Block Bootstrap Consistency Under Weak Assumptions

Gray Calhoun

Working Paper No. 11017
October 2014

IOWA STATE UNIVERSITY
Department of Economics
Ames, Iowa, 50011-1070
Block bootstrap consistency
under weak assumptions

Gray Calhoun∗
6 October, 2014

Abstract
This paper weakens the size and moment conditions needed for typical block bootstrap methods (i.e. the Moving Blocks, Circular Blocks, and Stationary Bootstraps) to be valid for the sample mean of Near-Epoch-Dependent (NED) functions of mixing processes; they are consistent under the weakest conditions that ensure the original NED process obeys a Central Limit Theorem (those of de Jong, 1997, Econometric Theory). In doing so, this paper extends de Jong’s method of proof, a blocking argument, to hold with random and unequal block lengths. This paper also proves that bootstrapped partial sums satisfy a Functional CLT under the same conditions.

JEL Classification: C12, C15

Keywords: Resampling, Time Series, Near Epoch Dependence, Functional Central Limit Theorem

∗Economics Department, Iowa State University, Ames, IA 50011. Telephone: (515) 294-6271. Email: gcalhoun@iastate.edu. Web: http://www.econ.iastate.edu/~gcalhoun. I would like to thank Helle Bunzel, Dimitris Politis, Robert Taylor, and three anonymous referees for their comments and feedback on earlier versions of this paper.
Block bootstraps, e.g. the Moving Blocks (Kunsch, 1989; Liu and Singh, 1992), Circular Block (Politis and Romano, 1992), and Stationary Bootstraps (Politis and Romano, 1994), have become popular in Economics, partly because they do not require the researcher to make parametric assumptions about the data generating process. They are valid under general weak dependence and moment conditions. Some recent papers (Gonçalves and White, 2002; Gonçalves and de Jong, 2003) relax the dependence and moment conditions of the original papers to fit with those commonly used in Econometrics based on Near-Epoch-Dependence (NED). But these conditions are still stronger than required for a CLT to hold; de Jong (1997) has established the CLT under $L_{2-NED}$ with smaller size and moment restrictions. This paper shows that these block bootstrap methods consistently estimate the distribution of the sample mean under de Jong’s (1997) assumptions, and show that an fclt holds as well. It also relaxes Gonçalves and White’s (2002) and Gonçalves and de Jong’s (2003) requirement that the expected block length be $o(n^{1/2})$ to the original papers’ requirement that it be $o(n)$.

The proof exploits the conditional independence of the blocks in each bootstrap. Each bootstrap proceeds by drawing blocks of $M$ consecutive observations from the original time series, and then pasting these blocks together to create the new bootstrap time series. The Moving Blocks bootstrap does exactly that; the Circular Block bootstrap “wraps” the observations, so that $(X_{n-1}, X_n, X_1, X_2)$, for example, is a possible block of length four (letting $X_t$ denote the original time series). The Stationary Bootstrap wraps the observations and also draws $M$ at random for each block; Politis and Romano (1994) suggest drawing $M$ from the geometric distribution. As the name suggests, the series produced by the Stationary Bootstrap are strictly stationary, while those produced by the other methods are not. Although the Stationary Bootstrap was believed to be much less efficient than other block bootstrap methods due to results of Lahiri (1999), Nordman (2009) has shown that it is only slightly less efficient than the other block-bootstrap methods discussed in this paper, and

---

1 Gonçalves and White (2002) show that these bootstrap methods can be applied to heterogeneous $L_{2+\delta\text{-NED}}$ processes of size $-2(r-1)/(r-2)$ on a strong mixing sequence of size $-(2+\delta)/(r-2)$, where $r > 2$ and $\delta > 0$. The original series has uniformly bounded $3r$-moments. Gonçalves and de Jong (2003) relax these conditions to $L_{2+\delta\text{-NED}}$ of size $-1$ and $r + \delta$ moments for the original series, and size $-(2+\delta)(r+\delta)/(r-2)$ for the underlying mixing series. Both papers require that the expected block length grow with $n$ and be $o(n^{1/2})$. Gonçalves and Politis (2011) discuss these issues further.

2 An array $\{X_{nt}\}$ is an $L_\varphi$-NED process on a mixing array $\{V_{nt}\}$ if

$$\|X_{nt} - E(X_{nt} | V_{n,t-m}, \ldots, V_{n,t+m})\|_\varphi \leq d_{nt}v_m$$

with $v_m \to 0$ as $m \to \infty$ and $\{d_{nt}\}$ an array of positive constants. It is of size $-\gamma$ if $v_m = O(m^{-\gamma-\delta})$ for all $\delta > 0$. Dropping the index “$n$” gives the series definition. Note that strong and uniform mixing arrays are not required to be stationary.

3 de Jong (1997) proves that the CLT holds for averages of $L_{2\text{-NED}}$ processes of size $-1/2$ on a strong mixing series of size $-r/(r-2)$, $r > 2$ and the original series having bounded $r$-moments.

4 Radulović (1996) proves consistency for the Moving Blocks Bootstrap for any stationary strong mixing sequence that satisfies the CLT. This paper uses a similar method of proof to his, but also accommodates nonstationary sequences and the Stationary Bootstrap.
has efficiency identical to that of the non-overlapping block bootstrap. Consequently, there has been renewed interest in the Stationary Bootstrap since stationarity of the bootstrap samples can be a useful property in theoretical research. Kreiss and Paparoditis (2011a) provides a recent review of the bootstrap for time-series processes.\(^5\) and Gonçalves and Politis (2011) further discuss recent developments in block-bootstraps.

Theorem 1 presents the main result, asymptotic normality of the distribution of bootstrapped partial sums. This paper adopts the standard notation that \(E^*, \text{var}^*, \) etc. are the usual operators with respect to the probability measure induced by the bootstrap and will use explicit stochastic array notation for precision. Also note that all results are presented for the scalar case but generalize immediately to random vectors. All of the proofs are presented in the appendix; only proofs for the Stationary Bootstrap are presented, since proofs for the other methods are similar and easier to follow. All limits are taken as \(n \to \infty \) unless otherwise noted and \(\|\cdot\|\) denotes the \(L_r\)-norm.

**Theorem 1.** Suppose the following conditions hold.

1. \(X_{nt}\) is \(L_2\)-NED of size \(-1/2\) on an array \(V_{nt}\) that is either strong mixing of size \(-r/(r-2)\) or uniform mixing of size \(-r/2(r-1)\), with \(r > 2\). The NED magnitude indices are denoted \(\{d_{nt}\}\).

2. The array \(\mu_{nt} - \bar{\mu}_n\) is uniformly bounded and \(\sum_{t=1}^{n} (\mu_{nt} - \bar{\mu}_n)^2 \to 0\), where \(E X_{nt} = \mu_{nt}\) and \(\bar{\mu}_n = n^{-1} \sum_{t=1}^{n} \mu_{nt}\). Moreover, \(\sqrt{n} \|X_n - \bar{\mu}_n\|_2 \to \sigma > 0\), with \(\bar{X}_n = (1/n) \sum_{t=1}^{n} X_{nt}\).

3. There exists an array of positive real numbers \(\{c_{nt}\}\) such that \((X_{nt} - \mu_{nt})/c_{nt}\) is uniformly \(L_r\)-bounded and \(c_{nt}\) and \(d_{nt}/c_{nt}\) are uniformly bounded in \(n\) and \(t\).

4. \(X_{nt}^*\) is generated by the Stationary Bootstrap with geometric block lengths with success probability \(p_0, p_n = cn^{-a}\) and \(a, c \in (0, 1)\), or by the Moving or Circular Block bootstrap with block length \(M_n\) such that \(M_n \sim n^a\) for \(a \in (0, 1)\). Let \(M_{ni}\) be the block length of the \(i\)th block, \(i = 1, \ldots, J_n\), and define \(K_{n0} = 0\) and \(K_{nj} = \sum_{i=1}^{j} M_{ni}\). The last block, \(M_{n,J_n}\), is defined as \(K_{n,J_n} - K_{n,J_n-1}\), so \(K_{n,J_n} = n\ a.s.\)

Then \(\hat{\sigma}^2 \to p \sigma^2\),

\[
\sup_x \left| \Pr^* \left[ \sqrt{n} (\bar{X}_n - E* \bar{X}_n^*) \leq x \right] - \Pr \left[ \sqrt{n} (\bar{X}_n - E \bar{X}_n) \leq x \right] \right| \to p 0, \quad (2)
\]

and

\[
\sup_x \left| \Pr^* \left[ \sqrt{n} (\bar{X}_n - E* \bar{X}_n^*)/\hat{\sigma}_n^* \leq x \right] - \Phi(x) \right| \to p 0, \quad (3)
\]

\(^5\)Also see the discussion papers: Dahlhaus (2011); Gonçalves and Politis (2011); Horowitz (2011); Jentsch and Mammen (2011); Kreiss and Paparoditis (2011b).
where Φ is the CDF of the Standard Normal distribution and
\[
\hat{\sigma}^*_n^2 = \frac{1}{n} \sum_{j=1}^{J_n} \left\{ \sum_{t=K_{n,j}-1+1}^{K_{n,j}} (X^*_nt - \bar{X}_n)^2 \right\}. \tag{4}
\]

As mentioned earlier, these are the same size and mixing conditions used by de Jong (1997). Note that de Jong does allow a little bit more flexibility in the conditions on the array \( \{c_{nt}\} \) (see also Davidson, 1993); essentially, he allows for a single set of blocks with the maximal \( \{c_{nt}\} \) over each block well-behaved, while this paper requires this condition to hold for every possible partition of blocks. This additional restriction is required because the Stationary Bootstrap will select the blocks randomly and is similar to de Jong and Davidson’s (2000) requirement for the FCLT. Similarly, our assumption on the dispersion of the individual means, \((\mu_{nt} - \bar{\mu}_n)^2\), is slightly stronger than Gonçalves and White’s (2002) and Gonçalves and de Jong’s (2003) to accommodate larger block sizes.

Theorem 1 relies on a general insight about the variance of the sample mean under the bootstrap-induced distribution. It is well-known that a key step in proving the CLT for arbitrary dependent processes is demonstrating that the squared elements converge to a positive and finite limit; i.e. if \( \{Z_{nj}; j = 1, \ldots, J_n\} \) is a representative stochastic array, \( \sum_j Z_{nj}^2 \rightarrow^p \sigma^2 \) is an important necessary condition for \( \sum_j Z_{nj} \rightarrow^d N(0, \sigma^2) \). (See Section 3.2 of Hall and Heyde, 1980, for further discussion.) For martingale difference arrays, each \( Z_{nj} \) is one of the original random variables \( X_{nt} \) (typically normalized by \( 1/\sqrt{n} \)), but for other forms of dependence (NED or mixingale arrays, for example, as in de Jong, 1997) each \( Z_{nj} \) is a contiguous block of the original random variables,

\[
Z_{nj} = \frac{1}{\sqrt{n}} \sum_{t=(j-1)M_n+1}^{jM_n} (X_{nt} - \mu_{nt})
\]

that adds up to the original summation, plus potentially a negligible residual, so \( \sum_j Z_{nj} = (1/\sqrt{n}) \sum_t (X_{nt} - \mu_{nt}) + o_p(1) \). In de Jong (1997), for example, the CLT for mixingale arrays assumes that there exists a sequence of blocks such that \( \sum_j Z_{nj}^2 \) converges i.p., and the NED CLT establishes conditions under which such a \( Z_{nj} \) exists.

Our insight is that the expectation of squared blocks of the bootstrap process can be expressed as a sequence of contiguous blocks of the original process, so the arguments that establish convergence of the original squared blocks can be applied with only minor changes to the bootstrapped blocks. Consider the Moving Blocks Bootstrap,\(^\text{6}\) for example, and let

\[
Z^*_nj = \frac{1}{\sqrt{n}} \sum_{t=(j-1)M_n+1}^{jM_n} (X^*_nt - \bar{X}_n).
\]

\(^6\)To make this presentation as simple as possible, assume for now that \( n = M_n J_n \) exactly.
Then, conditional on the data, the $Z_{nj}^2$ are independent can be expected to obey an LLN, so $\sum_j (Z_{nj}^2 - E^* Z_{nj}^2) \rightarrow^p 0$ and the CLT for the bootstrapped array requires $\sum_j E^* Z_{nj}^2$ to converge to a positive and finite limit. But, since $E^*$ only averages over the starting point of each block, we have

$$E^* Z_{nj}^2 = \frac{1}{n} \sum_{\tau=0}^{n-M_n} \left( \frac{1}{\sqrt{n}} \sum_{t=\tau+1}^{\tau+M_n} (X_{nt} - \bar{X}_n) \right)^2$$

= $\frac{1}{n} \sum_{\tau_0=0}^{M_n-1} \sum_{j=1}^{J_n-1} \left( \frac{1}{\sqrt{n}} \sum_{t=(j-1)M_n+\tau_0+1}^{jM_n+\tau_0} (X_{nt} - \bar{X}_n) \right)^2$

after grouping blocks separated by $M_n$ periods. For each $\tau_0$, the summation can be expected to converge in probability through the same arguments that were used to establish the CLT for the original array.\(^7\)\(^8\) A similar representation is available for the Circular and Stationary bootstraps.

In short, the basic approach that we use to prove Theorem 1 is based on a fundamental connection between the second moments of the bootstrap process and the sum of squared blocks of the original array. Even though the details of our proof rely on specific techniques for NED arrays, this connection implies that block bootstraps are typically consistent when the original dependent array obeys the CLT, and the connection should be useful for proving consistency of the bootstrap under other dependence conditions.

For some uses, a CLT is inadequate and an FCLT is necessary. Theorem 1 can be extended to give an FCLT (using arguments from de Jong and Davidson, 2000). We show in Theorem 2 that the partial sum of the bootstrapped process obeys an FCLT and can be used to derive critical values for other test statistics under the same assumptions as Theorem 1.

**Theorem 2.** Suppose that the conditions of Theorem 1 hold, let $W$ be standard Brownian Motion, and define

$$W_n^*(\gamma) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[\gamma n]} (X_{nt}^* - \mu_n^*) / \sigma_n^*$$

and

$$\hat{W}_n^*(\gamma) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[\gamma n]} (X_{nt}^* - \hat{\mu}_n^*) / \hat{\sigma}_n^*.$$  

Then $Pr^*[d(W_n^*, W) > \delta] \rightarrow 0$ i.p. and $Pr^*[d(\hat{W}_n^*, W) > \delta] \rightarrow 0$ i.p. for any positive $\delta$ and distance $d$ that metricizes weak convergence.

\(^7\) Some of the details of the argument will typically need to change, since the original CLT only requires convergence for $\tau_0 = 0$, but these details are often incidental to the original argument.

\(^8\) Lemmas 3 and 5 are particularly strong demonstrations of this argument.
Obviously, Theorem 2, has natural corollaries that allow it to be used to approximate the behavior of partial sums of the original series. But applying this result requires stronger assumptions than Theorem 1. The bootstrapped processes are (essentially) homoskedastic, so they are unable to match patterns of heteroskedasticity in the original series and we can not estimate the covariance process of the original partial sum in the bootstrap. This necessitates a separate consistent estimator of the covariance process, which we did not need for the sample mean. Other methods, such as the Local Block Bootstrap (Dowla et al., 2003; Paparoditis and Politis, 2002), may be able to capture this additional heterogeneity with the bootstrap alone, but we do not pursue that possibility further.

**Corollary 1.** Suppose the conditions of Theorem 1 hold and that \( \sup_{t=1,...,n} |\mu_{nt} - \bar{\mu}_n| = o(1/\sqrt{n}) \). Define

\[
\sigma_n^2(\gamma) = \frac{1}{n} \sum_{s,t=1}^{[\gamma n]} \text{cov}(X_{ns}, X_{nt}),
\]

let \( \hat{\sigma}_n^2(\cdot) \) be an estimator of \( \sigma_n^2(\cdot) \) such that \( \sup_{\gamma} |\sigma_n^2(\gamma)/\hat{\sigma}_n^2(\gamma) - 1| \to^p 0 \), and let

\[
\hat{W}_n(\gamma) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[\gamma n]} (X_{nt} - \bar{\mu}_n)/\hat{\sigma}_n(\gamma).
\]

Then \( \Pr^*[d(W^*_n, \hat{W}_n) > \delta] \to^p 0 \) and \( \Pr^*[d(\hat{W}^*_n, \hat{W}_n) > \delta] \to^p 0 \) for any positive \( \delta \) and any distance \( d \) that metricizes weak convergence.

The rest of the paper presents the mathematical proofs in detail.

**A Proof of main results**

For both results, we only present a proof for the stationary bootstrap. The moving blocks and circular block bootstrap follow the same general argument but are simpler. First we will define some notation that will be used repeatedly in our results.

Define the \( \sigma \)-field

\[
\mathcal{M}_n = \sigma(J_n, M_{n1}, \ldots, M_{nJ_n})
\]

and the conditional probability \( \Pr_{\mathcal{M}}(\cdot) = \Pr[\cdot \mid \mathcal{M}_n] \). (And define \( \Pr^*_\mathcal{M}(\cdot) = \Pr^*[\cdot \mid \mathcal{M}_n] \), \( E_{\mathcal{M}}(\cdot) = E^*(\cdot \mid \mathcal{M}_n) \), etc.) An important property is that \( \mathcal{M}_n \) is independent of the \( X_{nt} \), so we can treat any \( M_{nj} \) and \( J_n \) terms as constants within \( E_{\mathcal{M}}(\cdot) \) and \( \Pr_{\mathcal{M}}(\cdot) \) and integrate over the unconditional distributions of the \( X_{nt} \). This property is especially important because it allows us to use maximal inequalities and other moment inequalities for mixingale arrays without modification to construct almost sure bounds on the conditional moments \( E_{\mathcal{M}} \).
Also define

\[
I_n(\tau, m) = \begin{cases} 
\{\tau + 1, \ldots, \tau + m\} & \text{if } 0 \leq \tau \leq n - m \text{ and } 1 \leq m \\
\{1, \ldots, m - n + \tau\} \cup \{\tau + 1, \ldots, n\} & \text{if } n - m < \tau \leq n \text{ and } 1 \leq m \\
\emptyset & \text{if } m \leq 0,
\end{cases}
\]

so each \(I_n(\tau, m)\) defines a potential block of length \(m\) of the original observations that could be chosen by the bootstrap.\(^9\) By convention, summations over empty index sets will be considered equal to zero. Also let

\[
Z_n(\tau, m) = \frac{1}{\sqrt{n}} \sum_{t \in I_n(\tau, m)} (X_{nt} - \bar{X}_n) 
\]

(11)

\[
Z_n^*(\tau, m) = \frac{1}{\sqrt{n}} \sum_{t \in I_n(\tau, m)} (X_{nt}^* - \bar{X}_n)
\]

(12)

and

\[
Z_{nj}^* = Z_n^*(K_{nj-1}, M_{nj}),
\]

(13)

and define the corresponding demeaned terms

\[
Z_n'(\tau, m) = \frac{1}{\sqrt{n}} \sum_{t \in I_n(\tau, m)} (X_{nt} - \mu_{nt})
\]

(14)

\[
Z_n^*'(\tau, m) = \frac{1}{\sqrt{n}} \sum_{t \in I_n(\tau, m)} (X_{nt}^* - \mu_{nt}^*)
\]

(15)

and

\[
Z_{nj}^*' = Z_n^*'(K_{nj-1}, M_{nj}),
\]

(16)

where \(\mu_{nt}^*\) is the expected value of the observation in the original dataset corresponding to the \(t\)th observation in the bootstrapped dataset. Further, define the filtration

\[
\mathcal{G}_{nj} = \sigma(Z_{n1}^*, \ldots, Z_{nj}^*, X_{n1}, \ldots, X_{nn}, \mathcal{M}_n)
\]

(17)

so that \(\{Z_{nj}^*/\sigma_{nj}, \mathcal{G}_{nj}\}\) is a martingale difference array.

By construction,

\[
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (X_{nt}^* - \bar{X}_n) = \sum_{j=1}^{J_n} Z_{nj}^*
\]

(18)

\(^9\)The index sets \(I_n(\tau, m)\) are designed to “wrap around” and use the first observations when \(\tau + m > n\), matching the defining aspect of the stationary and circular block bootstraps.
and

\[ E_M^* g(Z_{nj}^*, \ldots, Z_{nk}^*) = \frac{1}{n^{k-j+1}} \sum_{\tau_1=0}^{n-1} \cdots \sum_{\tau_{k-j+1}=0}^{n-1} g(Z_n(\tau_1, M_{nj}), \ldots, Z_n(\tau_{k-j+1}, M_{nk})) \]  (19)

almost surely for any function \( g \) and any \( j \leq k \). Equation (19) conditions on the lengths of each block, but averages over their starting points.

**Proof of Theorem 1**

First we prove that

\[ \sup_x \left| \Pr^* \left[ \frac{\sqrt{n}(\bar{X}_n^* - \bar{X}_n)}{\sigma_n^*} \leq x \right] - \Phi(x) \right| \rightarrow_p 0 \]  (20)

where \( \sigma_n^2 = n \, E^*(\bar{X}_n^* - \bar{X}_n)^2 \). Rewrite \( \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \) as in Equation (18), so

\[ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (X_{nt}^* - \bar{X}_n) = \sum_{j=1}^{J_n} Z_{nj}^* \]

and \( \{Z_{nj}^*/\sigma_n^*, G_{nj}\} \) is a martingale difference array. Moreover,

\[ \Pr_M^* \left[ \sum_{j=1}^{J_n} Z_{nj}^*/\sigma_n^* \leq x \right] - \Phi(x) \rightarrow_p 0, \]  (21)

for all \( x \) if \( \sigma_n^2 \rightarrow_p \sigma^2 \) (which ensures that \( \sigma_n^2 \) is uniformly a.s. positive and holds by Lemma 3) and the following two conditions hold for all positive \( \epsilon \):

\[ \sum_{j=1}^{J_n} E_M^*(Z_{nj}^* 1\{Z_{nj}^* > \epsilon\}) \rightarrow_p 0 \]  (22)

and

\[ \Pr_M^* \left[ \left| \sum_{j=1}^{J_n} Z_{nj}^2 - \sigma_n^2 \right| > \epsilon \right] \rightarrow_p 0 \]  (23)

since (22) and (23) ensure that \( Z_{nj}^*/\sigma_n^* \) obeys a martingale difference CLT (e.g. Hall and Heyde, 1980, Theorem 3.3).\(^{10}\)

\(^{10}\)Conditional on \( X_{n1}, \ldots, X_{nn}, J_n, \) and \( M_{n1}, \ldots, M_{nk} \), the only stochastic components of \( \sum_{j=1}^{J_n} Z_{nj}^*/\sigma_n \) are the start periods of each block, which are discrete uniform(1, \ldots, \( n \)) and are independent of all of the other random variables in the information set used for conditioning. Consequently, \( \Pr_M^* \) is a regular conditional probability and arguments like Hall and Heyde’s (1980) Theorem 3.3 apply without modification on this probability measure. See also Section 23.2 of Van der Vaart (2000). Also note that Hall and Heyde’s Theorem 3.3 as stated imposes an additional restriction on the sigma-fields. However, as Hall and Heyde discuss on pages 59 and 63–64, that condition is unnecessary here because the \( \sigma_n^* \) is measurable with respect to all of the \( G_{nj} \).
For (23), since $Z^*_{nj}$ and $Z^*_{nk}$ ($k \neq j$) are uncorrelated conditional on $X_{n1}, \ldots, X_{nn}$, and $M_n$, we have
\[
\sum_{j=1}^{J_n} Z^*_{nj} \sigma_n^2 = \sum_{j=1}^{J_n} \left( Z^*_{nj} - (1/J_n) E^* \sum_{j=1}^{J_n} Z^*_{nj} \right)
\]
almost surely. But
\[
\left\{ \frac{n}{M_{nj}} \left( Z^*_{nj} - (1/J_n) E^* \sum_{j=1}^{J_n} Z^*_{nj} \right), G_{nj} \right\}
\]
is a uniformly-integrable martingale difference array by Lemma 5, and satisfies the LLN (e.g., Davidson, 1994, Theorem 19.7), so this sum converges to zero in conditional probability, proving (23).

To prove (22), it suffices to show that
\[
E_M \sum_{j=1}^{J_n} E^*_M \left( Z^*_{nj}^2 1\{Z^*_{nj}^2 > \epsilon \} \right) \to^p 0.
\]
(24)

Lemma 5 implies that there exists a finite, monotone function $B(\cdot)$ such that $B(x) \to 0$ as $x \to \infty$ and
\[
E_M((nZ^*_{nj}^2/M_{nj}) 1\{nZ^*_{nj}^2/M_{nj} > \epsilon n/M_{nj}\}) \leq B(\epsilon n/M_{nj})
\]
almost surely for large enough $n$ and all $j$ and $\tau$. Consequently,
\[
E_M \sum_{j=1}^{J_n} E^*_M(1\{Z^*_{nj}^2 > \epsilon \}) \leq \sum_{j=1}^{J_n} (M_{nj}/n) B(\epsilon n/M_{nj})
\]
\[
\leq \max_{j=1, \ldots, J_n} B(\epsilon n/M_{nj}) \sum_{i=1}^{J_n} M_{ni}/n
\]
\[
= B(\epsilon n/ \max_{j=1, \ldots, J_n} M_{nj})
\]
\[
\to^p 0 \text{ as } n \to \infty,
\]
where the equality holds by monotonicity of $B$ and convergence in probability holds by Lemma 1, completing the proof of (22). The Dominated Convergence Theorem then ensures that
\[
\Pr^* \left[ \sum_{j=1}^{J_n} Z^*_{nj}/\sigma_n^* \leq x \right] - \Phi(x) \to^p 0.
\]
(25)

follows from (21). (Also see Lemma 2.)

Lemma 3 implies that $\sigma_n^*$ and $\hat{\sigma}_n^*$ both converge to $\sigma^2$ in probability. This convergence then implies that
\[
\Pr^* \left[ \sqrt{n} (\bar{X}_n^* - \bar{X}_n) \leq x \right] \to^p \Phi(x/\sigma)
\]
(26)
and
\[ \Pr^* \left[ \sqrt{n}(\bar{X}_n^* - \bar{X}_n) / \hat{\sigma}_n^* \leq x \right] \to^p \Phi(x) \]  
(27)
for any \( x \). These results are sufficient for (2) and (3) though an argument attributed to Polyà that proceeds as follows. Let \( k \) be a finite integer and define \( x_i = \sigma \Phi^{-1}(i/k) \) for \( i = 0, \ldots, k \) (so \( x_0 = -\infty \) and \( x_k = +\infty \)). For any \( x \in [x_i, x_{i+1}] \),
\[ \Pr^* \left[ \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x \right] - \Phi(x/\sigma) \leq \Pr^* \left[ \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x_{i+1} \right] - \Phi(x_i/\sigma) \]
and
\[ \Pr^* \left[ \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x \right] - \Phi(x/\sigma) \geq \Pr^* \left[ \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x_i \right] - \Phi(x_{i+1}/\sigma) \]
almost surely. Then
\[ \sup_{x \in (-\infty, +\infty)} |\Pr^* \left[ \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x \right] - \Phi(x/\sigma)| \leq \sup_{i=0, \ldots, k} |\Pr^* \left[ \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x_i \right] - \Phi(x_i/\sigma)| + 1/k \]
almost surely, and (26) ensures that
\[ \sup_{i=0, \ldots, k} |\Pr^* \left[ \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x_i \right] - \Phi(x_i/\sigma)| + 1/k \to^p 1/k \]
for any finite \( k \). Since \( k \) is arbitrary, (20) holds. Since Theorem 1’s assumptions ensure that the original array obeys the CLT, (2) holds (de Jong, 1997, Theorem 2).

A similar argument applies to the asymptotic distribution of \( \sqrt{n}(\bar{X}_n^* - \bar{X}_n) / \hat{\sigma}_n^* \), completing the proof. \( \square \)

**Proof of Theorem 2**

Define \( W_n^{\gamma*} \) as
\[ W_n^{\gamma*}(\gamma) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor \gamma n \rfloor} (X_{nt}^* - \bar{X}_n) / \sigma = \sum_{j = 1}^{\lfloor \gamma J_n \rfloor} Z_{nj}^*/\sigma + Z_n^*(K_{n,\lfloor \gamma J_n \rfloor}, [\gamma n] - K_{n,\lfloor \gamma J_n \rfloor}) / \sigma. \]

Lemma 3 ensures that \( \sigma_n^2 \) and \( \hat{\sigma}_n^2 \) both converge in probability to \( \sigma^2 \), so it suffices to prove that \( \Pr^*[d(W_n^{\gamma*}, W) > \delta] \to^p 0 \); we can also assume without loss of generality that \( \sigma^2 = 1 \). Moreover, Lemma 2 ensures that it suffices to prove unconditional convergence, so we will establish \( \Pr[d(W_n^{\gamma*}, W) > \delta] \to 0 \).
Theorem 1 implies that, for any fixed $\gamma$, $W_n^t(\gamma)$ is asymptotically normal and, as a partial sum of an MDS, $W_n^t$ has asymptotically independent increments. To see this, choose $\gamma, \gamma' \in [0, 1]$ and $\delta, \delta' > 0$ so that $\delta + \gamma \leq \gamma'$. Since the increments of $W_n^t$ are uncorrelated conditional on $X_n, \ldots, X_{n-1}$, and $M_n$, we have

$$E[(W_n^t(\delta' + \gamma') - W_n^t(\gamma'))(W_n^t(\delta + \gamma) - W_n^t(\gamma))] = 0$$

for large enough $n$ if $\gamma' > \gamma + \delta$ and, if $\gamma' = \gamma + \delta$,

$$E[(W_n^t(\delta' + \gamma') - W_n^t(\gamma'))(W_n^t(\delta + \gamma) - W_n^t(\gamma))] = E_{\delta M} \{- Z_n^*([\gamma' n], K_n, \gamma J_n) - Z_n^*(K_n, \gamma J_n) \} = 0
$$

But this second quantity can be bounded:

$$E_{\delta M} \{- Z_n^*([\gamma' n], K_n, \gamma J_n) - Z_n^*(K_n, \gamma J_n) \} \leq \{ E_{\delta M} [Z_n^*([\gamma' n], K_n, \gamma J_n) - Z_n^*(K_n, \gamma J_n)] \}^2 \times \{ E_{\delta M} Z_n^*(K_n, \gamma J_n) - Z_n^*(K_n, \gamma J_n) \}^{1/2} \leq C M_n, \gamma J_n / n$$

for some constant $C$ by Lemma 5. This term converges to zero by Lemma 1.

Consequently, as in de Jong and Davidson (2000), the result follows from stochastic equicontinuity (Billingsley, 1968, Theorems 15.4 and 15.5) namely that

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \Pr[\sup_{\gamma \in [0, 1]} \sup_{\gamma' \in [\gamma - \delta, \gamma + \delta]} |W_n^t(\gamma) - W_n^t(\gamma')| > \varepsilon] = 0$$

for any positive $\varepsilon$. Fix $\delta > 0$ such that $D = 2/\delta$ is a positive integer and let $\gamma_d = d/D$ for $d = 0, \ldots, D$. Then, mimicking the argument in de Jong and Davidson (2000),

$$\Pr[\sup_{\gamma \in [0, 1]} \sup_{\gamma' \in [\gamma - \delta, \gamma + \delta]} |W_n^t(\gamma) - W_n^t(\gamma')| > \varepsilon] \leq \Pr[\sup_{d = 1, \ldots, D} \sup_{\gamma \in [0, \delta]} |W_n^t(\gamma + \gamma_d) - W_n^t(\gamma_d)| > \varepsilon/2] \leq (4/\varepsilon^2) \sum_{d = 1}^D E(\sup_{\gamma \in [0, \delta]} |W_n^t(\gamma + \gamma_d) - W_n^t(\gamma_d)|^2 \times 1\{|W_n^t(\gamma + \gamma_d) - W_n^t(\gamma_d)|^2 > \varepsilon^2/4\}) \leq (4/\varepsilon^2) \max_{d = 1, \ldots, D} E(\sup_{\gamma \in [0, \delta]} |W_n^t(\gamma + \gamma_d) - W_n^t(\gamma_d)|^2 \times 1\{|W_n^t(\gamma + \gamma_d) - W_n^t(\gamma_d)|^2 > \varepsilon^2/4\})}.$$
Lemma 5 implies that

\[ \sup_{\gamma \in [0, \delta]} (1/\delta)|W_n'(\gamma + \gamma d) - W_n'(\gamma d)|^2 \]

is uniformly integrable, so there exists a finite and positive function \( B \) such that \( B(x) \to 0 \) as \( x \to \infty \) and

\[
E[ \sup_{\gamma \in [0, \delta]} (1/\delta)|W_n'(\gamma + \gamma d) - W_n'(\gamma d)|^2 1\{ \sup_{\gamma \in [0, \delta]} (1/\delta)|W_n'(\gamma + \gamma d) - W_n'(\gamma d)|^2 > x \}] \leq B(x)
\]

for all \( d \) and \( \delta \) and all large enough \( n \).

As a result,

\[
\lim_{\delta \to 0} \limsup_{n \to \infty} \Pr[ \sup_{\gamma \in [0,1]} \sup_{\gamma' \in [\gamma-\delta, \gamma+\delta]} |W_n'(\gamma) - W_n'(\gamma')| > \epsilon ] \leq \lim_{\delta \to 0} \frac{4/\epsilon^2}{B(\epsilon^2/4\delta)} = 0,
\]

completing the proof.

B Supplemental results

**Lemma 1.** Suppose that \( M_{n1}, M_{n2}, \ldots \) are i.i.d. geometric random variables with success parameter \( p_n = cn^{-a} \) with \( a, c \in (0, 1) \), and that \( \ell_n = (p_n \log p_n^{-1})^{-1} \) and define \( J_n \) so that

1. \( \max_{i=1, \ldots, J_n} M_{ni}/n \to 0 \) for any positive \( C \),
2. \( \max_{i=1, \ldots, J_n} M_{ni}/n \to 0 \),
3. \( \max_{i=1, \ldots, J_n} M_{ni}/\ell_n^{1+\epsilon} \to 0 \) as \( n \to \infty \) for any positive \( \epsilon \) and \( C \),
4. \( \max_{i=1, \ldots, J_n} M_{ni}/\ell_n^{1+\epsilon} \to 0 \) as \( n \to \infty \) for any positive \( \epsilon \), and
5. \( \sum_{i=1}^{J_n} M_{ni}^2/n^2 \to 0 \).

**Proof of Lemma 1.** To prove part 1, observe that, for any increasing positive sequence \( x_n \) such that \( x_np_n \to \infty \),

\[
\Pr \left[ \max_{i=1, \ldots, [Cnp_n]} M_{ni} \leq x_n \right] = (1 - (1 - p_n)^{x_n})^{[Cnp_n]} \to \lim \exp(-Cnp_n(1 - p_n)^{x_n})
\]

and \( Cnp_n(1 - p_n)^{x_n} \to \lim Cnp_ne^{-x_np_n} \). Now, let \( x_n = nx \) for any positive number \( x \). Then

\[
\Pr \left[ \max_{i=1, \ldots, [Cnp_n]} M_{ni}/n \leq x \right] \to \lim \exp(-Cnp_ne^{-np_nx}) = \exp(0) = 1.
\]
Since \( x \) is arbitrary, \( \max_{i=1,\ldots,J_n} M_{ni}/n \to^p 0 \).

For part 2, take \( C \) to be an arbitrary constant strictly greater than one. For any \( x \),
\[
\Pr \left[ \max_{i=1,\ldots,J_n} M_{ni} > x \right] \leq \Pr \left[ \max_{i=1,\ldots,[Cnp_n]} M_{ni} > x \text{ or } J_n > [Cnp_n] \right]
\]
\[
\leq \Pr \left[ \max_{i=1,\ldots,[Cnp_n]} M_{ni} > x \right] + \Pr \left[ \sum_{i=1}^{[Cnp_n]} M_{ni} < n \right]
\]
The first term converges to zero by part 1 and the second term by the LLN.

For part 3, let \( x_n = \ell_n^{1+\epsilon} x \) and note that
\[
\Pr^* \left[ \sigma^2_n > \ell_n^{1+\epsilon} x \right] \leq \Pr^* \left[ \left| \sigma^2_n - \sigma^2 \right| > \epsilon \right] \to^p 1
\]
as well.

The proof of part 4 is the same as part 2, making the obvious substitutions. Part 5 holds because \( \sum_{i=1}^{J_n} M_{ni}^2/n^2 \leq \max_{i=1,\ldots,J_n} M_{ni}/n \) which converges to zero in probability by part 2.

\begin{proof}
\end{proof}

\textbf{Lemma 2.} If \( \{A_n\} \) is a sequence of events then the following are equivalent:
\[
\Pr[A_n] \to 0, \quad \Pr^*[A_n] \to 0 \text{ in } L_1, \quad \text{and} \quad \Pr^*_M[A_n] \to 0 \text{ in } L_1.
\]

\textbf{Proof.} Since \( |\Pr[A_n]| = \E|\Pr^*[A_n]| = \E|\Pr^*_M[A_n]| \) these conditions are equivalent by definition.

\textbf{Lemma 3.} Under the conditions of Theorem 1,
\[
\sum_{j=1}^{J_n} \left( Z_{nj}^* - \E^*_M Z_{nj}^* \right) \to^p 0, \tag{28}
\]
\[
\sum_{j=1}^{J_n} \left( Z_{nj}^2 - Z_{nj}^* Z_{nj}^* \right) \to^p 0, \tag{29}
\]
and \( \Pr^*[|\sigma^2_n - \sigma^2| > \epsilon] \to^p 0 \). If, in addition, \( \bar{X}^* - \bar{X}_n = O_p(1/\sqrt{n}) \) then \( \Pr^*[|\hat{\sigma}^2_n - \sigma^2| > \epsilon] \to^p 0 \).
Proof. For (29), \((Z_{nj}^* - E_M Z_{nj}^*) \cdot (n/M_{nj})\) is a uniformly integrable martingale difference array, by Lemma 5, and satisfies the LLN. (See Davidson, 1994, Theorem 19.7.) For (29), observe that

\[
\left| \sum_{j=1}^{J_n} (Z_{nj}^* - Z_{nj}^*) \right| \leq 2 \left( \sum_{j=1}^{J_n} Z_{nj}^* \right)^{1/2} \left( \frac{1}{n} \sum_{t=1}^{n} (\mu_{nt} - \bar{X}_n)^2 \right)^{1/2} + \frac{1}{n} \sum_{t=1}^{n} (\mu_{nt} - \bar{X}_n)^2
\]

after several applications of the Cauchy-Schwarz inequality. Lemma 4, along with (19) and (28), implies that \(\sum_{j} Z_{nj}^2 \sim O_p(1)\); \((1/n) \sum_{t} (\mu_{nt} - \bar{X})^2\) converges to zero in probability by assumption on \(\mu_{nt} - \bar{\mu}_n\) and because \(\bar{X}\) itself obeys the LLN.

To show that \(\sigma_n^2\) converges, we can write

\[
\sigma_n^2 - \sigma^2 = E^* \left\{ \sum_{j=1}^{J_n} (Z_{nj}^* - Z_{nj}^*) \right\} + \frac{1}{n} \sum_{\tau=0}^{n-1} E^* \left\{ \sum_{j=1}^{J_n} (Z_{n\tau}^*(\tau, M_{nj})^2 - E_M Z_{n\tau}^*(\tau, M_{nj})^2) \right\}
\]

\[
+ \frac{1}{n} \sum_{\tau=0}^{n-1} E^* \left\{ \sum_{j=1}^{J_n} E_M Z_{n\tau}^*(\tau, M_{nj})^2 - \sigma^2 \right\}.
\]

Uniform integrability ensures that the convergence in (28) holds in \(L_1\) as well and Lemma 2 then implies that the first term converges to zero in probability. Lemma 4 proves that the second and third summation converge to zero in probability as well.

Next,

\[
\hat{\sigma}_n^2 - \sigma^2 = \sum_{j=1}^{J_n} \left( Z_{nj}^* + \left( M_{nj}/\sqrt{n} \right)(\bar{X}_n - \bar{X}_n^*) \right) - E^* \sum_{j=1}^{J_n} Z_{nj}^2
\]

so, in light of the previous arguments, \(\hat{\sigma}_n^2 \rightarrow^p \sigma^2\) if

\[
(\bar{X}_n - \bar{X}_n^*)^2 \sum_{j=1}^{J_n} M_{nj}^2/n \rightarrow^p 0,
\]

which holds by Lemma 1 and assumption. 

**Lemma 4.** If the conditions of Theorem 1 hold then

\[
Pr \left[ \left| \frac{1}{n} \sum_{\tau=0}^{n-1} \sum_{j=1}^{J_n} \left[ Z_{n\tau}^*(\tau, M_{nj})^2 - E_M Z_{n\tau}^*(\tau, M_{nj})^2 \right] \right| > \epsilon \right] \rightarrow 0,
\]

and

\[
Pr \left[ \left| \frac{1}{n} \sum_{\tau=0}^{n-1} \sum_{j=1}^{J_n} E_M Z_{n\tau}^*(\tau, M_{nj})^2 - \sigma^2 \right| > \epsilon \right] \rightarrow 0.
\]
For these two proofs, let $\ell_n = (p_n \log p_n^{-1})^{-1}$ and let $L_{nj} = \lfloor n/M_{nj} \rfloor$; $\ell_n$ represents a smaller block size that satisfies $\ell_n J_n/n \to^p 0$.

Proof of (31). We can express this summation as

$$\frac{1}{n} \sum_{\tau=0}^{n-1} \sum_{j=1}^{J_n} \left\{ Z_n'(\tau, M_{nj})^2 - E_M Z_n'(\tau, M_{nj}) \right\}$$

$$= \frac{1}{n} \sum_{j=1}^{J_n} \sum_{\tau=0}^{M_{nj}-1} \sum_{i=0}^{L_{nj}-1} \left\{ [Z_n'(\tau + iM_{nj} - \ell_n) + Z_n'(\tau + (i+1)M_{nj} - \ell_n, \ell_n)]^2 \right.$$

$$- E_M \left[ Z_n'(\tau + iM_{nj} - \ell_n) + Z_n'(\tau + (i+1)M_{nj} - \ell_n, \ell_n) \right]^2 \}

$$+ \frac{1}{n} \sum_{j=1}^{J_n} \sum_{\tau=M_{nj}L_{nj}}^{n-1} \left\{ Z_n'(\tau, M_{nj})^2 - E_M Z_n'(\tau, M_{nj})^2 \right\} \quad (33)$$

almost surely.

By Lemma 6 (Equation 37), for any $\delta > 0$ there exist positive $C$ and $\epsilon$ such that

$$\left\| \frac{1}{n} \sum_{j=1}^{J_n} \sum_{\tau=0}^{M_{nj}-1} \sum_{i=0}^{L_{nj}-1} \left\{ Z_n'(\tau + iM_{nj} - \ell_n)^2 - E_M (Z_n'(\tau + iM_{nj} - \ell_n)^2) \right\} \right\|_1$$

$$\leq E \left\| \frac{1}{n} \sum_{j=1}^{J_n} \sum_{\tau=0}^{M_{nj}-1} E_M \sum_{i=0}^{L_{nj}-1} \left\{ Z_n'(\tau + iM_{nj} - \ell_n)^2 - E_M (Z_n'(\tau + iM_{nj} - \ell_n)^2) \right\} \right\|$$

$$\leq E \left( \frac{1}{n} \sum_{j=1}^{J_n} \sum_{\tau=0}^{M_{nj}-1} (2\delta + C \frac{M_{nj}}{n^{1/2}\tau^{1/2}}) \right)$$

for large enough $n$, which converges to $2\delta$ by Lemma 1. Also, Lemma 7 ensures that there exists a value $C$ (possibly different from the value above) such that

$$E \left( \frac{1}{n} \sum_{j=1}^{J_n} \sum_{\tau=0}^{M_{nj}-1} \sum_{i=0}^{L_{nj}-1} Z_n'(\tau + (i+1)M_{nj} - \ell_n, \ell_n)^2 \right) \leq C E \left( \sum_{j=1}^{J_n} L_{nj} \ell_n M_{nj}/n^2 \right)$$

and

$$E \left( \frac{1}{n} \sum_{j=1}^{J_n} \sum_{\tau=M_{nj}L_{nj}}^{n-1} Z_n'(\tau, M_{nj})^2 \right) \leq C E \left( \sum_{j=1}^{J_n} M_{nj}^2/n^2 \right)$$

for large enough $n$, both of which converge to zero in $L_1$ as $n \to \infty$ by Lemma 1. These three convergence results imply that the RHS of (33) converges to zero, completing the proof. \qed
Proof of (32). After using similar arguments to in the previous part of the proof, the conclusion holds if

\[ \frac{1}{n} \sum_{j=1}^{J_n} \sum_{\tau=0}^{M_{nj}-1} \sum_{i=0}^{\ell_n} E_M Z_n^\tau (\tau + iM_{nj}, M_{nj} - \ell_n)^2 \rightarrow^p \sigma^2, \]

which is a direct implication of Lemma 6.

Lemma 5. Under the conditions of Theorem 1,

\[
\lim_{C \to 0} \limsup_{n \to \infty} \sup_{m'=1,\ldots,n} E \left( \left( \sum_{\tau=0}^{n-1} \max_{m=1,\ldots,m'} Z_n(\tau, m)^2/m' \right) \right) \\
\times 1 \left\{ \sum_{\tau=0}^{n-1} \max_{m=1,\ldots,m'} Z_n(\tau, m)^2/m' > C \right\} = 0. \quad (34)
\]

and

\[
\lim_{C \to \infty} \limsup_{n \to \infty} \sup_{m'=1,\ldots,n} E \left( \max_{m=1,\ldots,m'} (n Z_n^*(\tau, m)^2/m') \right) \\
\times 1 \left\{ \max_{m=1,\ldots,m'} n Z_n^*(\tau, m)^2/m' > C \right\} = 0, \quad (35)
\]

making the families of random variables \( \{ \max_{m=1,\ldots,m'} Z_n(\tau, m)^2/m' \} \) and \( \{ \max_{m=1,\ldots,m'} Z_n^*(\tau, m)^2/m' \} \) uniformly integrable.

Proof. For (34), take \( \epsilon > 0 \). The sum

\[ \sum_{\tau=0}^{n-1} \max_{m=1,\ldots,m'} Z_n(\tau, m)^2/m' \]

is uniformly integrable if and only if it is uniformly \( L_1 \)-bounded and there exists an \( \epsilon' \) such that

\[ \sup_{n,m} E \left( 1(A) \sum_{\tau=0}^{n-1} \max_{m=1,\ldots,m'} Z_n(\tau, m)^2/m' \right) < \epsilon \]

for every event \( A \) satisfying \( \Pr(A) < \epsilon' \). (Davidson, 1994, 12.9.) Lemma 7 ensures that there is an \( \epsilon' \) such that

\[ \sup_{n,m',\tau} E \left( 1(A) \max_{m=1,\ldots,m'} Z_n^*(\tau, m)^2/m' \right) \leq \epsilon/9 \]
for all $A$ with $\Pr(A) \leq \epsilon'$. Then, writing

$$Z_n(\tau, m) = Z'_n(\tau, m) + \frac{1}{\sqrt{n}} \sum_{t \in I_n(\tau, m)} ((\mu_{nt} - \bar{\mu}_n) + (\bar{\mu}_n - \bar{X}_n))$$

$$= Z'_n(\tau, m) + \frac{1}{\sqrt{n}} \sum_{t \in I_n(\tau, m)} (\mu_{nt} - \bar{\mu}_n) - mZ'_n(0, n)$$


gives the relations

$$E\left( 1(A) \sum_{\tau=0}^{n-1} \max_{m=1,\ldots,m'} Z_n(\tau, m)^2 / m' \right)$$

$$\leq 3 \sum_{\tau=0}^{n-1} E\left( 1(A) \max_{m=1,\ldots,m'} Z'_n(\tau, m)^2 / m' \right)$$

$$+ (3/nm') E\left( 1(A) \sum_{\tau=0}^{n-1} \max_{m=1,\ldots,m'} \left( \sum_{t \in I_n(\tau, m)} (\mu_{nt} - \bar{\mu}_n)^2 \right) \right)$$

$$+ (3/m') \sum_{\tau=0}^{n-1} E(1(A) \max_{m=1,\ldots,m'} mZ'_n(0, n)^2)$$

$$\leq (2\epsilon/3) + \Pr(A)(3/n) \sum_{t=1}^{n} (\mu_{nt} - \bar{\mu}_n)^2$$

where the last inequality holds as a consequence of $A$’s construction. Since $(1/n) \sum_{t=1}^{n} (\mu_{nt} - \bar{\mu}_n)^2 \rightarrow 0$, choose $n$ large enough that this average is less than $\epsilon/9$. The proof that that the first moment is bounded is similar, which completes the proof of (34).

For (35), we’ll first show that, for every $\epsilon > 0$, there exists an $\epsilon' > 0$ with the property that

$$E_M\left( 1(A) \times \max_{m=1,\ldots,m'} Z^*_n(\tau, m)^2 n/m' \right) < \epsilon$$

(36)

for any event $A \in G \equiv \sigma(X_{11}, \ldots, X_{nn}; M_{n1}, \ldots, M_{nJ_n}; J_n)$ with $\Pr(A) \leq \epsilon'$. Then we will show that this property implies uniform integrability.\(^{11}\)

First, take an arbitrary $\epsilon > 0$ and a value of $\epsilon' > 0$ so that

$$E_M\left( 1(\bar{A}) \sum_{\tau=0}^{n-1} \frac{1}{m'} \max_{m=1,\ldots,m'} Z_n(\tau, m)^2 \right) < \epsilon/4$$

\(^{11}\)The difference between this proof and the proof of (34) is the measurability requirement. Here we need to finesse measurability of the sets $A$. 17
for any $A \in \mathcal{G}$ with $\Pr(A) \leq \epsilon'$, which is shown to exist in the first part of this Lemma. Then define $J(x)$ to be the block index such that $K_{n,J(x)} - 1 < x \leq K_{n,J(x)}$. We have the relationships

\[
E_M \left( 1(A) \times \max_{m=1,\ldots,m'} n Z_n^*(\tau, m)^2/m' \right)
\]
\[= E_M 1(A) E_M \left[ \max_{m=1,\ldots,m'} \left( \sum_{j=J(\tau)+1}^{J(\tau+m')-1} Z_{n,j}^* + Z_n^*(\tau, K_{n,J(\tau)} - \tau) + Z_n^*(K_{n,J(\tau+m')-1}, m - K_{n,J(\tau+m')-1}) \right)^2 \right]
\]
\[\leq 4 E_M 1(A) \frac{1}{m'} \sum_{u=0}^{n-1} \left( \sum_{j=J(\tau)+1}^{J(\tau+m')-1} Z_n(u, M_{n,j})^2 + Z_n(u, K_{n,J(\tau)} - \tau)^2 + Z_n(u, m' - K_{n,J(\tau+m')-1})^2 \right)
\]
\[= \left( \frac{4}{m'} \right) \sum_{j=J(\tau)+1}^{J(\tau+m')-1} E_M \left[ 1(A) \sum_{u=0}^{n-1} Z_n(u, M_{n,j})^2 \right]
\]
\[+ \left( \frac{4}{m'} \right) E_M \left[ 1(A) \sum_{u=0}^{n-1} Z_n(\tau, M_{n,J(\tau)})^2 \right]
\]
\[+ \left( \frac{4}{m'} \right) E_M \left[ 1(A) \sum_{u=0}^{n-1} Z_n(u, m' - K_{n,J(\tau+m')})^2 \right]
\]
\[\leq \epsilon.
\]

The first inequality is a consequence of Doob’s maximal inequality for martingales (see Davidson, 1994, Theorem 15.15, for example), since $Z_{n,j}^*$ is a martingale difference array, and the second equality and last inequality follow by construction of $A$. A similar argument also implies that $n Z_n^*(\tau, m)^2/m'$ has finite first moment.

Now, for uniform integrability, take $\epsilon > 0$, choose $\epsilon'$ so that (36) holds for all $A \in \mathcal{G}$ s.t. $\Pr(A) \leq \epsilon'$, and choose $C_0$ large enough that

\[
\Pr_M \left[ \max_{m=1,\ldots,m'} n Z_n^*(\tau, m)^2/m' > C_0 \right] \leq \epsilon'.
\]

Markov’s inequality guarantees the existence of this $C_0$. Then

\[
E_M \left( \max_{m=1,\ldots,m'} n Z_n^*(\tau, m)^2/m' \times 1 \left\{ \max_{m=1,\ldots,m'} n Z_n^*(\tau, m)^2/m' > C \right\} \right) \leq \epsilon
\]

for all $C > C_0$. Since $\epsilon$ is arbitrary, this completes the proof. \qed
Lemma 6. Suppose the conditions of Theorem 1 hold. For any positive \( \delta \), there exist positive and finite constants \( C \), \( n_0 \), and \( \varepsilon \) such that for all \( n > n_0 \), \( m = 1, \ldots, n \), \( \tau = 0, \ldots, m \), and \( \ell = 1, \ldots, m - 1 \):

\[
E \left| \sum_{i=0}^{\lfloor n/m \rfloor - 1} [Z_n'(\tau + im, m - \ell)^2 - E(Z_n'(\tau + im, m - \ell)^2)] \right| \leq 2\delta + C \cdot \left( \frac{m}{n} \right)^{1/2} \left( \frac{m}{\ell + 1} \right)^{1/2}. \quad (37)
\]

Also, there exists a constant \( C \) and a finite function \( D(x) \) such that \( D(x) \to 0 \) as \( x \to \infty \) such that, for large enough \( n \),

\[
E \left| \sum_{i=0}^{\lfloor n/m \rfloor - 1} E(Z_n'(\tau + im, m - \ell)^2 - \sigma^2) \right| \leq C D(\ell). \quad (38)
\]

Results (37) and (38) are direct extensions of de Jong’s (1997) Lemmas 5 and 4, respectively, replacing de Jong’s implicit use of inequalities with explicit inequalities. The supplemental appendix presents a proof of (38) to show the basic idea.

Lemma 7. Under the conditions of Theorem 1,

\[
\lim_{C \to 0} \limsup_{n \to \infty} \sup_{\tau=0,\ldots,n-1} \max_{m'=1,\ldots,n} E\left( \left( \max_{m=1,\ldots,m'} Z_n'(\tau, m)^2 n/m' \right) \times 1\{ \max_{m=1,\ldots,m'} Z_n'(\tau, m)^2 n/m' > C \} \right) = 0. \quad (39)
\]

Proof. See supplemental appendix for the proof of (39), which follows McLeish (1975, Lemma 6.5) and McLeish (1977, Lemma 3.5) almost exactly and is also presented as Theorem 16.13 in Davidson (1994).

References


