^*_n(t) \exp(\|t\|^2/2) = 1 + \sum_{1k} E_n(\sqrt{-1}\, t^{*\prime}_n Z^*_{1k})^3/3! - \sqrt{-1} \sum_{1j} \sum_{1k} E_n V^*_{jk}(t) \cdot (t^{*\prime}_n Z^*_{1j}) \cdot (t^{*\prime}_n Z^*_{1j}) + \sqrt{-1} \sum_{1k} E_n V^*_{kk}(t)$. Note that T^*_{1n} is defined in terms of *independent* variables ξ^*_k, $1 \leqslant k \leqslant k_o$. Define T^*_{2n} by deleting all variables based on ξ^*_k, $1 \leqslant k \leqslant a_1$, from all but the first term of T^*_{1n} (cf. the definition of T_{1n} and T_{2n} above). Now, adapting the techniques of Götze [10] and using the results of Bhattacharya and Ranga Rao [4], one can show that for all $|\alpha| \leqslant p+1$,

$$\int_{\Gamma_{2n}} |D^{\alpha}(Q^*_n(t) - \Psi^*_n(t))|\, dt = O_p(v_{3n}), \tag{5.7}$$

where $\Gamma_{2n} = \{t \colon \|t\|^2 < n l^{-1} (\log n)^{-10}\}$ and $Q^*_n(t) = E_n f(t' T^*_{2n})$. For details of the arguments involved, see (5.15) through (5.27) of [15]. Next, using Lemma 3.1 and a discretizing argument as in the proof of Lemma 4.2 of Babu and Singh [1], one can show that

$$\operatorname{Sup} \left\{ \left| E_n f(t' Z^*_{1k}) - Ef\left(t' \sum_{j=1}^{l} d_{(k-1)l+j} \psi(\varepsilon_i) \right) \right| \colon \right.$$

$$\left. 1 \leqslant k \leqslant a_1,\ \|t\| \leqslant n^{5/8} \right\} = O_p(n^{-1/8}). \tag{5.8}$$

Furthermore, from the proof of Lemma 3.2 of [GH], it follows that for any random vector X and any sub-σ-field $\mathscr{C} \subseteq \mathscr{F}$, it there exist $\eta_1 \in (0, 1)$ and $\eta_2 > 0$ such that $|E(f(t'X) \mid \mathscr{C})| \leqslant 1 - \eta_1$ for $\eta_2 \leqslant \|t\| \leqslant 2\eta_2$, then for all $\|t\| < \eta_2$, $|(f(t'X) \mid \mathscr{C})| \leqslant \exp(-\eta_1 \|t\|^2/(8\eta_2^2))$. Hence, by (A.3) (iv), (A.4), and (5.8),

$$\prod \{ |E_n f(t^{*\prime}_n Z^*_{1k})| : k \in J,\ |J| \geqslant a_1 - p - 1,\ J \subseteq \{1, ..., a_1\} \}$$

$$\leqslant \exp(-C(\rho, \lambda)(1 - o_p(1))\, a_1\, \|t\|^2/n)$$

$$\leqslant \exp(-C(\rho, \lambda)(1 - o_p(1)(\log n)^2)$$

for all $n^{1/2}l^{-1/2}(\log n)^{-5} \leqslant \|t\| \leqslant C(\rho, \lambda)\, n^{1/2}$. So, by (5.8) and (A.3),

$$\max_{|\alpha| \leqslant p+1} \int_{\Gamma_{3n}} |D^{\alpha}\hat{Q}_n^*(t)| = O_p(v_{3n}),$$

where $\Gamma_{3n} = \Gamma_n \backslash \Gamma_{2n}$. This proves the validity of the Edgeworth expansion for T_n^*. Next, using Lemma 3.7 and similar arguments, and comparing (4.10) and Ψ_n^*, one gets (2.8). See Lahiri [15] for details.

References

[1] Babu, G. J., and Singh, K. (1984). On one term Edgeworth correction by Efron's bootstrap. *Sankhya Ser. A* **46** 219–232.

[2] Bhattacharya, R. N. (1985). Some recent results on Cramér–Edgeworth expansions with applications. *Multivariate Analysis IV* (P. R. Krishnaiah, Ed.), pp. 57–75. Elsevier Science, New York/Amsterdam.

[3] Bhattacharya, R. N., and Ghosh, J. K. (1978). On the validity of formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

[4] Bhattacharya, R. N., and Ranga Rao, R. (1986). *Normal Approximations and Asymptotic Expansions.* Krieger, Malabar, FL.

[5] Bickel, P. J., and Freedman, D. (1983). Bootstrapping regression models with many parameters. In *A Festschrift for E. L. Lehmann* (P. J. Bickel, K. A. Docksum, and J. L. Hodges, JR., Eds.), pp. 28–48. Wadsworth, Belmont, CA.

[6] Bickel, P. J., Götze, F., and Van Zwet, W. R. (1986). The Edgeworth expansion for U-statistics of degree two. *Ann. Statist.* **14** 1463–1484.

[7] Efron, B. (1979). Bootstrap methods: Another look at the Jackknife. *Ann. Statist.* **7** 1–26.

[8] Freedman, D. (1981). Bootstrapping regression models. *Ann. Statist.* **9** 1218–1228.

[9] Fuk, D. K., and Nagaev, S. V. (1971). Probability inequalities for sums of independent random variables. *Theory Probab. Appl.* **16** 643–660.

[10] Götze, F. (1987). Approximations for multivariate U-Statistics. *J. Multivariate Anal.* **22** 212–229.

[11] Götze, F., and Hipp, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **64** 211–239.

[12] Götze, F., and Künsch, H. R. (1990). Blockwise bootstrap for dependent observations: Higher order approximations for Studentized statistics (Abstract). *Inst. Math. Statist. Bull.* **19** 443.

[13] Huber, P. J. (1981). *Robust Statistics.* Wiley, New York.

[14] Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241.

[15] Lahiri, S. N. (1993). On Edgeworth expansion and moving block bootstrap for Studentized M-estimators in multiple linear regression models. Preprint, Department of Statistics, Iowa State University.

[16] Lahiri, S. N. (1992). Bootstrapping M-estimators of a multiple linear regression parameter. *Ann. Statist.* **20** 1548–1570.

[17] Lahiri, S. N. (1994). On second-order correctness of Efron's bootstrap without Cramér-type conditions in linear regression models. *Math. Methods Statist.* **3** 130–148.

[18] Lahiri, S. N. (1991). Second-order optimality of stationary bootstrap. *Statist. Probab. Lett.* **11** 335–341.

[19] Lahiri, S. N. (1994). Two-term Edgeworth expansion and bootstrap approximation for multivariate Studentized M-estimators, *Sankhya Ser. A* **56** 201–226.

[20] Liu, R. Y., and Singh, K. (1992). Moving blocks Jackknife and bootstrap capture weak dependence. In *Exploring the limits of bootstrap* (R. LePage and L. Billard, Eds.), pp. 225–248. Wiley, New York.

[21] Qumsiyeh, M. B. (1990). Bootstrapping and empirical Edgeworth expansions in multiple linear regression models. Preprint, Department of Mathematics, Indiana University.

[22] Shorack, G. (1982). Bootstrapping robust regression. *Comm. Statist. A* **11** 961–972.

[23] Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

[24] Tiro, A. M. (1991). *Edgeworth Expansion and Bootstrap Approximation for M-Estimators of Linear Regression Parameters with Increasing Dimensions*. Unpublished Ph.D. dissertation, Department of Statistics, Iowa State University.