A Uniform-in-$P$ Edgeworth Expansion under Weak Cramér Conditions

Kyungchul Song
Vancouver School of Economics, University of British Columbia

Abstract. This paper provides a finite sample bound for the error term in the Edgeworth expansion for a sum of independent, potentially discrete, nonlattice random vectors, using a uniform-in-$P$ version of the weaker Cramér condition in Angst and Poly (2017). This finite sample bound is used to derive a bound for the error term in the Edgeworth expansion that is uniform over the joint distributions $P$ of the random vectors, and eventually to derive a higher order expansion of resampling-based distributions in a unifying way. As an application, we derive a uniform-in-$P$ Edgeworth expansion of bootstrap distributions and that of randomized subsampling distributions, when the joint distribution of the original sample is absolutely continuous with respect to Lebesgue measure.

Key words. Edgeworth Expansion; Normal Approximation; Bootstrap; Randomized Subsampling Distributions; Weak Cramér Condition

AMS 1991 Classification: 62G20, 62E20

1. Introduction

Suppose that $\{X_{i,n}\}_{i=1}^n$ is a triangular array of independent random vectors taking values in $\mathbb{R}^d$ and have mean zero. Let $\mathcal{P}_n$ be the collection of the joint distributions for $(X_{i,n})_{i=1}^n$. Define

$$V_{n,P} \equiv \frac{1}{n} \sum_{i=1}^n Var_P(X_{i,n}),$$

Date: June 6, 2018.

I acknowledge that this research was supported by Social Sciences and Humanities Research Council of Canada. Corresponding address: Kyungchul Song, Vancouver School of Economics, University of British Columbia, Vancouver, BC, Canada. Email address: kysong@mail.ubc.ca.
where $\text{Var}_P(X_{i,n}) \equiv \text{E}(X_{i,n}X'_{i,n})$. We are interested in the Edgeworth expansion of the distribution of

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} V_{n,P}^{-1/2} X_{i,n},
$$

which is uniform over $P \in \mathcal{P}_n$.

The Edgeworth expansion has long received attention in the literature. See Bhattacharya and Rao (2010) for a formal review of the results. The validity of the classical Edgeworth expansion is obtained under the Cramér condition which says that the average of the characteristic functions over the sample units stays bounded below 1 in absolute value as the function is evaluated at a sequence of points which increase to infinity. The Cramér condition fails when the random variable has a support that consists of a finite number of points, and hence does not apply to resampling-based distributions such as bootstrap.\(^1\) A standard approach to deal with this issue is to derive an Edgeworth expansion separately for the bootstrap distribution by using the expansion of the empirical characteristic functions. (e.g. Singh (1981) and Hall (1992).)

The main contribution of this paper is to provide a finite sample bound for the remainder term in the Edgeworth expansion for a sum of independent random vectors. Using the finite sample bound, one can immediately obtain a uniform-in-$P$ Edgeworth expansion, where the error bound for the remainder term in the expansion is uniform over a collection of probabilities. A notable feature of the Edgeworth expansion is that it admits random vectors which can be discrete, non-lattice distributions. From this result, as shown in the paper, a uniform-in-$P$ Edgeworth expansion for various resampling-based discrete distributions follows as a corollary. To obtain such an expansion, this paper uses a uniform-in-$P$ version of weak Cramér conditions introduced by Angst and Poly (2017) and obtains a finite sample bound for the error term in the Edgeworth expansion by following the proofs of Theorems 20.1 and 20.6 of Bhattacharya and Rao (2010) and Theorem 4.3 of Angst and Poly (2017). This paper’s finite sample bound reveals that we obtain a uniform-in-$P$ Edgeworth expansion, whenever we have a uniform-in-$P$ bound for the moment of the same order in the Edgeworth expansion.

A uniform-in-$P$ asymptotic approximation is naturally required in a testing set-up with a composite null hypothesis. By definition, a composite null hypothesis involves a collection of probabilities, and the size of a test in this case is its maximal rejection probability over all the probabilities admitted in the null hypothesis. Asymptotic control of size requires

\(^1\)Despite the failure of the Cramér condition, the Edgeworth expansion for lattice distributions is well known. See Bhattacharya and Rao (2010), Chapter 5. See Kolassa and McCullagh (1990) for a general result of Edgeworth expansion for lattice distributions. Booth, Hall, and Wood (1994) provide the Edgeworth expansion for discrete yet non-lattice distributions.
uniform-in-$P$ asymptotic approximation of the test statistic’s distribution under the null hypothesis. One can apply the same notion to the coverage probability control of confidence intervals as well.

As for uniform-in-$P$ Gaussian approximation, one can obtain the result immediately from a Berry-Esseen bound once appropriate moments are bounded uniformly in $P$. It is worth noting that uniform-in-$P$ Gaussian approximation of empirical processes was studied by Giné and Zinn (1991) and Sheehy and Wellner (1992). There has been a growing interest in uniform-in-$P$ inference in various nonstandard set-ups in the literature of econometrics in connection with the finite sample stability of inference. (See Mikusheva (2007), Linton, Song, and Whang (2010), and Andrews and Shi (2013), among others.)

When a test based on a resampling procedure exhibits higher order asymptotic refinement properties, the uniform-in-$P$ Edgeworth expansion can be used to establish a higher order asymptotic size control for the test. A related work is found in Hall and Jing (1995) who used the uniform-in-$P$ Edgeworth expansion to study the asymptotic behavior of the confidence intervals based on a studentized t statistic. They used a certain smoothness condition for the distributions of the random vectors which excludes resampling-based distributions. A recent paper by the author (Song (2018)) uses this paper’s result to compare two different testing procedures based on randomized subsampling inference, when observations are locally dependent with unknown dependence ordering.

2. A Uniform-in-$P$ Edgeworth Expansion

2.1. Uniform-in-$P$ Weak Cramér Conditions

Angst and Poly (2017) (hereafter, AP) introduced what they called a weak Cramér condition and a mean weak Cramér condition which are weaker than the classical Cramér condition. They showed that through their weakening of the latter condition, we can obtain a classical Edgeworth expansion which accommodates the distribution of discrete random variables that arise in resampling methods in statistics. In this paper, we introduce their uniform-in-$P$ versions. Let us prepare notation. Let $\| \cdot \|$ be the Euclidean norm in $\mathbb{R}^d$, i.e., $\|a\|^2 = tr(a'a)$. The following definition modifies the weak Cramér condition introduced by AP into a condition for a collection of probabilities.\footnote{The original definition of the weak Cramér condition in AP specifies the bounds in (1) and (2) in the form $1 - c/\|t\|^b$. Since we can simply set $c = 1$ in applications, we use this choice throughout this paper.}

**Definition 2.1.** (i) Given $b, R > 0$, a collection of the distributions $\mathcal{P}$ of a random vector $W$ taking values in $\mathbb{R}^d$ and having characteristic function $\phi_{W,P}$ under $P \in \mathcal{P}$ is said to
satisfy the weak Cramér condition with parameter \((b, R)\), if for all \(t \in \mathbb{R}^d\) with \(\|t\| > R\),

\[
\sup_{P \in \mathcal{P}} |\phi_{W,P}(t)| \leq 1 - \frac{1}{\|t\|^b}.
\]

(ii) Given \(b, R > 0\), a collection of the joint distributions \(\mathcal{P}_n\) of a triangular array of random vectors \((W_{i,n})_{i=1}^n\) with each \(W_{i,n}\) taking values in \(\mathbb{R}^d\) and having characteristic function \(\phi_{W_{i,n},P}\) under \(P \in \mathcal{P}_n\) is said to satisfy the mean weak Cramér condition with parameter \((b, R)\), if for all \(t \in \mathbb{R}^d\) with \(\|t\| > R\),

\[
\sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^n |\phi_{W_{i,n},P}(t)| \leq 1 - \frac{1}{\|t\|^b}.
\]

As noted by AP, the weak Cramér condition is useful for dealing with distributions obtained from a resampling procedure. To clarify this in our context, let us introduce some notation. For any integers \(d, p \geq 1\), any given sequence of vectors \(u = (u_j)_{j=1}^p\) where \(u_j \in \mathbb{R}^d\), let

\[
\mathcal{M}_{d,p}(u) \equiv \left\{ \sum_{j=1}^p c_j \delta_{u_j} : (c_j)_{j=1}^p \in (0, \infty)^p, \sum_{j=1}^p c_j = 1 \right\},
\]

where \(\delta_{u_j}\) denotes the Dirac measure at \(u_j\). Let \(\mathcal{U}_{d,p}(b, R)\) be the collection of \(u\)’s such that \(\mathcal{M}_{d,p}(u)\) does not satisfy the weak Cramér condition with parameter \((b, R)\). The following proposition is due to AP. (See Proposition 2.4 there.)

**Proposition 2.1 (Angst and Poly (2017)).** Suppose that \(p \geq 3\) and \(b > 1/(p - 2)\). Then

\[
\lambda \left( \bigcup_{R > 0} \mathcal{U}_{d,p}(2b, R) \right) = 0,
\]

where \(\lambda\) is Lebesgue measure on \(\mathbb{R}^p\).

Therefore, the weak Cramér condition is generically satisfied by \(\mathcal{M}_{d,p}(u)\) for some \(R > 0\) for almost all \(u\)’s. Angst and Poly (2017) give the proof only for the case of \(d = 1\). The proof for the general case with \(d \geq 1\) is provided in the appendix of this paper.

Let us illustrate how this proposition can be used to establish uniform-in-\(P\) inference based on resampling. Let \(U_n \equiv \{U_{j,n}\}_{j=1}^n\) be a triangular array of random vectors with \(U_{j,n} \in \mathbb{R}^d\). Let the collection of the joint distributions for \(U_n\) be denoted by \(\mathcal{P}_n\) and assume that each \(P \in \mathcal{P}_n\) is dominated by Lebesgue measure on \(\mathbb{R}^{nd}\). Let us assume that \(X_{i,n}, i = 1, \ldots, n\), are i.i.d. draws from the empirical measure \(\frac{1}{n} \sum_{j=1}^n \delta_{U_{j,n}}\) of \(\{U_{i,n}\}_{i=1}^n\). Thus,

\[
\frac{1}{n} \sum_{j=1}^n \delta_{u_j} \in \mathcal{M}_{d,p}(u),
\]
with \( p = n \). Let \( \mathcal{U}_n \) be the \( \sigma \)-field generated by \( (U_{j,n})_{j=1}^{n} \) and the conditional distribution of

\[
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_{i,n} - \mathbb{E}[X_{i,n}|\mathcal{U}_n])
\]
given \( \mathcal{U}_n \) be denoted by \( Q_n(\cdot|\mathcal{U}_n) \), and let \( \tilde{Q}_n(\cdot|\mathcal{U}_n) \) be a signed conditional measure which we would like to show to be approximating \( Q_n(\cdot|\mathcal{U}_n) \). In particular, we are interested in showing that for any collection of convex subsets \( A(c) \) indexed by \( c \in \mathbb{R} \), and for some decreasing sequence \( \varepsilon_n \to 0 \),

(3) \[
\sup_{P \in \mathscr{P}_n} \mathbb{P} \left\{ \left| Q_n(A(c)|\mathcal{U}_n) - \tilde{Q}_n(A(c)|\mathcal{U}_n) \right| > \varepsilon_n \right\} \to 0,
\]
as \( n \to \infty \). For example, if we take \( A(c) = \{ x \in \mathbb{R}^d : \|x\|^2 \leq c \} \), \( Q_n(A(c)|\mathcal{U}_n) \) denotes the conditional CDF of the quadratic form of \( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_{i,n} - \mathbb{E}[X_{i,n}|\mathcal{U}_n]) \).

Now, let us see how Proposition 2.1 can be useful here. Suppose that for each outcome \( \omega \in \Omega \) such that \( U_n(\omega) = u \) for some \( u \) with \( \mathcal{P}_n(u) \) satisfying the weak Cramér condition with parameter \( (2b, R) \), we have

(4) \[
|Q_n(A(t)|\mathcal{U}_n)(\omega) - \tilde{Q}_n(A(t)|\mathcal{U}_n)(\omega)| \leq h_n(\mathcal{U}_n(\omega)),
\]
for some sequence of Borel measurable functions \( h_n \) for each \( n \geq 1 \), such that

(5) \[
\varepsilon_n^{-1} \sup_{P \in \mathcal{P}_n} \mathbb{E}_P[h_n(\mathcal{U}_n)] \to 0,
\]
as \( n \to \infty \), where \( \mathbb{E}_P \) denotes the expectation under \( P \). Now, for each \( b > 0 \), let \( E_n(b, R) \) be the event that \( \mathcal{P}_n(U_n) \) satisfies the weak Cramér condition with parameter \( (2b, R) \). We bound the supremum in (3) by

\[
\sup_{P \in \mathcal{P}_n} \mathbb{P} \left\{ \left| Q_n(A(t)|\mathcal{U}_n) - \tilde{Q}_n(A(t)|\mathcal{U}_n) \right| > \varepsilon_n \right\} \cap E_n(b, R)
\]
\[
+ \sup_{P \in \mathcal{P}_n} \mathbb{P} \left\{ \left| Q_n(A(t)|\mathcal{U}_n) - \tilde{Q}_n(A(t)|\mathcal{U}_n) \right| > \varepsilon_n \right\} \cap E_n^c(b, R).
\]

By Proposition 2.1, the second term is zero. Using Markov’s inequality, (4), and (5), the leading term vanishes to zero as \( n \to \infty \), establishing (3). Thus for the uniform-in-\( P \) approximation of \( Q_n(\cdot|\mathcal{U}_n) \) by \( \tilde{Q}_n(\cdot|\mathcal{U}_n) \), it is useful to obtain an explicit finite sample bound in (4). For such a result, a uniform-in \( P \) Edgeworth expansion is helpful.

### 2.2. A Uniform-in-\( P \) Edgeworth Expansion under Weak Cramér Conditions

In this section, we present the main result that gives a finite sample bound for the error term in the Edgeworth expansion. Let us prepare notation first. Given each multi-index \( \nu = (\nu_1, \ldots, \nu_d) \), with \( \nu_k \) being a nonnegative integer, we let \( \bar{X}_{\nu,P} \) be the average of the
\[ \nu - \text{th cumulant of } V_{n,P}^{-1/2} X_{i,n}. \] For each \( j = 1, 2, \ldots, \) let \( \tilde{P}_j(z : \{\tilde{\chi}_{\nu,P}\}) \) be a polynomial in \( z \) as given in (7.3) of Bhattacharya and Rao (2010) (BR, hereafter). This polynomial has degree \( 3j \), the smallest order of the terms in the polynomial is \( j + 2 \), and the coefficients in the polynomial involve only \( \tilde{\chi}_{\nu,P} \)'s with \( |\nu| \leq j + 2 \). (Lemma 7.2 of BR, p.52.) Following the convention, we define the derivative operators as follows:

\[
D_k \equiv \frac{\partial}{\partial x_k}, \quad k = 1, \ldots, d, \quad \text{and} \quad D^\alpha \equiv D_1^{\alpha_1} \cdots D_d^{\alpha_d},
\]

for \( \alpha = (\alpha_1, \ldots, \alpha_d) \in \{0, 1, 2, \ldots\}^d \) and \( D = (D_1, \ldots, D_d) \). For each \( P \in \mathcal{P}_n \), let \( Q_{n,P} \) be the distribution of \( \frac{1}{\sqrt{n}} \sum_{i=1}^n V_{n,P}^{-1/2} X_{i,n} \), and define a signed measure \( \tilde{Q}_{n,s,P} \) as follows: for any Borel set \( A \subset \mathbb{R}^d \),

\[
\tilde{Q}_{n,s,P}(A) \equiv \int_A \sum_{j=0}^{s-2} n^{-j/2} \tilde{P}_j(-D : \{\tilde{\chi}_{\nu}\}) \phi(x) dx,
\]

where \( \phi \) is the density of the standard normal distribution on \( \mathbb{R}^d \). For each \( s \geq 1 \), define

\[
\rho_{n,s,P} \equiv \frac{1}{n} \sum_{i=1}^n E_P \|X_{i,n}\|^s.
\]

Let \( \phi_{X_{i,n},P}(t) = E_P[\exp(it'X_{i,n})] \), i.e., the characteristic function of \( X_{i,n} \), and define for any \( 0 < r < R, \) and \( n_0 \geq 1 \),

\[
\bar{\phi}(r, R, n_0) \equiv \sup_{n \geq n_0} \sup_{r \leq ||x|| \leq R} \sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^n |\phi_{X_{i,n},P}(t)|.
\]

We introduce notation for modulus of continuity for functions: for any Borel measurable function \( f \) on \( \mathbb{R}^d \), and any measure \( \mu \), we define for \( \varepsilon > 0 \) and \( x \in \mathbb{R}^d \),

\[
\omega_f(x; \varepsilon) \equiv \sup \{ |f(z) - f(y)| : z, y \in B(x; \varepsilon) \}, \quad \text{and} \quad \bar{\omega}_f(\varepsilon; \mu) \equiv \int \omega_f(x; \varepsilon) \mu(dx),
\]

where \( \Phi \) denotes the distribution function of \( N(0, I_d) \) and \( B(x; \varepsilon) \) is the \( \varepsilon \)-open ball in \( \mathbb{R}^d \) around \( x \). For any measurable function \( f \) and for \( s > 0 \), we define

\[
M_s(f) \equiv \sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{1 + ||x||^s}.
\]

Define for any constant \( a > 0 \) and integers \( s \geq 1 \),

\[
\bar{\Delta}_{n,s,P}(a) \equiv \frac{1}{n} \sum_{i=1}^n E_P \left[ ||X_{i,n}||^s I \{||X_{i,n}|| \geq a\sqrt{n}\} \right].
\]

The theorem below is the main result of this paper which is a modification of Theorem 4.3 of AP with the bound made explicit in finite samples.
Theorem 2.1. Suppose that for each \( P \in \mathcal{P}_n \), \( V_{n,P} \) is positive definite, and that there exist \( n_1 \geq 1 \) and \( b > 0 \) such that the following two conditions hold.

(i) There exists a number \( \bar{\rho} > 0 \) such that

\[
\sup_{n \geq n_1} \sup_{P \in \mathcal{P}_n} \rho_{s,n,P} \leq \bar{\rho} < \infty,
\]

for some \( s \geq 3 \).

(ii) \( \mathcal{P}_n \) satisfies the mean weak Cramér condition with parameter \( (b, R) \) for all \( n \geq n_1 \), with \( b \) and \( R \) satisfying that for \( s \) and \( \bar{\rho} \) in (i),

\[
\frac{s - 3}{2} < \frac{1}{b}, \quad \text{and} \quad R > \frac{1}{16\bar{\rho}}.
\]

Then, for any Borel measurable function \( f \) on \( \mathbb{R}^d \) such that \( M_{s'}(f) < \infty \) for some \( 0 \leq s' \leq s \), and for all \( \delta > 0 \) such that

\[
\frac{s - 2}{2} < \delta < \frac{1}{b} + \frac{1}{2},
\]

there exist constants \( n_0 \geq 1 \) and \( C > 0 \) such that for all \( n \geq n_0, P \in \mathcal{P}_n, \xi \in (0,1] \), and we have

\[
\left| \int f d(Q_{n,P} - \tilde{Q}_{n,s,P}) \right| \leq Cn^{-(s-2)/2} (1 + M_{s'}(f)) \{ \xi + \bar{\Delta}_{n,s,P}(\xi) \}
\]

\[
\quad + C\tilde{\omega}_f(n^{-\delta}; \Phi),
\]

where \( \Phi \) denotes the distribution of \( N(0, I_d) \), and \( n_0 \) and \( C \) depend only on \( s', s, d, \bar{\rho}, b, R, n_1, \delta, \) and \( \bar{\phi}(1/(16\bar{\rho}), R, n_1) \).

The bound in the above depends on \( P \in \mathcal{P}_n \) only through \( \bar{\Delta}_{n,s,P}(\xi) \). By choosing a sequence \( \xi_n \to 0 \) with \( \sqrt{n}\xi_n \to \infty \), and replacing \( \bar{\Delta}_{n,s,P}(\xi_n) \) by \( \sup_{P \in \mathcal{P}_n} \bar{\Delta}_{n,s,P}(\xi_n) \), we obtain a bound for the error term in the Edgeworth expansion that is \( o(n^{-(s-2)/2}) \) uniformly in \( P \in \mathcal{P}_n \), as \( n \to \infty \). Thus it is revealed that the uniform-in-\( P \) Edgeworth expansion of \( Q_{n,P} \) is essentially obtained by strengthening the mean weak Cramér condition to the same condition but with uniformity in \( P \) and by strengthening the moment condition to uniformity in \( P \) as in (6).

When we take \( f \) to be the indicator function of convex subsets of \( \mathbb{R}^d \), we obtain the following corollary which is a version of Corollary 20.15 of BR with the finite sample bound made explicit here.

Corollary 2.1. Suppose that the conditions of Theorem 2.1 hold, and let \( \mathcal{C} \) be the collection of convex subsets of \( \mathbb{R}^d \). Then, there exist constants \( n_0 \geq 1 \) and \( C > 0 \) such that for all \( n \geq n_0 \),
\[ \xi \in (0, 1], P \in \mathcal{P}_n, \]
\[ \sup_{A \in \mathcal{E}} \left| Q_{n,P}(A) - \tilde{Q}_{n,s,P}(A) \right| \leq Cn^{-(s-2)/2} \{ \xi + \Delta_{n,s,P}(\xi) \} + Cn^{-\delta}, \]
where \( n_0 \) and \( C \) depend only on \( s, d, \bar{\rho}, b, R, n_1, \delta \), and \( \bar{\phi}(1/(16\bar{\rho}), R, n_1) \).

The last term \( Cn^{-\delta} \) follows because for any indicator \( f(x) = 1_A(x) \) on a convex set \( A \), we have \( \bar{\omega}_f(\varepsilon; \Phi) \leq C \varepsilon \), where \( C \) is a constant that depends only on \( s \) and \( d \). (See Corollary 3.2 of BR, p.24.)

The following result for the case of indicator functions on sets defined by polynomials is useful for establishing an Edgeworth expansion of a studentized sample mean. Later we use this result to establish a uniform-in-\( P \) Edgeworth expansion for the bootstrap distribution of the studentized sample mean. Let us define a set as follows:

\[ A_{n,\nu,\delta}(t, a, c_\nu) \equiv \{ x \in \mathbb{R}^d : \|x\| \leq a \log n, \text{ and } |r_{n,\nu}(x; c_\nu) - t| \leq n^{-\delta} \}, \]
where for \( x \in \mathbb{R}^d \),

\[ r_{n,\nu}(x; c_\nu) = \sum_{k_1} c_{k_1} x_{k_1} + n^{-1/2} \sum_{k_1} \sum_{k_2} c_{k_1,k_2} x_{k_1} x_{k_2} + \cdots 
+ n^{-\nu/2} \sum_{k_1} \sum_{k_2} \cdots \sum_{k_s} c_{k_1,k_2,\ldots,k_\nu} x_{k_1} x_{k_2} \cdots x_{k_\nu}, \]
with \( c_\nu = \{ c_{k_1}, c_{k_1,k_2}, \ldots, c_{k_1,k_2,\ldots,k_\nu} \in \mathbb{R} : k_1, k_2, \ldots, k_\nu \in \{1, \ldots, d\} \} \). Let

\( \mathcal{A}_{n,\nu,\delta}(a, \bar{c}) \equiv \{ A_{n,\nu,\delta}(t, a, c_\nu) : \bar{c}_\nu \leq \bar{c}, c_k > 1/\bar{c}, \text{ for some } k = 1, \ldots, d, \text{ and } t \in \mathbb{R} \}, \)
where \( \bar{c}_\nu = \max_{k_1,k_2,\ldots,k_\nu \in \{1,\ldots,d\}} \max \{ |c_{k_1}|, |c_{k_1,k_2}|, \ldots, |c_{k_1,k_2,\ldots,k_\nu}| \} \).
Then, we obtain the following result as a corollary from Theorem 2.1.

**Corollary 2.2.** Suppose that the conditions of Theorem 2.1 hold. Then, for any \( R > 1/(16\bar{\rho}) \) and \( a, \bar{c} > 0 \), there exist constants \( n_0 \geq 1 \) and \( C > 0 \) such that for all \( n \geq n_0, P \in \mathcal{P}_n, \xi \in (0, 1], \)

\[ \sup_{A \in \mathcal{A}_{n,\nu}(a, \bar{c})} \left| Q_{n,P}(A) - \tilde{Q}_{n,s,P}(A) \right| \leq Cn^{-(s-2)/2} \{ \xi + \Delta_{n,s,P}(\xi) \} + Cn^{-\delta}, \]
where \( C \) and \( n_0 \) depend only on \( s, d, \bar{\rho}, b, R, n_1, \delta, \bar{\phi}(1/(16\bar{\rho}), R, n_1), a, \) and \( \bar{c} \).

The result immediately follows from Theorem 2.1 above, after we apply Lemma 5.3 of Hall (1992), p.254, to bound \( \bar{\omega}_f(n^{-\delta}; \Phi) \) by \( Cn^{-\delta} \) with \( f(x) = 1 \{ x \in A_{n,\nu,\delta}(t, a, c_\nu) \} \).
3. Applications

3.1. Nonparametric Bootstrap Distributions

Let us illustrate how the previous results can be applied to obtain a uniform-in-$P$ Edgeworth expansion of a bootstrap distribution of a sum of independent random variables when the random variables are continuous. The Edgeworth expansion of a bootstrap distribution for the i.i.d. random variables is well known in the literature. (Hall (1992)). The result in this paper is distinct for two reasons. First, the Edgeworth expansion is uniform in $P$, where $P$ runs over the distribution of the random variable. Second, the Edgeworth expansion follows directly from the Edgeworth expansion for a sum of i.i.d. random variables, due to the use of the weak Cramér condition. On the other hand, this paper’s result assumes that the random variables are continuous, whereas the standard bootstrap result requires only the classical Cramér condition for the random variables. This is due to our reliance on Proposition 2.1.

Suppose that $\{W_{i,n}\}_{i=1}^n$ is a triangular array of continuous random variables which are i.i.d. drawn from a common distribution. Let us assume that this distribution belongs to the collection $\mathcal{P}_n$ of distributions. Let $\{W_{i,n,b}\}_{i=1}^n, b = 1, ..., B$, be the bootstrap sample drawn with replacement from the empirical distribution of $\{W_{i,n}\}_{i=1}^n$. Define the sample variance:

$$s_{n,b}^2 = \frac{1}{n} \sum_{i=1}^n (W_{i,n,b} - \bar{W}_{n,b})^2,$$

where $\bar{W}_{n,b} = \frac{1}{n} \sum_{i=1}^n W_{i,n,b}$. Then, we are interested in the uniform-in-$P$ Edgeworth expansion of the bootstrap distribution of the following:

$$T_{n,b}^\ast = \frac{1}{s_{n,b} \sqrt{n}} \sum_{i=1}^n (W_{i,n,b}^\ast - \bar{W}_n),$$

where $\bar{W}_n = \frac{1}{n} \sum_{i=1}^n W_{i,n}$. For this, we adopt the approach in Chapter 5 of Hall (1992), and in deriving the finite sample bound, we apply Corollary 2.2. Let $\bar{X}_{n,b}^\ast = (W_{n,b}^\ast, \frac{1}{n} \sum_{i=1}^n W_{i,n,b}^\ast, 2)$.

It is not hard to see that we can write

$$T_{n,b}^\ast = g_n(\bar{X}_{n,b}^\ast),$$

where $g_n(x) = \sqrt{n}(x_1 - \bar{W}_n)/\sqrt{x_2 - x_1^2}$, $x = (x_1, x_2)$, with $x_2 > x_1^2$.

More generally, suppose that we have a triangular array of random vectors $X_{i,n}$ taking values in $\mathbb{R}^d$, and the bootstrap sample $\{X_{i,n,b}\}_{i=1}^n, b = 1, ..., B$, and let the $\sigma$-field generated by $\{X_{i,n}\}_{i=1}^n$ be $\mathcal{F}_n$. Our focus is on the Edgeworth expansion of the bootstrap distribution of the test statistic of the form in (10) for a generic function $g_n(x)$ which is
s + 3 times continuously differentiable at \( x = \bar{X}_n \), with \( \bar{X}_n = (\bar{W}_n, \frac{1}{n} \sum_{i=1}^{n} W_{i,n}^2) \), and is \( \mathcal{F}_n \)-measurable.

The bootstrap distribution of \( T_{n,b}^s \) (defined in (10)) is the conditional distribution of \( T_{n,b}^s \) given \( \mathcal{F}_n \) which we denote by \( Q_n(\cdot | \mathcal{F}_n) \). Let \( \chi_{\nu}^s \) be the \( \nu \)-th cumulant of the conditional distribution of \( X_{i,n}^s \) given \( \mathcal{F}_n \). Define for each Borel \( A \subset \mathbb{R}^d \),

\[
\tilde{Q}_n(A|\mathcal{F}_n) = \int_A \sum_{j=0}^{s-2} n^{-j/2} \tilde{P}_j(-D : \{\chi_{\nu}^s\}) \phi(x) dx.
\]

Our purpose is to obtain a finite sample bound for the error term in the approximation of the bootstrap measure \( Q_n(\cdot | \mathcal{F}_n) \) by an Edgeworth expansion. We define the following three events: for constants \( s, c_1, c_2, c_3 > 0 \), \( Z_+ \equiv \{0, 1, 2, \ldots\} \),

\[
\mathcal{E}_{n,1}(s, c_1) \equiv \left\{ \max_{\ell_1, \ldots, \ell_d \in \mathbb{Z}_n, |\ell_1| + \cdots + |\ell_d| \leq s - 2} \left| \frac{1}{n} \sum_{i=1}^{n} X_{ik,n}^\ell \right| \leq c_1 \right\},
\]

\[
\mathcal{E}_{n,2}(s, c_2) \equiv \left\{ \max_{|\alpha| \leq s + 3} |D^\alpha g_n(\bar{X}_n)| \leq c_2 \right\},
\]

and \( \mathcal{E}_{n,3}(c_3) \) be the event of all the eigenvalues of \( \tilde{V}_n \) lying in \((c_3^{-1}, c_3)\), where

\[
\tilde{V}_n = \frac{1}{n} \sum_{i=1}^{n} (X_{i,n} - \bar{X}_n)(X_{i,n} - \bar{X}_n)'.
\]

Let \( \mathcal{E}_{n,4}(\tilde{\rho}) \) be the event where \( \frac{1}{n} \sum_{i=1}^{n} \|X_{i,n}\|^* \leq \tilde{\rho} \). Define

\[
\mathcal{E}_n(s, c_1, c_2, c_3, \tilde{\rho}) \equiv \mathcal{E}_{n,1}(s, c_1) \cap \mathcal{E}_{n,2}(s, c_2) \cap \mathcal{E}_{n,3}(c_3) \cap \mathcal{E}_{n,4}(\tilde{\rho}).
\]

**Theorem 3.1.** Suppose that \( X_{i,n}'s \) are i.i.d., and each \( X_{i,n} \) is absolutely continuous with respect to Lebesgue measure. Let \( s \geq 3 \) be a given integer and \( \tilde{\rho} > 0 \) a constant. Then, there exist constants \( n_0 \geq 1, (s - 2)/2 < \delta < (s - 1)/2, \) and \( C > 0 \) such that for all \( n \geq n_0, \xi \in (0, 1), \) all \( \delta \) that satisfies (8) and for all \( c_1, c_2, c_3 > 0, \) on the event \( \mathcal{E}_n(s, c_1, c_2, c_3, \tilde{\rho}), \)

\[
\sup_{t \in \mathbb{R}} |Q_n((-\infty, t] | \mathcal{F}_n) - \tilde{Q}_n((-\infty, t] | \mathcal{F}_n)| \leq C n^{-(s-2)/2} \left\{ \xi + \hat{\Delta}_{n,s}(\xi) \right\} + C n^{-\delta}, \text{ a.e.,}
\]

where \( C \) and \( n_0 \) depend only on \( d, s, \tilde{\rho}, \delta \) and \( c_1, c_2, c_3 > 0, \) and

\[
\hat{\Delta}_{n,s}(\xi) \equiv \frac{1}{n} \sum_{i=1}^{n} \|X_{i,n}\|^* 1 \left\{ \|X_{i,n}\| \geq \xi \sqrt{n} \right\}.
\]

As for the probability \( P(\mathcal{E}_n^c(s, c_1, c_2, c_3)) \), we can find its finite sample bound using the standard arguments, and show it to be \( O(n^{-(s-2)/2}) \) uniformly over \( P \in \mathcal{P}_n \), when there is a uniform bound for the population moments, and a uniform lower and upper bound for
the eigenvalues of $V_n \equiv E[(X_{i,n} - E X_{i,n})(X_{i,n} - E X_{i,n})^\prime]$. Details can be furnished following the same arguments in Chapter 5 of Hall (1992).

3.2. Randomized Subsampling Distributions

Let $X_n \equiv \{X_{i,n}\}_{i=1}^n$ be a triangular array of continuous random vectors taking values in $\mathbb{R}^d$, having joint distribution $P \in \mathcal{P}_n$. Let us denote $Z_n$ to be the $\sigma$-field generated by $X_n$. Let $\tau_n(\cdot) : \mathbb{R}^{db_n} \to \mathbb{R}^d_1$ be a $Z_n$-measurable map, and $\Pi$ be the collection of the permutations on $[n] \equiv \{1, 2, ..., n\}$. Then, a test statistic built from randomized subsampling is based on the following form of the sum of conditionally i.i.d. random vectors: for each random vector of permutations $\tilde{\pi} = (\pi_1, ..., \pi_{R_n}) \in \Pi^{R_n}$,

$$S_{n,\tilde{\pi}} = \frac{1}{\sqrt{R_n}} \sum_{r=1}^{R_n} Y_{n,r},$$

where $Y_{n,r} = \tau_n((X_{\pi_r(i),n})_{i=1}^{b_n})$ and $b_n < n$ is the subsample size. Suppose that our test statistic is centered, so that

$$E[Y_{n,r}|Z_n] = 0, \text{ a.e.}$$

Our main focus in this section is an Edgeworth expansion of the conditional distribution of $S_{n,\tilde{\pi}}$ given $Z_n$, where $\pi_r$’s are drawn i.i.d. from the uniform distribution on $\Pi$.

Let us enumerate

$$\{\tau_n((X_{\pi(i),n})_{i=1}^{b_n}) : \pi \in \Pi\} = \{W_{j,n} : j = 1, ..., p_n\},$$

with $p_n$ denoting the cardinality of the set on the left hand side, and let the empirical measure of $\{W_{j,n} : j = 1, ..., p_n\}$ be denoted by $H_n$, i.e., $H_n$ is a discrete distribution which gives a point mass of $1/p_n$ at $W_{j,n}$ for each $j = 1, ..., p_n$. Then $Y_{n,r}, r = 1, ..., R_n$, are i.i.d. draws from the distribution $H_n$. Let the conditional distribution of $S_{n,\tilde{\pi}}$ given $Z_n$ be denoted by $Q_n$. We call $Q_n$ a randomized subsampling distribution.

The randomized subsampling distribution is different from the subsampling distribution proposed in (Politis and Romano (1994)), in the sense that conditional on $Z_n$, the distribution is the sum of i.i.d. random variables from an empirical distribution. In this sense, it is closer to the $m$ out of $n$ bootstrap distribution. The $m$ out of $n$ bootstrap focuses on the conditional distribution of $Y_{n,r}$ given $Z_n$ (with subsample size $b_n$ playing the role of $m$ here). (Bickel, Götze, and van Zwet (1997)) In contrast, our focus is on the conditional distribution of $S_{n,\tilde{\pi}}$. Closely related concept is the bag of little bootstraps (BLB) recently
proposed by Kleiner, Talwalker, Sarkar, and Jordan (2014). Unlike the randomized sub-
sampling distribution, the BLB bootstrap method focuses on the CDF (or its functionals)
\[
\frac{1}{R_n} \sum_{r=1}^{R_n} 1\{Y_{n,r} \leq t\}
\]
as an approximation of the CDF of \( \tau_n((X_{i,n})_{i=1}^{n}) \). The main motivation for BLB is to reduce
the computational costs which are substantial when \( n \) is large and the computation of
\( \tau_n((X_{i,n})_{i=1}^{n}) \) is complex. The use of randomized subsampling distribution is proposed by
Song (2018), as a device for inference on a parameter, when the observat-
ions are locally
dependent but the dependence ordering is not known to the researcher.

Given each multi-index \( \nu = (\nu_1, \ldots, \nu_d) \), with \( \nu_k \) being a nonnegative integer, we let \( \bar{\chi}_\nu \) be
the average of the \( \nu \)-th cumulant of \( H_n \). Define
\[
\bar{\Omega}_n \equiv \frac{1}{p_n} \sum_{j=1}^{p_n} W_{j,n} W'_{j,n}.
\]

As before, for each \( j = 1, 2, \ldots \), let \( \bar{P}_j(z : \{\bar{\chi}_\nu\}) \) be a polynomial in \( z \) as given in (7.3) of
BR, p.52. Let us define a signed measure \( \bar{Q}_n \) on the Borel \( \sigma \)-field of \( \mathbb{R}^d \) as follows: for each
Borel \( B \subset \mathbb{R}^d \),
\[
\bar{Q}_n(B) \equiv \int_B \sum_{j=1}^{s-1} R_n^{-j/2} \bar{P}_j(-D; \{\bar{\chi}_\nu\}) \phi_{0,\hat{\Omega}_n}(x) dx,
\]
where \( \phi_{0,\hat{\Omega}_n} \) denotes the multivariate normal density with mean 0 and variance matrix \( \hat{\Omega} \).
The measure \( \bar{Q}_n \) is the Edgeworth expansion of \( Q_n \). Our first focus is on finding a finite
sample bound for the error in the approximation of \( Q_n \) by \( \bar{Q}_n \). For this, let us use the
following assumptions.

**Assumption 3.1.** There exist integers \( n_1, s \geq 3 \) and \( \delta, \bar{\rho} > 0 \) such that for all \( n \geq n_1 \),
\[
\sup_{P \in \mathcal{P}_n} \frac{1}{p_n} \sum_{j=1}^{p_n} \mathbb{E}[\|W_{j,n}\|^{s+a}] \leq \bar{\rho},
\]
where \( p_n \geq s - 2 \) for all \( n \geq n_1 \) and
\[
a = \frac{4\delta - (s - 2)}{s - 1 - 2\delta}, \quad \text{and} \quad \frac{s - 2}{2} < \delta < \frac{s - 1}{2}.
\]

**Assumption 3.2.** The \( p_n \) dimensional random vector, \( (W_{j,n})_{j=1}^{p_n} \), is absolutely continuous
with respect to Lebesgue measure.

Assumption 3.1 requires a uniform moment bound. Assumption 3.2 is stronger than the
classical Cramér condition for the original sampling distribution. The latter condition is
often used for proving higher order refinements for bootstrap t tests. (See Hall (1992).) Using Corollary 2.1, we provide an explicit finite sample bound for the error term, because we need to obtain a bound that is uniform over the distributions of $X_{i,n}$'s.

**Theorem 3.2.** Suppose that Assumptions 3.1 and 3.2 hold. Then there exist a constant $C > 0$ and an integer $n_0 \geq 1$ such that for any $n \geq n_0$, 

$$
\sup_{B \in \mathcal{C}_d} \left| Q_n(B) - \tilde{Q}_n(B) \right| \leq CR_n^{-\delta} + C_1 D_n,
$$

where $\mathcal{C}_d$ denotes the set of the convex subsets of $\mathbb{R}^d$, constants $C$ and $n_0$ depend only on $n_1$ and $d$, $s$, $\bar{\rho}$, $\delta$, and $D_n$ is a nonnegative random variable such that 

$$
E[D_n] \leq R_n^{-\delta}.
$$

Our next result is related to the modulus of continuity of the randomized subsampling distribution $Q_n$. As the distribution $Q_n$ is discrete, the modulus of continuity around convex sets can be established only up to $R_n^{-1/2}$ when we use a normal approximation of $Q_n$ (under independence or proper local dependence assumption on the observations.) However, using a higher order approximation of $Q_n$ by $\tilde{Q}_n$, we can achieve the modulus of continuity up to the order $R_n^{-\delta}$ with $(s - 2)/2 < \delta < (s - 1)/2$, depending on the moment conditions. This result can be useful when we compare two tests in terms of their power properties where their empirical sizes have a higher order accuracy. (See, e.g. Song (2018).)

**Assumption 3.3.** There exist $n_2 \geq 1$, $\varepsilon_1$, $\nu > 0$, and $M_1 > 0$ such that for all $n \geq n_2$, 

$$
P\left\{ \lambda_{\min}(\hat{\Omega}) \leq \varepsilon_1 \right\} \leq \nu, \quad \text{and} \quad P \left\{ \frac{1}{p_n} \sum_{j=1}^{p_n} \|W_{j,n}\|^{s+2} \geq M_1 \right\} \leq \nu,
$$

where $\lambda_{\min}(\hat{\Omega})$ denotes the minimum eigenvalue of $\hat{\Omega}$.

These assumptions can be shown to be satisfied under appropriate moment conditions. For example in the case where $X_{i,n}$’s are independent across $i$’s, we may use Rosenthal’s inequality to derive the the desired result.

**Corollary 3.1.** Suppose that the assumptions of Theorem 3.2 and Assumption 3.3 holds. Then there exist a constant $C > 0$ and an integer $n_0 \geq 1$ such that for any $n \geq n_0$ and for any $\eta > 0$, 

$$
\sup_{t \geq 0} \left| Q_n(\hat{A}(t)) - Q_n(\hat{A}(t + \eta)) \right| \leq C \left( R_n^{-\delta} + \nu + \eta \right) + 2C_1 D_n,
$$

where $\mathcal{C}_d$ denotes the set of the convex subsets of $\mathbb{R}^d$, constants $C$ and $n_0$ depend only on $n_1$ and $d$, $s$, $\bar{\rho}$, $\delta$, and $D_n$ is a nonnegative random variable such that 

$$
E[D_n] \leq R_n^{-\delta}.$$
where constants $C$ and $n_0$ depend only $n_1, n_2$ and $d, s, \bar{\rho}, \delta$ and $M_1$, and $D_n$ is a nonnegative random variable in Theorem 3.2, and

$$\hat{A}(t) \equiv \left\{ x \in \mathbb{R}^d : x'\hat{\Omega}x \leq t \right\}.$$ 

4. Proofs

We introduce notation. For a measure or a signed measure $\mu$ and a measurable function $f$, we write

$$\mu(f) = \int f \, d\mu,$$

for brevity. Recall that the Fourier-Stieltjes transform of a signed measure $\mu$ on $\mathbb{R}^d$ is a complex-valued function on $\mathbb{R}^d$:

$$\hat{\mu}(t) = \int_{\mathbb{R}^d} e^{it'x} \mu(dx), \quad t \in \mathbb{R}^d.$$

Recall that $\| \cdot \|$ denotes the Euclidean norm on $\mathbb{R}^d$, i.e., $\|a\|^2 = a'a, \ a \in \mathbb{R}$. We also define $|a| = \sum_{k=1}^d |a_k|$, for any $a = (a_1, \ldots, a_d)' \in \mathbb{R}^d$. For a given $d \times d$ matrix $T$, we define the operator norm:

$$\|T\| \equiv \sup_{x \in \mathbb{R}^d : \|x\| \leq 1} \|Tx\|.$$

**Proof of Proposition 2.1:** Let us write $\mathcal{M}_d(u) = \mathcal{M}_{d,p}(u)$. We let $\mathcal{U}_d = \mathcal{U}_d(2b, R)$. We use mathematical induction. The case with $d = 1$ is proven in Angst and Poly (2017). Now, suppose that the statement of Proposition 2.1 holds with $d = d' - 1$ for $d' \geq 2$. Given $u = (u_j)_{j=1}^p$ with $u_j \in \mathbb{R}^{d'}$, let $\mathcal{M}'(u), \ u = (u_1, u_2), \ \text{with} \ u_1 \in \mathbb{R}^{d'-1}$ and $u_2 \in \mathbb{R}$, be a collection where each member is a product measures between a measure, say, $P_1$, from $\mathcal{M}_{d'-1}(u_1)$ and a measure, say, $P_2$, from $\mathcal{M}_1(u_2)$. Let the set of $u$'s such that $\mathcal{M}'(u)$ does not satisfy the weak Cramér condition with parameter $(2b, R)$ be denoted by $\mathcal{U}'$. We first show that

$$(12) \quad \mathcal{U}' \subset \mathcal{U}_{d'-1} \times \mathcal{U}_1.$$ 

Choose $u = (u_1, u_2) \in \mathbb{R}^{d'-1} \times \mathbb{R}$ such that either $\mathcal{M}_{d'-1}(u_1)$ or $\mathcal{M}_1(u_2)$ satisfies the weak Cramér condition with parameter $(2b, R)$. Let the characteristic functions of $P_1 \in \mathcal{M}_{d'-1}(u_1)$ and $P_2 \in \mathcal{M}_1(u_2)$ be denoted by $\phi_1$ and $\phi_2$ respectively, and the characteristic function of product measure $P_1 \times P_2$ be denoted by $\phi_{12}$. Then, for $t_1 \in \mathbb{R}^{d'-1}$ and $t_2 \in \mathbb{R}$
such that \( \| t_1 \| > R \) and \( \| t_2 \| > R \),
\[
|\phi_{12}((t_1, t_2))| \leq |\phi_1(t_1)||\phi_2(t_2)| \leq \max\left\{ 1 - \frac{1}{\| t_1 \|^b}, 1 - \frac{1}{\| t_2 \|^b} \right\}
\leq 1 - \frac{1}{\| (t_1, t_2) \|^b}.
\]

Since \( \| (t_1, t_2) \| \geq R \), \( \mathcal{M}'(u) \) satisfies the weak Cramér condition with parameter \((2b, R)\). This means that if \( \mathcal{M}'(u) \) does not satisfy the weak Cramér condition with parameter \((2b, R)\), neither of \( \mathcal{M}_{d-1}(u_1) \) and \( \mathcal{M}_1(u_2) \) with \( u = (u_1, u_2) \) does. This proves (12). Thus we have
\[
\lambda (\mathcal{U}') \leq \lambda (\mathcal{U}_{d-1}) \times \lambda (\mathcal{U}_1) = 0,
\]
by the hypothesis of the induction.

Since \( \delta_{u_j} = \delta_{u_{j1}} \times \cdots \times \delta_{u_{jd}} \) for \( u_j = (u_{j1}, \ldots, u_{jd}) \) in general, we find that \( \mathcal{M}'(u) \subset \mathcal{M}(u) \) for each \( u \). This means that if \( \mathcal{M}(u) \) does not satisfy the weak Cramér condition with parameter \((2b, R)\), neither does \( \mathcal{M}'(u) \) with the same parameter. This means that \( \mathcal{U}' \) contains \( \mathcal{U}_d \), completing the proof. ■

Define
\[
Y_{i,n} \equiv X_{i,n,1}\{\|X_{i,n}\| \leq \sqrt{n}\}, \text{ and } Z_{i,n} \equiv Y_{i,n} - E[Y_{i,n}],
\]
and let \( \phi_{i,n,P}(t) \equiv E_P[\exp(it'Z_{i,n})] \). Let us begin with an auxiliary lemma which is a uniform-in-\( P \) modification of Proposition 2.10 and Lemma 5.4 of AP.

**Lemma 4.1.** Let \( \{X_{i,n}\}_{i=1}^n \) be a triangular array of independent random vectors with values in \( \mathbb{R}^d \), and let \( \mathcal{P}_n \) be the collection of the joint distributions of \( \{X_{i,n}\}_{i=1}^n \) such that \( \mathcal{P}_n \) satisfies the mean weak Cramér condition with parameter \((b, R)\). Furthermore, assume that
\[
\sup_{n \geq 1} \sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^n E_P[\|X_{i,n}\|] < \infty.
\]

Then there exist \( \varepsilon > 0 \) and \( n_0 \) such that for all \( 0 < r < R \) and all \( n \geq n_0 \),
\[
\sup_{r \leq \|u\| \leq R} \sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^n |\phi_{X_{i,n},P}(u)| \leq 1 - \varepsilon,
\]
where \( \varepsilon > 0 \) and \( n_0 \) depend only on \( b, r, R \).

Suppose further that for some \( s \geq 3 \) and \( \bar{\rho} > 0 \),
\[
\sup_{n \geq 1} \sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^n E_P[\|X_{i,n}\|^s] \leq \bar{\rho}.
\]
Then, there exists \( \varepsilon > 0 \) and \( n_0 \) such that for all \( 0 < r < R \) and all \( n \geq n_0 \),

\[
\sup_{r \leq \|t\| \leq R} \sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^{n} |\phi_{i,n,P}(t)| \leq 1 - \varepsilon,
\]

where \( \varepsilon > 0 \) and \( n_0 \) depend only on \( s, \bar{\rho}, b, r, R \).

Furthermore, there exists \( \bar{n} \geq 1 \) such that for all \( R > 0 \), \( n \geq \bar{n} \), and for all \( t \in \mathbb{R}^d \) satisfying \( R < \|t\| \leq (1/(4\bar{\rho}))^{1/b} n^{s/(2b)} \),

\[
\sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^{n} |\phi_{i,n,P}(t)| \leq 1 - \frac{1}{2\|t\|^b},
\]

**Proof:** The first statement is obtained by following the proof of Proposition 2.10 of AP. More specifically, suppose to the contrary that we have a subsequence \( n(k) \) such that

\[
\lim_{k \to \infty} \sup_{r \leq \|t\| \leq R} \sup_{P \in \mathcal{P}_{n(k)}} \frac{1}{n(k)} \sum_{i=1}^{n(k)} |\phi_{i,n(k),P}(t)| = 1.
\]

Since \( \frac{1}{n(k)} \sum_{i=1}^{n(k)} |\phi_{X_i(n(k)),P}(u)| \) is uniformly continuous on \( \mathbb{R}^d \), for each \( P \in \mathcal{P}_{n(k)} \), so is \( \sup_{P \in \mathcal{P}_{n(k)}} \frac{1}{n(k)} \sum_{i=1}^{n(k)} |\phi_{X_i,n(k),P}(u)| \) on \( \mathbb{R}^d \). Take a sequence \( u_k \) such that \( r \leq \|u_k\| \leq R \) and

\[
\sup_{P \in \mathcal{P}_{n(k)}} \frac{1}{n(k)} \sum_{i=1}^{n(k)} |\phi_{X_i,n(k),P}(u_k)| = \sup_{r \leq \|t\| \leq R} \sup_{P \in \mathcal{P}_{n(k)}} \frac{1}{n(k)} \sum_{i=1}^{n(k)} |\phi_{i,n(k),P}(t)|.
\]

By the definition of supremum, we can take a further subsequence \( \{n_{k_j}\} \) of \( \{n_k\} \) such that

\[
\lim_{j \to \infty} \frac{1}{n(k_j)} \sum_{i=1}^{n(k_j)} |\phi_{X_i(n(k_j)),P_n(u_k)}| = \lim_{k \to \infty} \sup_{P \in \mathcal{P}_{n(k)}} \frac{1}{n(k)} \sum_{i=1}^{n(k)} |\phi_{X_i,n(k),P}(u_k)|.
\]

We can follow the rest of the proof in AP, p.6-7, to arrive at a contradiction to the weak Cramér condition.

As for the second result, by (5.17) of AP,

\[
\sup_{r \leq \|t\| \leq R} \sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^{n} |\phi_{i,n,P}(t)| \leq \sup_{r \leq \|t\| \leq R} \sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^{n} |\phi_{i,n,P}(t)| + \frac{2\bar{\rho}}{n^{s/2}}.
\]

Thus the desired result follows from the first result.

As for the last result, we take supremum over \( P \in \mathcal{P}_n \) on both sides of (5.17) in Lemma 5.4 of AP and follow the same arguments in the proof of the lemma using Definition 2.1.

**Proof of Theorem 2.1:** We walk through the steps in the proofs of Theorem 20.1 of BR and Theorem 4.3 of AP, making bounds in the error explicit in finite samples. Throughout the proof, we assume without loss of generality that \( V_n,P = I_d \). While the proof of Theorem
20.1 and Theorem 4.3 assumes that \( X_{i,n}'s \) are i.i.d., here we make explicit that they are allowed to be non-identically distributed. In the proof, as in BR, we use notation \( c_j(x) \) to denote a constant that depends only on \( x \), so that, for example, \( c_1(s,d) \) is a constant that depends only on \( s \) and \( d \). Let \( Q''_{n,p} \) and \( Q''_{n,p} \) be the distributions of \( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n} \) and \( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_{i,n} \) respectively. We take \( a_n \equiv \sqrt{n}E[Y_{i,n}] \), and define \( f_{a_n}(x) = f(x + a_n) \). Using the identity \( Q''_{n,p}(f) = Q''_{n,p}(f_{a_n}) \), we bound

\[
|Q_{n,p}(f) - \tilde{Q}_{n,s,p}(f)| \leq A_{1n} + A_{2n} + A_{3n},
\]

where

\[
A_{1n} \equiv |Q'_{n,p}(f) - Q''_{n,p}(f)|,
\]

\[
A_{2n} \equiv |Q'_{n,p}(f_{a_n}) - \tilde{Q}_{n,s,p}(f_{a_n})|,
\]

and

\[
A_{3n} \equiv |\tilde{Q}_{n,s,p}(f_{a_n}) - \tilde{Q}_{n,s,p}(f)|.
\]

By (20.8) of BR, p.208, and (20.12) of BR, p.209, we find that

\[
(13) \quad A_{1n} + A_{3n} \leq c_1(s', s, d) M_{s'}(f) n^{-(s-2)/2} \Delta_{n,s,p}(1).
\]

Let us focus on finding a bound for \( A_{2n} \). Let \( D_{n,p} \equiv \frac{1}{n} \sum_{i=1}^{n} \text{Var}_p(Z_{i,n}) \) and define a signed measure \( \tilde{Q}_{n,s,p}' \) as

\[
(14) \quad \tilde{Q}_{n,s,p}'(A) = \int_{A} \sum_{j=0}^{s-2} n^{-j/2} \tilde{P}_j(-D; \{\tilde{X}_{\nu,p}\}) \phi_{0,D_{n,p}}(x) dx, \quad \text{for any Borel set } A,
\]

where \( \phi_{0,D_{n,p}} \) denotes the density of \( N(0, D_{n,p}) \). We bound

\[
(15) \quad A_{2n} \leq |\tilde{Q}'_{n,s,p}(f_{a_n}) - \tilde{Q}_{n,s,p}'(f_{a_n})| + A_{2n}',
\]

\[
\leq c_2(s', s, d) M_{s'}(f) n^{-(s-2)/2} \Delta_{n,s,p}(1) + A_{2n}',
\]

where \( A_{2n}' \equiv |Q'_{n,p}(f_{a_n}) - \tilde{Q}'_{n,s,p}(f_{a_n})| \), and the last inequality comes from (20.13) of BR, p.209. Let us focus on \( A_{2n}' \). Define

\[
H_{n,p} \equiv Q_{n,p} - \tilde{Q}_{n,s,p}.
\]

For any positive number \( 0 < \varepsilon < 1 \), we define \( K_{\varepsilon} \) to be a probability measure such that

\[
K_{\varepsilon}(\{x \in \mathbb{R}^d : \|x\| < \varepsilon\}) = 1,
\]

and

\[
|D^\varepsilon K_{\varepsilon}(t)| \leq c_3(s, d) \varepsilon^{\alpha_1} \exp \left(-\frac{\varepsilon \|t\|^2}{2}\right)
\]

for any \( \varepsilon > 0 \).
for all $t \in \mathbf{R}^d$ and all $|\alpha| \leq s + d + 1$. Then, by (20.17) of BR, p.210, we find that

$$|H_{n,P}(f_{a_n})| \leq M_s'(f) \int \left(1 + (||x|| + \varepsilon + ||a||)^s) |H_{n,P} \ast K_{\varepsilon}|(dx)
+ \omega_{f_{a_n}}(2\varepsilon; |\tilde{Q}_{n,s+d,P}|) \right) \equiv B_{1n} + B_{2n},$$

say.

From (20.19) of BR, p.210, we bound

$$B_{1n} \leq c_4(s', s, d) \max_{0 \leq |\beta| \leq s+d+1} \int |D^\beta(\hat{H}_{n,P}\hat{K}_{\varepsilon})(t)| dt.$$

We write

$$D^\beta(\hat{H}_{n,P}\hat{K}_{\varepsilon}) = \sum_{0 \leq \alpha \leq \beta} c_5(\alpha, \beta)(D^{\beta-\alpha}\hat{H}_{n,P})(D^\alpha\hat{K}_{\varepsilon}).$$

Now, let us focus on finding a bound for

$$\int \left| (D^{\beta-\alpha}\hat{H}_{n,P}(t))(D^\alpha\hat{K}_{\varepsilon}(t)) \right| dt.$$

Define for some constant $c_6(s, d)$ which depends only on $s$ and $d$,

$$A_n \equiv c_6(s, d)n^{1/2} \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{E}_P \|D_{n,P}^{-1/2}Z_{i,n}\|^{|s+d+1|} \right)^{1/(s+d+1)} \Lambda_n^{1/2},$$

where $\Lambda_n$ denotes the maximum eigenvalue of $D_{n,P}$, and bound the integral in (17) by

$$\int_{||t|| \leq A_n} \left| (D^{\beta-\alpha}\hat{H}_{n,P}(t))(D^\alpha\hat{K}_{\varepsilon}(t)) \right| dt$$

+ $$\int_{||t|| > A_n} \left| (D^{\beta-\alpha}\hat{H}_{n,P}(t))(D^\alpha\hat{K}_{\varepsilon}(t)) \right| dt.$$

By (20.21) of BR, p.210, we can choose $c_6(s, d)$ in the definition of $A_n$ so that the leading term in (19) is bounded by

$$c_7(s, d)n^{-(s+d-1)/2} \frac{1}{n} \sum_{i=1}^{n} \mathbf{E}_P \left[ \|D_{n,P}^{-1/2}Z_{i,n}\|^{|s+d+1|} \right]$$

$$\leq c_7(s, d)n^{-(s+d-1)/2} \left\| D_{n,P}^{-1/2} \right\|^{s+d+1} \frac{1}{n} \sum_{i=1}^{n} \mathbf{E}_P \left[ \|Z_{i,n}\|^{|s+d+1|} \right]$$

$$= c_7(s, d)n^{-(s+d-1)/2} \left\| D_{n,P}^{-1/2} \right\|^{(s+d+1)/2} \frac{1}{n} \sum_{i=1}^{n} \mathbf{E}_P \left[ \|Z_{i,n}\|^{|s+d+1|} \right]$$
From (14.23) and (14.24) of BR, p.125, and our assumption that \( V_{n,P} = I_d \), and because \( D_{n,P} \) is symmetric and \( \|D_{n,P}\| \) denotes the operator norm of \( D_{n,P} \), we find that

(21) \[ \|D_{n,P} - I_d\| \equiv \sup_{\|t\| \leq 1} t'(D_{n,P} - I_d)t \leq 2dn^{-(s-2)/2}\bar{\Delta}_{n,s,P}(1), \]

and

(22) \[ \|D_{n,P}^{-1}\| = \|(I_d + (D_{n,P} - I_d))^{-1}\| \leq \frac{1}{1 - \|D_{n,P} - I_d\|} \leq \frac{1}{1 - 2dn^{-(s-2)/2}\bar{\Delta}_{n,s,P}(1)}. \]

As noted in (20.23) of BR, p.211,

(23) \[ E_P\|Z_{i,n}\|^{s+d+1} \leq 2^{s+d+1}E_P\|Y_{i,n}\|^{s+d+1}. \]

For any \( \xi > 0 \), we can bound

(24) \[ \frac{1}{n} \sum_{i=1}^{n} E_P\|Y_{i,n}\|^{s+d+1} \leq (\xi n^{1/2})^{d+1} \frac{1}{n} \sum_{i=1}^{n} E_P \left[ 1\{\|X_{i,n}\| \leq \xi n^{1/2}\} \|X_{i,n}\|^s \right] + n^{(d+1)/2} \frac{1}{n} \sum_{i=1}^{n} E_P \left[ 1\{\xi n^{1/2} < \|X_{i,n}\| \leq n^{1/2}\} \|X_{i,n}\|^s \right] \leq n^{(d+1)/2} \left\{ \xi^{d+1} \rho_{n,s,P} + \bar{\Delta}_{n,s,P}(\xi) \right\}. \]

This gives a bound for \( \frac{1}{n} \sum_{i=1}^{n} E_P\|Z_{i,n}\|^{s+d+1} \) through (23). Using this bound and (22), we find that for any \( \xi > 0 \),

(25) \[ \frac{1}{n} \sum_{i=1}^{n} E_P \left[ \|D_{n,P}^{-1/2} Z_{i,n}\|^{s+d+1} \right] \leq 2^{s+d+1} n^{(d+1)/2} \left( \xi^{d+1} \rho_{n,s,P} + \bar{\Delta}_{n,s,P}(\xi) \right) \left(1 - 2dn^{-(s-2)/2}\bar{\Delta}_{n,s,P}(1)\right)^{(s+d+1)/2}, \]

and from (20), we obtain that

(26) \[ \int_{\|t\| \leq A_n} \left| (D^{\beta-\alpha}\hat{H}_{n,P}(t))(D^{\alpha}\hat{K}_s(t)) \right| dt \leq c_7(s,d)n^{-(s-2)/2} \left( \xi^{d+1} \rho_{n,s,P} + \bar{\Delta}_{n,s,P}(\xi) \right) \left(1 - 2dn^{-(s-2)/2}\bar{\Delta}_{n,s,P}(1)\right)^{(s+d+1)/2}. \]

Let us turn to the second integral in (19). Note first that

\[ \Lambda_n \leq \|D_{n,P}\| \leq \|D_{n,P} - I_d\| + \|I_d\| \leq 1 + 2dn^{-(s-2)/2}\bar{\Delta}_{n,s,P}(1), \]

and

\[ \int_{\|t\| \leq A_n} \left| (D^{\beta-\alpha}\hat{H}_{n,P}(t))(D^{\alpha}\hat{K}_s(t)) \right| dt \leq c_7(s,d)n^{-(s-2)/2} \left( \xi^{d+1} \rho_{n,s,P} + \bar{\Delta}_{n,s,P}(\xi) \right) \left(1 - 2dn^{-(s-2)/2}\bar{\Delta}_{n,s,P}(1)\right)^{(s+d+1)/2}. \]
by (21). Using this and noting (25), we find from the definition of \( A_n \) in (18) that

\[
A_n \geq \frac{c_6(s,d)n^{s-2}}{(\epsilon^{d+1}p_\nu \Delta n,s,P(\epsilon))^{s+d-1}/2} \left( 1 - 2dn^{-(s-2)/2} \Delta n,s,P(1) \right)^{s+d+1/2}.
\]

if we confine \( \xi \in (0,1) \). We take

\[
q_n = \frac{\sqrt{n}}{16\rho},
\]

and, as in (20.26) of BR, p.211, let us bound

\[
\int_{\|t\| > A_n} \left| (D^{\beta-\alpha} \mathcal{H}_{n,P}(t)) (D^{\alpha} \mathcal{K}_\epsilon(t)) \right| dt \leq I_1 + I_2 + I_3,
\]

where

\[
I_1 = \int_{\|t\| > q_n} \left| (D^{\beta-\alpha} \mathcal{Q}_{n,P}(t)) (D^{\alpha} \mathcal{K}_\epsilon(t)) \right| dt
\]

\[
I_2 = c_8(s,d) \int_{\|t\| > A_n} (1 + \|t\|^{\beta-\alpha}) \exp \left( -\frac{5}{24} \|t\|^2 \right) dt,
\]

\[
I_3 = \int_{\|t\| > A_n} D^{\beta-\alpha} \sum_{j=0}^{s+d-2} n^{-j/2} \mathcal{P}_j(it; \chi_{\nu,P}) \exp \left( -\frac{1}{2} t' D_{n,P} t \right) dt.
\]

Let us deal with \( I_2 \) first. Choose any large number \( w \geq 1 \). We bound

\[
I_2 \leq c_8(s,d) A_n^{-w} \int_{\|t\| > A_n} (1 + \|t\|^{\beta-\alpha}) \|t\|^w \exp \left( -\frac{5}{24} \|t\|^2 \right) dt
\]

The last integral is bounded for any fixed \( w \geq 1 \). In the light of the lower bound for \( A_n \) in (27), we find that for any \( w \geq 1 \), there exist a constant \( C \) and a number \( n_{0,1} \) which depend only on \((s,d,\rho,w)\) such that for all \( n \geq n_{0,1} \),

\[
I_2 \leq C n^{-\frac{(s-2)w}{2(s+d-1)}}.
\]

Since the integrand in \( I_3 \) involves \( \exp(-t'D_{n,P}t/2) \), we obtain the same conclusion for \( I_3 \) as well.

The rest of the proof focuses on finding a finite sample bound for \( I_1 \). In doing so, we switch to the proof of Theorem 4.3 of AP. By (5.15) of AP, p.19, we bound

\[
I_1 \leq c_9(s,d) \epsilon^{[\beta-\alpha]} \sum_{\gamma \in (\mathbb{Z}_+^d)^n \ |\gamma|=|\beta-\alpha|} \binom{\beta-\alpha}{\gamma} J_\gamma(n, \epsilon),
\]
where

\[ J_\gamma(n, \varepsilon) \equiv n^{d/2} \int_{\|t\| \geq q_n/\sqrt{n}} \prod_{i=1, \gamma_i = 0}^n \phi_{i,n,P}(t) \left| e^{-\sqrt{n\varepsilon}\|t\|^{1/2}} \right| dt, \]

and

\[ \left( \frac{\beta - \alpha}{\gamma} \right) = \frac{|\beta - \alpha|!}{\prod_{i=1}^n |\gamma_i|!}. \]

We let

\[ r = 1/(16\bar{\rho}), \]

so that \( q_n/\sqrt{n} = r \) for all \( n \geq 1 \). With this choice of \( r > 0 \), from (5.16) of AP, p.20, we deduce that for all \( n \geq n_1 \),

\[ J_\gamma(n, \varepsilon) \leq n^{d/2} \exp \left( (n - |\gamma|) \log \left( \frac{n}{n - |\gamma|} \right) \right) K_\gamma(n, \varepsilon), \]

where

\[ K_\gamma(n, \varepsilon) = \int_{\|t\| \geq r} \exp \left( (n - |\gamma|) \log \left( \frac{1}{n} \sum_{i=1}^n |\phi_{i,n,P}(t)| \right) \right) e^{-\left( \varepsilon \sqrt{n\|t\|} \right)^{1/2}} dt \]

By Lemma 4.1, there exists \( \bar{n} \geq 1 \) such that for all \( n \geq \bar{n} \) and all \( t \) satisfying \( R < \|t\| \leq (1/(4\bar{\rho}))^{1/b} n^{s/(2b)} \),

\[ \sup_{P \in \mathcal{P}_n} \frac{1}{n} \sum_{i=1}^n |\phi_{i,n,P}(t)| \leq 1 - \frac{1}{2\|t\|^{b_0}}. \]

Take \( n'_0 = \max\{\bar{n}, n_{0,1}\} \), and write

\[ K_\gamma(n, \varepsilon) = K^1_\gamma(n, \varepsilon) + K^2_\gamma(n, \varepsilon) + K^3_\gamma(n, \varepsilon), \]

where

\[ K^1_\gamma(n, \varepsilon) \equiv \int_{r \leq \|t\| \leq R} \exp \left( (n - |\gamma|) \log \left( \frac{1}{n} \sum_{i=1}^n |\phi_{i,n,P}(t)| \right) \right) e^{-\left( \varepsilon \sqrt{n\|t\|} \right)^{1/2}} dt \]

\[ K^2_\gamma(n, \varepsilon) \equiv \int_{R < \|t\| \leq (1/(4\bar{\rho}))^{1/b} n^{s/(2b)}} \exp \left( (n - |\gamma|) \log \left( \frac{1}{n} \sum_{i=1}^n |\phi_{i,n,P}(t)| \right) \right) e^{-\left( \varepsilon \sqrt{n\|t\|} \right)^{1/2}} dt \]

\[ K^3_\gamma(n, \varepsilon) \equiv \int_{(1/(4\bar{\rho}))^{1/b} n^{s/(2b)} < \|t\|} \exp \left( (n - |\gamma|) \log \left( \frac{1}{n} \sum_{i=1}^n |\phi_{i,n,P}(t)| \right) \right) e^{-\left( \varepsilon \sqrt{n\|t\|} \right)^{1/2}} dt. \]

By (5.20) of AP, p.21, for all \( n \geq n'_0 \),

\[ K^1_\gamma(n, \varepsilon) \leq \frac{\pi^{d/2}}{\Gamma((d/2) + 1)} R^d e^{-\left( (n - |\gamma|) \eta \right)} \]
where \( \Gamma \) is the gamma function, and
\[
\eta \equiv 1 - \tilde{\phi}(1/(16\bar{\rho}), R, n_0).
\]
Hence
\[
n^{d/2} \exp \left( (n - |\gamma|) \log \left( \frac{n}{n - |\gamma|} \right) \right) K_1^1(n, \varepsilon) \\
\leq c_0(d)n^{d/2} \exp \left( (n - |\gamma|) \left( \log \left( \frac{n}{n - |\gamma|} \right) - \eta \right) \right) R^d \\
\leq c_0(d)n^{d/2} \exp \left( (n - s - d - 1) \left( \log \left( \frac{n}{n - s - d - 1} \right) - \eta \right) \right) R^d.
\]
Let us consider \( K_2^2(n, \varepsilon) \). Let
\[
R(n, \varepsilon) \equiv \varepsilon \left( \frac{1}{4\bar{\rho}} \right)^{1/b} n^{\frac{s}{2b} + \frac{1}{2}}.
\]
By (5.21) of AP, p.22, we find that for all \( n \) such that \( R(n, \varepsilon) > 4d^2 \),
\[
K_2^3(n, \varepsilon) \leq c_0(d)(\varepsilon \sqrt{n})^{-d} R(n, \varepsilon)^d \varepsilon^{-\sqrt{R(n, \varepsilon)}}.
\]
Note that \( R(n, \varepsilon) \) increases to infinity at the polynomial rate in \( n \). Lastly, by (5.23) and (5.24) on p.22 of AP, for any \( a > 0 \) such that \( 1 - ab > 0 \), we have
\[
K_2^2(n, \varepsilon) \leq c_0(d)(\varepsilon \sqrt{n})^{-d} \exp \left( -n^{1-ab} (1 + o(1)) \right) + c_1(d)(\varepsilon \sqrt{n})^{-d} \int_{u \geq \varepsilon n^{a+1/2}} \exp^{-\sqrt{u}} u^{d-1} du,
\]
where \( o(1) \) is a sequence of numbers decreasing to zero as \( n \to \infty \) and depends only on \( s, d, b, R, \varepsilon, a \) only. Thus, we conclude as in AP that if we choose \( \varepsilon = n^{-\delta} \) with \( \delta > 0 \) satisfying (8), for any \( w \geq 1 \), there exist a constant \( C \) and a number \( n_{0,2} \) which depend only on \( s, d, \bar{\rho}, b, R, \delta, w \) such that for all \( n \geq n_{0,2} \),
\[
I_1 + I_2 + I_3 \leq C n^{-w}.
\]
Thus, we conclude that there exist a constant \( C \) and \( n_0 \) which depend only on \( s, d, \bar{\rho}, b, R, \delta, w \) such that for all \( n \geq n_0 \), the first integral in (19) dominates the second one. Collecting the bounds in (13), (15), (16), and (26), we find that
\[
|Q_{n,P}(f) - \bar{Q}_{n,s,P}(f)| \leq c_{21}(s', s, d) M_{s'}(f) n^{-(s-2)/2} \Delta_{n,s,P}(1) + \bar{\omega}_{fan} \left( 2n^{-\delta}; |\bar{Q}'_{n,s+d,P} | \right).
\]
Recall the definition $\tilde{Q}'_{n,s,P}$ in (14) and define
\[
P_r(-\Phi_0,D_n,P;\{\tilde{X}_n,P\})(A) \equiv \int_A \tilde{P}_j(-D;\{\tilde{X}_n,P\})\phi_{0,D_n,P}(x)dx, \text{ for any Borel set } A.
\]

As for the last term, we follow (20.36) and (20.37) of BR, p.213, to find that for absolute constants $C, C' > 0$, for any $0 \leq j \leq s - 2$,
\[
\begin{align*}
\bar{\omega}_{\rho_n}(2n^{-\delta}; n^{-j/2}|P_r(-\Phi_0,D_n,P;\{\tilde{X}_n,P\})| & \leq C\rho_{n,s,P} \left(M_{s'}(f) \int (1 + \|x\|^{s'})|\phi_{a_n,D_n,P}(x) - \phi(x)||dx + \bar{\omega}_f(2n^{-\delta}; \Phi)\right) \\
& \quad + Cn^{-j/2}\rho_{n,s,P} \int_{\|x\| > n^{1/6}} \left(1 + \|x\|^{3j+s'}\right)\phi_{a_n,D_n,P}(x)dx \\
& \leq c_{13}(s,d,j)M_{s'}(f)n^{-(s-2)/2}\Delta_{n,s,P}(1) + C'\rho_{n,s,P}\bar{\omega}_f(2n^{-\delta}; \Phi),
\end{align*}
\]
where the last inequality uses Lemma 14.6 of BR, p.131. Also, from (20.39) of BR on p.213, for $s - 1 \leq j \leq s + d - 2$,
\[
\begin{align*}
\bar{\omega}_{\rho_n}(2n^{-\delta}; n^{-j/2}|P_r(-\Phi_0,D_n,P;\{\tilde{X}_n,P\})| & \leq c_{14}(s,d,j)n^{-(s-2)/2} \frac{1}{n} \sum_{i=1}^{n} E_P \|Z_{i,n}\|^{j+2}M_{s'}(f) \int (1 + \|x\|^{3j+s'})\phi_{a_n,D_n,P}(x)dx.
\end{align*}
\]

From (23) and (24), the last term is bounded by
\[
\begin{align*}
c_{15}(s,d,j)M_{s'}(f)n^{(-s+j+2)-j/2} \left\{\xi^{s+j+2}\rho_{n,s,P} + \Delta_{n,s,P}(\xi)\right\} & \leq c_{16}(s,d,j)M_{s'}(f)n^{-(s-2)/2} \left\{\xi\rho_{n,s,P} + \Delta_{n,s,P}(\xi)\right\},
\end{align*}
\]
because $0 < \xi \leq 1$. Hence
\[
\begin{align*}
\bar{\omega}_{\rho_n}(2n^{-\delta}; |\tilde{Q}'_{n,s+d,P}|) & \leq c_{17}(s,d)M_{s'}(f)n^{-(s-2)/2} \left\{\xi\rho_{n,s,P} + \Delta_{n,s,P}(\xi) + \Delta_{n,s,P}(1)\right\} \\
& \quad + C'\rho_{n,s,P}\bar{\omega}_f(2n^{-\delta}; \Phi).
\end{align*}
\]
Combining this with (28), we obtain the desired result. \[\square\]

**Proof of Theorem 3.1:** We set $b = 1/(s-2)$ and $R = 1/(8\bar{b})$. Let $\mathcal{U}$ be the set of $u \in \mathbb{R}^{2n}$ such that the discrete measure $\frac{1}{n} \sum_{j=1}^{n} \delta_{u_j}$ (with $\delta_{u_j}$ denoting Dirac measure at $u_j \in \mathbb{R}^2$) fails to satisfy the weak Cramér condition in Definition 2.1 with parameter $(2b, R)$. We define
\[
\mathcal{E}_{n,0} \equiv \{(X_{j,n})_{j=1}^{n} \in \mathcal{U}\}.
\]
By Lemma 2.1 and the assumption that $X_{i,n}$'s are i.i.d. continuous random variables, $P\mathcal{E}_{n,0} = 0$. Focusing on the event $\mathcal{E}_{n,0}$, we apply Theorem 2.1 with the choice of $\delta$ such that $(s - 2)/2 < \delta < (s - 1)/2$. It remains to obtain a bound for $\bar{\omega}_{f_n}(n^{-\delta}; \Phi)$, where $f_n(t) = 1\{x \in \mathbb{R} : g_n(x) \leq t\}$. For this, we follow the arguments on pages 253-254 of Hall (1992). More specifically, we write
\[
g_n(\bar{X}_{n,b}^*) = g_{n,1}(\bar{X}_{n,b}^*) + g_{n,2}(\bar{X}_{n,b}^*),
\]
where
\[
g_{n,1}(x) = \sum_{r=1}^{s} \frac{\partial^r g_n(\bar{X}_n)(x - \bar{X}_n)^r}{r!}
\]
and $|g_{n,2}(x)| \leq C'n^{-(s-1)/2}(\log n)^{s+1}$, for some constant $C' > 0$, if $|x - \bar{X}_n| \leq \log n$. Let
\[
f_{n,1}(t) = 1\{x \in \mathbb{R} : |g_{n,1}(x) - t| \leq (1 + C')n^{-\delta}, |x - \bar{X}_n| \leq \log n\}, \text{ and}
\]
\[
f_{n,2}(t) = 1\{x \in \mathbb{R} : |x - \bar{X}_n| > \log n\}.
\]
Then
\[
\bar{\omega}_{f_n}(n^{-\delta}; \Phi) \leq \Phi(f_{n,1}) + \Phi(f_{n,2}).
\]
The last term on the right hand side is bounded by a term that vanishes faster than any polynomial rate in $n$. We apply Lemma 5.3 of Hall (1992) to bound the leading term on the right hand side by $Cn^{-\delta}$ for some constant $C$ that depends only on $\delta$ and $s$. ■

**Proof of Theorem 3.2:** The matrix $\hat{\Omega}$ is positive semidefinite everywhere, and by Assumption 3.2, the minimum eigenvalue of $\hat{\Omega}$ is positive almost everywhere. Therefore, without loss of generality, we will assume that $\hat{\Omega}$ is the identity matrix with dimension $d$. As in the proof of Theorem 3.1, we set $b = 1/(s - 2)$ and $R = 1/(8\rho)$, and let $\mathcal{U}$ be the set of $u \in \mathbb{R}^{dp_n}$ such that the discrete measure $\frac{1}{p_n} \sum_{j=1}^{p_n} \delta_{u_j}$ (with $\delta_{u_j}$ denoting Dirac measure at $u_j \in \mathbb{R}^d$) fails to satisfy the weak Cramér condition in Definition 2.1 with parameter $(2b, R)$. We define two events:
\[
\mathcal{E}_{n,1} \equiv \{(W_{j,n})_{j=1}^{p_n} \in \mathcal{U}\}, \text{ and } \mathcal{E}_{n,2} \equiv \left\{\frac{1}{p_n} \sum_{j=1}^{p_n} \|W_{j,n}\|^s < R_n^\delta\right\}.
\]
We bound
\[
|Q_n(B) - \tilde{Q}_n(B)| \leq A_{1n} + A_{2n} + A_{3n},
\]
where

\[ A_{1n} \equiv |Q_n(B) - \tilde{Q}_n(B)|1_{\mathcal{E}_{n,1} \cap \mathcal{E}_{n,2}} \]

\[ A_{2n} \equiv |Q_n(B) - \tilde{Q}_n(B)|1_{\mathcal{E}_{n,1}}, \quad \text{and} \]

\[ A_{3n} \equiv |Q_n(B) - \tilde{Q}_n(B)|1_{\mathcal{E}_{n,2}}. \]

Let us consider \( A_{2n} \) first. We bound \( E[A_{2n}] \) by \( P\mathcal{E}_{n,1}^c \) which is zero by Proposition 2.1 and Assumption 3.2.

Let us turn to \( A_{1n} \). By Corollary 2.1,

\[ A_{1n} \leq CR_n^{-\frac{(s-2)}{2}} \left( \xi + \hat{\Delta}_{n,s}(\xi) \right) + CR_n^{-\delta}, \]

for any \( \xi \in (0, 1] \), where

\[ \hat{\Delta}_{n,s}(\xi) \equiv \frac{1}{p_n} \sum_{j=1}^{p_n} \|W_{j,n}\|^s \cdot 1\{\|W_{j,n}\| \geq \xi \sqrt{R_n}\}. \]

Note that on the event \( \mathcal{E}_{n,2} \),

\[ \hat{\Delta}_{n,s}(\xi) \leq \left( \xi \sqrt{R_n} \right)^{-a} \frac{1}{p_n} \sum_{j=1}^{p_n} \|W_{j,n}\|^s \leq \left( \xi \sqrt{R_n} \right)^{-a} R_n^\delta. \]

We take \( \xi = R_n^{\left(\frac{\delta-\frac{s}{2}}{1+s}\right)\frac{1}{1+s}} \) to deduce that

\[ A_{1n} \leq 2CR_n^{-\frac{s-2}{2}} R_n^{\frac{2s-a}{1+s}} + CR_n^{-\delta} = 3CR_n^{-\delta}, \]

using the definition that \( a = \frac{4\delta - (s-2)}{(s-1-2\delta)} \) in Assumption 3.1. (By the condition \( 2\delta \in (s-2, s-1) \), we have \( a > 0 \).)

As for \( A_{3n} \), note that by Markov’s inequality,

\[ P\mathcal{E}_{n,2}^c \leq \frac{1}{p_n} \sum_{j=1}^{p_n} E\|W_{j,n}\|^{s+a} \cdot R_n^{-\delta} \leq C_1 R_n^{-\delta} , \]

again by Assumption 3.1. Thus, we find that

\[ E[A_{3n}] \leq C_1 R_n^{-\delta} . \]

This concludes the proof of the theorem. ■
Proof of Corollary 3.1: First, we bound
\[
\sup_{t \geq 0} \left| Q_n(\hat{A}(t)) - Q_n(\hat{A}(t + \eta)) \right|
\leq \sup_{t \geq 0} \left| Q_n(\hat{A}(t)) - \tilde{Q}_n(\hat{A}(t)) \right| + \sup_{t \geq 0} \left| Q_n(\hat{A}(t + \eta)) - \tilde{Q}_n(\hat{A}(t + \eta)) \right|
+ \sup_{t \geq 0} \left| \tilde{Q}_n(\hat{A}(t)) - \tilde{Q}_n(\hat{A}(t + \eta)) \right|
\leq CR_n^{-8} + 2C_1 D_n + \sup_{t \geq 0} \left| \tilde{Q}_n(\hat{A}(t)) - \tilde{Q}_n(\hat{A}(t + \eta)) \right|,
\]
by Theorem 3.2. Let us define the event
\[
\mathcal{E}_{n,3} \equiv \left\{ \lambda_{\min}(\hat{\Omega}) > \varepsilon_1, \text{ and } \frac{1}{p_n} \sum_{i=1}^{p_n} \|W_{i,n}\|^s+2 \leq M_1 \right\}.
\]
On the event \( \mathcal{E}_{n,3} \), \( \tilde{P}_j(-D; \{\tilde{\chi}_\nu\})\phi_{0,\tilde{\chi}_n}(x) \) is Lipschitz continuous in \( x \) with the Lipschitz coefficient depending only on \( d, s, \varepsilon_1, \) and \( M_1 \). Therefore, on the event \( \mathcal{E}_{n,3} \),
\[
\sup_{t \geq 0} \left| \tilde{Q}_n(\hat{A}(t)) - \tilde{Q}_n(\hat{A}(t + \eta)) \right| \leq C' \eta,
\]
for some constant \( C' \) which depends only on \( d, s, \varepsilon_1, \) and \( M_1 \). □

References

Andrews, D. W. K., and X. Shi (2013): “Inference Based on Conditional Moment Inequalities,” Econometrica, 81, 609–666.

Angst, J., and G. Poly (2017): “A Weak Cramér Condition and Application to Edgeworth Expansions,” Electronic Journal of Probability, pp. 1–24.

Bhattacharya, R. N., and R. R. Rao (2010): Normal Approximation and Asymptotic Expansion. SIAM, Philadelphia.

Bickel, P. J., F. Götze, and W. R. van Zwet (1997): “Resampling Fewer Than \( n \) Observations: Gains, Losses, and Remedies for Losses,” Statistica Sinica, 7.

Booth, J. G., P. Hall, and A. T. A. Wood (1994): “On the Validity of Edgeworth and Saddlepoint Approximations,” Journal of Multivariate Analysis, 51.

Giné, E., and J. Zinn (1991): “Gaussian Characterization of Uniform Donsker Classes of Functions,” Annals of Statistics, 19.

Hall, P. (1992): The Bootstrap and Edgeworth Expansion. Springer, New York.

Hall, P., and B.-Y. Jing (1995): “Uniform Coverage Bounds for Confidence Intervals and Berry-Esseen Theorems for Edgeworth Expansion,” Annals of Statistics, 1995, 363 – 375.
KLEINER, A., A. TALWALKER, P. SARKAR, AND M. I. JORDAN (2014): “A scalable bootstrap for massive data,” *Journal of the Royal Statistical Society, B.*, 76, 795–816.

KOLASSA, J. E., AND P. MCCULLAGH (1990): “Edgeworth Series for Lattice Distributions,” *Annals of Statistics*, 18, 981–985.

LINTON, O., K. SONG, AND Y.-J. WHANG (2010): “An Improved Bootstrap Test of Stochastic Dominance,” *Journal of Econometrics*, 154, 186–202.

MIKUSHEVA, A. (2007): “Uniform Inference in Autoregressive Models,” *Econometrica*, 75, 1441–1452.

POLITIS, D. N., AND J. P. ROMANO (1994): “Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions,” *Annals of Statistics*, 22, 2031–2050.

SHEEHY, A., AND J. A. WELLNER (1992): “Uniform Donsker Class of Functions,” *Annals of Statistics*, 20, 1983–2030.

SINGH, K. (1981): “On the Asymptotic Accuracy of Efron’s Bootstrap,” *Annals of Statistics*, pp. 1187–1195.

SONG, K. (2018): “Ordering-Free Inference for Locally Dependent Observations,” *arXiv:1604.00447v2 [stat.ME]*.

VANCOUVER SCHOOL OF ECONOMICS, UNIVERSITY OF BRITISH COLUMBIA, 6000 IONA DRIVE, VANCOUVER, BC, V6T 1L4, CANADA

*E-mail address: kysong@mail.ubc.ca*