The Annals of Statistics
2017, Vol. 45, No. 5, 1895–1919
DOI: 10.1214/16-AOS1512

GAUSSIAN APPROXIMATION FOR HIGH DIMENSIONAL TIME SERIES

BY DANNA ZHANG AND WEI BIAO WU

University of California, San Diego and University of Chicago

We consider the problem of approximating sums of high dimensional stationary time series by Gaussian vectors, using the framework of functional dependence measure. The validity of the Gaussian approximation depends on the sample size n, the dimension p, the moment condition and the dependence of the underlying processes. We also consider an estimator for long-run covariance matrices and study its convergence properties. Our results allow constructing simultaneous confidence intervals for mean vectors of high-dimensional time series with asymptotically correct coverage probabilities. As an application, we propose a Kolmogorov–Smirnov-type statistic for testing distributions of high-dimensional time series.

1. Introduction. During the past decade, there has been a significant development on high-dimensional data analysis with applications in many fields. In this paper, we shall consider simultaneous inference for mean vectors of high-dimensional stationary processes, so that one can perform family-wise multiple testing or construct simultaneous confidence intervals, an important problem in the analysis of spatial-temporal processes. To fix the idea, let (X_i) be a stationary process in $\mathbb{R}^p$ with mean $\mu = (\mu_1, \dots, \mu_p)^\top$ and finite second moment in the sense that $\mathbb{E}(X_i^\top X_i) < \infty$. In the scalar case in which $p = 1$ or when p is fixed, under suitable weak dependence conditions, we can have the central limit theorem (CLT):

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \Rightarrow N(0, \Sigma), \qquad \text{where } \Sigma = \sum_{k=-\infty}^{\infty} \mathbb{E}\big((X_0 - \mu)(X_k - \mu)^\top\big).$$

See, for example, [4, 14, 20, 37, 44] among others. In the high dimension case in which p can also diverge to infinity, [33] showed that the central limit theorem can fail for i.i.d. random vectors if $\sqrt{n} = o(p)$. In this paper, we shall consider an alternative form: Gaussian approximation for the largest entry of the sample mean vector $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$. For a vector $v = (v_1, \dots, v_p)^\top$, let $|v|_\infty = \max_{j \le p} |v_j|$. Specifically, our primary goal is to establish the Gaussian Approximation (GA)

Received April 2016; revised August 2016.
MSC2010 subject classifications. Primary 62M10; secondary 62E17.
Key words and phrases. Gaussian approximation, high-dimensional time series, Kolmogorov–Smirnov test, long run covariance matrix, simultaneous inference.

in $\mathbb{R}^p$

$$\sup_{u\geq 0}\bigl|\mathbb{P}(\sqrt{n}|\bar{X}_n-\mu|_\infty \geq u)-\mathbb{P}(|Z_j|_\infty \geq u)\bigr| \to 0, \tag{1.1}$$

where both $n, p \to \infty$. Here, the Gaussian vector $Z = (Z_1, \dots, Z_p)^\top \sim N(0, \Sigma)$. Chernozhukov, Chetverikov and Kato [10] studied the Gaussian approximation for independent random vectors. There has been limited research on high-dimensional inference under dependence. The associated statistical inference becomes considerably more challenging since the autocovariances with all lags should be considered. Zhang and Cheng [49] extended the Gaussian approximation in [10] to very weakly dependent random vectors which satisfy a uniform geometric moment contraction condition. The latter condition is also adopted in [8] for self-normalized sums. Chernozhukov, Chetverikov and Kato [11] did a similar extension to strong mixing random vectors. Here, we shall establish (1.1) for a wide class of high-dimensional stationary process under suitable conditions on the magnitudes of p, n and the mild dependence conditions on the process (X_i).

In Section 2, we shall introduce the framework of high-dimensional time series and some concepts about functional dependence measures that are useful for establishing an asymptotic theory. The main result for Gaussian approximation of the normalized mean vector and the choice of the normalization matrix is presented in Section 3. Depending on the moment and the dependence conditions, both high dimension and ultra high dimension cases are discussed. In Section 3.1, we apply our Gaussian approximation result to simultaneous inference of entries of sample covariance matrices of high-dimensional time series. In Section 4, we shall develop a Kolmogorov–Smirnov-type statistic for testing distributions of high-dimensional time series.

To perform statistical inference based on (1.1), one needs to estimate the long-run covariance matrix Σ. The latter problem has been extensively studied in the scalar and the low-dimensional case; see [1, 5, 23, 30, 32], among others. In Section 5, we study the batched-mean estimate of long-run covariance matrices and derive a large deviation result about quadratic forms of stationary processes. The latter tail probability inequalities allow dependent and/or non-sub-Gaussian processes under mild conditions, which are expected to be useful in other high-dimensional inference problems for dependent vectors. The consistency of the batched-mean estimate ensures the validity of the quantile estimates of $\mathcal{L}^\infty$ norms of sample means; see Section 5.1.

We provide in Section 6 some sharp inequalities for tail probabilities for high dimensional dependent processes in the polynomial tail case. The readers are referred to Appendix (supplementary material [48])C for the tail probability inequalities in the one-dimensional case under finite polynomial moment and exponential moment conditions, respectively. Part of the proofs are relegated to Section 7. Appendix D includes a simulation study.

We now introduce some notation. For a random variable X and $q > 0$, we write $X \in \mathcal{L}^q$ if $\|X\|_q := (\mathbb{E}|X_j|^q)^{1/q} < \infty$, and for a vector $v = (v_1, \dots, v_p)^\top$, let the norm-$s$ length $|v|_s = (\sum_{j=1}^p |v_j|^s)^{1/s}$, $s \ge 1$. Write the $p \times p$ identity matrix as Id_p. For two real numbers, set $x \vee y = \max(x, y)$ and $x \wedge y = \min(x, y)$. For two sequences of positive numbers (a_n) and (b_n), we write $a_n \asymp b_n$ (resp., $a_n \lesssim b_n$ or $a_n \ll b_n$) if there exists some constant $C > 0$ such that $C^{-1} \le a_n/b_n \le C$ (resp., $a_n/b_n \le C$ or $a_n/b_n \to 0$) for all large n. We use $C, C_1, C_2, \dots$ to denote positive constants whose values may differ from place to place. A constant with a symbolic subscript is used to emphasize the dependence of the value on the subscript. Throughout the paper, we assume $p = p_n \to \infty$ as $n \to \infty$.

2. High-dimensional time series. Let ε_i, $i \in \mathbb{Z}$, be i.i.d. random elements and $\mathcal{F}^i = (\dots, \varepsilon_{i-1}, \varepsilon_i)$; let (X_i) be a stationary process taking values in $\mathbb{R}^p$ that assumes the form

$$X_i = (X_{i1}, X_{i2}, \dots, X_{ip})^\top = G(\mathcal{F}^i), \tag{2.1}$$

where $G(\cdot) = (g_1(\cdot), \dots, g_p(\cdot))^\top$ is an $\mathbb{R}^p$-valued measurable function such that X_i is well-defined. In the scalar case with $p = 1$, (2.1) allows a very general class of stationary processes (cf. [35, 38, 40, 41, 43–45]). It includes linear processes as well as a large class of nonlinear time series models. For example, if ε_i, $i \in \mathbb{Z}$, are i.i.d. d-dimensional random vectors with mean 0 and $\mathbb{E}(\varepsilon_i^\top \varepsilon_i) < \infty$, and A_i, $i \ge 0$, are $p \times d$ coefficient matrices with real entries such that $\sum_{i=0}^\infty \mathrm{tr}(A_i^\top A_i) < \infty$, where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. Then by Kolmogorov's three-series theorem, the linear process

$$X_i = \sum_{l=0}^{\infty} A_l \varepsilon_{i-l} \tag{2.2}$$

exists, and it is of form (2.1) with a linear functional G. In particular, the vector AR(1) process $X_i = AX_{i-1} + \varepsilon_i$ has form (2.2) with $A_l = A^l$ if $\max_{j \le p} |\lambda_j(A)| < 1$, where A is a coefficient matrix and $\lambda_1(A), \dots, \lambda_p(A)$ are eigenvalues of A. Within this framework, (ε_i) can be viewed as independent inputs of a physical system and all the dependencies among the outputs (X_i) result from the underlying data-generating mechanism $G(\cdot)$. The function $g_j(\cdot)$, $1 \le j \le p$, is the jth coordinate projection of $G(\cdot)$. Unless otherwise specified, assume throughout the paper that $\mathbb{E}X_i = 0$ and $\max_{j \le p} \|X_{ij}\|_q < \infty$ for some $q \ge 2$. Let $\Gamma(l) = (\gamma_{jk}(l))_{j,k=1}^p = \mathbb{E}(X_i X_{i+l}^\top)$ be the autocovariance matrix and recall the long-run covariance matrix

$$\Sigma = (\sigma_{jk})_{j,k=1}^p = \sum_{l=-\infty}^{\infty} \Gamma(l) \tag{2.3}$$

if it exists. Note that $\sigma_{jj} = \sum_{l=-\infty}^{\infty} \gamma_{jj}(l)$, $1 \le j \le p$, is the long-run variance of the component process $X_{\cdot j} = (X_{ij})_{i \in \mathbb{Z}}$. For the latter process, following [44] we define the functional dependence measure:

$$\delta_{i,q,j} = \|X_{ij} - X_{ij,\{0\}}\|_q = \|X_{ij} - g_j(\mathcal{F}^{i,\{0\}})\|_q, \tag{2.4}$$

where $\mathcal{F}^{i,\{k\}} = (\ldots, \varepsilon_{k-1}, \varepsilon'_k, \varepsilon_{k+1}, \ldots, \varepsilon_i)$ is a coupled version of $\mathcal{F}^i$ with ε_k in $\mathcal{F}^i$ replaced by ε'_k, and $\varepsilon_i, \varepsilon'_l$, $i, l \in \mathbb{Z}$, are i.i.d. random elements. Note that $\mathcal{F}^{i,\{k\}} = \mathcal{F}^i$ if $k > i$. To account for the dependence in the process $X_{\cdot j}$, we define the dependence adjusted norm

$$\|X_{\cdot j}\|_{q,\alpha} = \sup_{m \ge 0} (m+1)^\alpha \Delta_{m,q,j}, \qquad \alpha \ge 0, \text{ where } \Delta_{m,q,j} = \sum_{i=m}^{\infty} \delta_{i,q,j}. \tag{2.5}$$

Due to the dependence, it may happen that $\max_{j \le p} \|X_{ij}\|_q < \infty$ while $\|X_{\cdot j}\|_{q,\alpha} = \infty$. Elementary calculations show that, if $X_{ij}, i \in \mathbb{Z}$, are i.i.d., then $\|X_{ij}\|_q \le \|X_{\cdot j}\|_{q,\alpha} \le 2\|X_{ij}\|_q$, suggesting that the dependence adjusted norm is equivalent to the classical $\mathcal{L}^q$ norm.

To account for high-dimensionality, we define

$$\Psi_{q,\alpha} = \max_{1 \le j \le p} \|X_{\cdot j}\|_{q,\alpha} \quad \text{and} \quad \Upsilon_{q,\alpha} = \left(\sum_{j=1}^{p} \|X_{\cdot j}\|_{q,\alpha}^q \right)^{1/q},$$

which can be interpreted as the uniform and the overall dependence adjusted norms of $(X_i)_{i \in \mathbb{Z}}$, respectively. The form (2.1) and its associated dependence measures provide a convenient framework for studying high-dimensional time series. Zhang and Cheng [49] considered the special case which imposes the stronger geometric moment contraction condition $\max_{1 \le j \le p} \Delta_{m,q,j} \le C\rho^m$ with $\rho \in (0, 1)$ and some constant C. This assumption can be fairly restrictive. In this paper $\Psi_{q,\alpha}$ can be unbounded in p. Additionally, we define the $\mathcal{L}^\infty$ functional dependence measure and its corresponding dependence adjusted norm for the p-dimensional stationary process (X_i)

$$\omega_{i,q} = \big\| |X_i - X_{i,\{0\}}|_\infty \big\|_q;$$

$$\big\| |X_\cdot|_\infty \big\|_{q,\alpha} = \sup_{m \ge 0} (m+1)^\alpha \Omega_{m,q}, \qquad \alpha \ge 0, \text{ where } \Omega_{m,q} = \sum_{i=m}^{\infty} \omega_{i,q}.$$

Clearly, we have $\Psi_{q,\alpha} \le \| |X_\cdot|_\infty \|_{q,\alpha} \le \Upsilon_{q,\alpha}$.

3. Gaussian approximations.

In this section, we shall present main results on Gaussian approximations. Theorem 3.2 concerns the finite polynomial moment case with both weaker and stronger temporal dependence. If the underlying process has finite dependence adjusted sub-exponential norms, Theorem 3.3 asserts that an ultra-high dimension p can be allowed. Theorem 7.4 in Section 7.1 provides a convergence rate of the Gaussian approximation.

Recall (2.3) for the long-run covariance matrix Σ. Let $\Sigma_0 = \mathrm{diag}(\Sigma)$ be the diagonal matrix of Σ, and $D_0 = \mathrm{diag}(\sigma_{11}^{1/2}, \dots, \sigma_{pp}^{1/2}) = \Sigma_0^{1/2}$. Assume $\mu = 0$. We consider the following normalized version of (1.1):

$$\rho_n := \sup_{u \ge 0} \bigl| \mathbb{P}\bigl(\sqrt{n} \bigl|D_0^{-1} \bar{X}_n\bigr|_\infty \ge u\bigr) - \mathbb{P}\bigl(\bigl|D_0^{-1} Z\bigr|_\infty \ge u\bigr)\bigr| \to 0. \tag{3.1}$$

ASSUMPTION 3.1. There exists a constant $c > 0$ such that $\min_{1 \le j \le p} \sigma_{jj} \ge c$.

To state Theorem 3.2, we need to define the following quantities:

$$\Theta_{q,\alpha} = \Upsilon_{q,\alpha} \wedge \bigl(\bigl\| |X_\cdot|_\infty \bigr\|_{q,\alpha} (\log p)^{3/2}\bigr), \qquad L_1 = \bigl(\Psi_{2,\alpha} \Psi_{2,0} (\log p)^2\bigr)^{1/\alpha},$$
$$W_1 = \bigl(\Psi_{3,0}^6 + \Psi_{4,0}^4\bigr)\bigl(\log(pn)\bigr)^7, \qquad W_2 = \Psi_{2,\alpha}^2 \bigl(\log(pn)\bigr)^4,$$
$$W_3 = \bigl(n^{-\alpha} \bigl(\log(pn)\bigr)^{3/2} \Theta_{q,\alpha}\bigr)^{1/(1/2-\alpha-1/q)},$$
$$N_1 = (n/\log p)^{q/2} / \Theta_{q,\alpha}^q, \qquad N_2 = n (\log p)^{-2} \Psi_{2,\alpha}^{-2},$$
$$N_3 = \bigl(n^{1/2} (\log p)^{-1/2} \Theta_{q,\alpha}^{-1}\bigr)^{1/(1/2-\alpha)}.$$

THEOREM 3.2. *Let Assumption* 3.1 *be satisfied*. (i) *Assume that* $\Theta_{q,\alpha} < \infty$ *holds with some* $q \ge 4$ *and* $\alpha > 1/2 - 1/q$ (*the weaker dependence case*),

$$\Theta_{q,\alpha} n^{1/q-1/2} \bigl(\log(pn)\bigr)^{3/2} \to 0 \tag{3.2}$$

and

$$L_1 \max(W_1, W_2) = o(1) \min(N_1, N_2). \tag{3.3}$$

Then the Gaussian approximation (3.1) *holds*. (ii) *Assume* $0 < \alpha < 1/2 - 1/q$ (*the stronger dependence case*). *Then* (3.1) *holds if* $\Theta_{q,\alpha} (\log p)^{1/2} = o(n^\alpha)$ *and*

$$L_1 \max(W_1, W_2, W_3) = o(1) \min(N_2, N_3). \tag{3.4}$$

REMARK 1. A careful check of the proof of Theorem 3.2 indicates that if it is further assumed that $\max_{1 \le j \le p} \sigma_{jj}$ is bounded from above, the Gaussian approximation is also valid for the nonnormalized maximum, that is, for both cases of Theorem 3.2,

$$\sup_{u \ge 0} \bigl| \mathbb{P}\bigl(\sqrt{n} |\bar{X}_n|_\infty \ge u\bigr) - \mathbb{P}\bigl(|Z_j|_\infty \ge u\bigr)\bigr| \to 0. \tag{3.5}$$

REMARK 2 (Optimality of our result on the allowed dimension p). Assume $\alpha > 1/2 - 1/q$. In the special case with $\Psi_{q,\alpha} \asymp 1$ and $\Theta_{q,\alpha} \asymp p^{1/q}$, (3.2) becomes

$$p \bigl(\log(pn)\bigr)^{3q/2} = o\bigl(n^{q/2-1}\bigr), \tag{3.6}$$

which by elementary manipulations implies (3.3), and hence the GA (3.1). It turns out that condition (3.6), or equivalently $p(\log p)^{3q/2} = o(n^{q/2-1})$, is optimal up to a multiplicative logarithmic term. Consider the special case in which X_{ij}, $i, j \in \mathbb{Z}$, are i.i.d. symmetric random variables with $\mathbb{E}(X_{ij}^2) = 1$ and the tail probability $\mathbb{P}(X_{ij} \geq u) = u^{-q}\ell(u)$, $u \geq u_0$, where $\ell(u) = (\log u)^{-2}$. By Theorem 1.9 of [29], we have the expansion: for a sequence $y_n \geq \sqrt{n}$, as $n \to \infty$,

$$\frac{\mathbb{P}(X_{11} + \cdots + X_{n1} \geq y_n)}{n y_n^{-q}\ell(y_n) + 1 - \Phi(y_n/\sqrt{n})} \to 1. \tag{3.7}$$

Let $M_n = X_{11} + \cdots + X_{n1}$, $Z = (Z_1, \ldots, Z_p)^\top \sim N(0, \mathrm{Id}_p)$ and assume

$$n^{q/2-1} = o\big(p(\log n)^{-2}(\log p)^{-q/2}\big). \tag{3.8}$$

Then the GA (3.1) *does not hold*. To see this, let $u = (2\log p)^{1/2}$. Then $p\mathbb{P}(|Z_1| \geq u) \to 0$, and, by (3.7) and (3.8), $p\mathbb{P}(M_n \geq \sqrt{n}u) \to \infty$. Hence, $\mathbb{P}^p(|M_n| \leq \sqrt{n}u) \to 0$ and $\mathbb{P}^p(|Z_1| \leq u) \to 1$, implying that

$$\begin{aligned}\rho_n &\geq \big|\mathbb{P}\big(\sqrt{n}|\bar{X}_n|_\infty \leq u\big) - \mathbb{P}\big(|Z_j|_\infty \leq u\big)\big| \\ &= \big|\mathbb{P}^p\big(|M_n| \leq \sqrt{n}u\big) - \mathbb{P}^p\big(|Z_1| \leq u\big)\big| \\ &= \big|\big[1 - 2\mathbb{P}(M_n \geq \sqrt{n}u)\big]^p - \mathbb{P}^p\big(|Z_1| \leq u\big)\big| \to 1.\end{aligned}$$

Note that (3.8) is equivalent to $n^{q/2-1} = o(p(\log p)^{-2-q/2})$, suggesting that (3.6) is optimal up to a logarithmic term.

Now suppose there exist $0 \leq \kappa_1 \leq \kappa_2$ such that $\Psi_{q,\alpha} \asymp p^{\kappa_1}$ and $\Theta_{q,\alpha} \asymp p^{\kappa_2}$, and $p^\tau \asymp n$. Elementary but tedious calculations show that, in the weaker dependence case $\alpha > 1/2 - 1/q$, if

$$\tau > \max\left\{\frac{\kappa_2}{1/2 - 1/q}, \frac{2\kappa_1}{\alpha} + 8\kappa_1, \frac{2}{q}\left(\frac{2\kappa_1}{\alpha} + 8\kappa_1\right) + 2\kappa_2\right\}, \tag{3.9}$$

then conditions in (i) of Theorem 3.2 are satisfied, while for the stronger dependence case with $0 < \alpha < 1/2 - 1/q$, a larger sample size n is required:

$$\tau > \max\left\{\frac{\kappa_2}{\alpha}, \frac{2\kappa_1}{\alpha} + 8\kappa_1, (1 - 2\alpha)\left(\frac{2\kappa_1}{\alpha} + 8\kappa_1\right) + 2\kappa_2\right\}. \tag{3.10}$$

The lower bounds in (3.9) and (3.10) are both nondecreasing of κ_1, κ_2 and nonincreasing in q, α.

Next we consider the sub-exponential case in which X_{ij} satisfies a stronger moment condition than the existence of finite qth moment. Assume that X_{ij} has finite moment with any order. For $\nu \geq 0$ and $\alpha \geq 0$, define the dependence adjusted sub-exponential norm

$$\|X_{\cdot j}\|_{\psi_\nu,\alpha} = \sup_{q \geq 2} \frac{\|X_{\cdot j}\|_{q,\alpha}}{q^\nu} \quad \text{and} \quad \Phi_{\psi_\nu,\alpha} = \max_{j \leq p} \|X_{\cdot j}\|_{\psi_\nu,\alpha}.$$

By this definition, if X_{ij}, $i \in \mathbb{Z}$ are i.i.d., $\|X_{\cdot j}\|_{\psi_\nu,\alpha}$ is equivalent to the sub-Gaussian norm ($\nu = 1$) or sub-exponential norm ($\nu = 1/2$), due to the equivalence of $\|X_{\cdot j}\|_{q,\alpha}$ and $\|X_{ij}\|_q$. The parameter ν measures how fast $\|X_{\cdot j}\|_{q,\alpha}$ increases with q.

To state Theorem 3.3, we let $\beta = 2/(1+2\nu)$ and define

$$L_2 = \big((\log p)^{1/\beta+1/2}\Phi_{\psi_\nu,\alpha}\big)^{1/\alpha}, \qquad N_4 = n(\log p)^{-1-2/\beta}\Phi^{-2}_{\psi_\nu,0},$$
$$W_4 = \big(\log(pn)\big)^{3+2/\beta}\Phi^2_{\psi_\nu,0} + \big(\log(pn)\big)^4.$$

THEOREM 3.3. *Let Assumption 3.1 be satisfied. Assume that $\Phi_{\psi_\nu,\alpha} < \infty$ for some $\nu \geq 0$, $\alpha > 0$ and*

$$\max(L_1, L_2)\max(W_1, W_4) = o(N_4), \qquad L_1^\alpha \max(W_1, W_4) = o(n). \tag{3.11}$$

Then the Gaussian approximation (3.1) holds.

If $\Phi_{\psi_\nu,\alpha} \asymp 1$, then the ultra high-dimensional case with $\log p = o(n^c)$ with some $c > 0$ is allowed, where specifically we can let

$$c = \begin{cases} 1/(8+2/\alpha+2/\beta), & 2/3 \leq \beta \leq 2, \\ 1/\big[7+(1/\beta+1/2)(1/\alpha+2)\big], & 1/2 \leq \beta < 2/3, \\ 1/\big[3+2/\beta+(1/\beta+1/2)(1/\alpha+2)\big], & 0 < \beta < 1/2. \end{cases} \tag{3.12}$$

3.1. *Simultaneous inference of covariances.* Let $X_1, \ldots, X_n$ be i.i.d. p-dimensional vectors with mean 0 and covariance matrix $\Gamma_0 = (\gamma_{jk})_{j,k=1}^p = \mathbb{E}(X_i X_i^\top)$. We estimate Γ_0 by the sample covariance matrix $\hat{\Gamma}_0 = (\hat{\gamma}_{jk})_{j,k=1}^p = n^{-1}\sum_{i=1}^n X_i X_i^\top$. To perform simultaneous inference on γ_{jk}, $1 \leq j, k \leq p$, one needs to derive the asymptotic distribution of the maximum deviation $\max_{j,k\leq p}|\hat{\gamma}_{jk} - \gamma_{jk}|$ or the normalized version $\max_{j,k\leq p}|\hat{\gamma}_{jk} - \gamma_{jk}|/\tau_{jk}$; cf. equation (2) in [46]. The former is also referred to as the mutual coherence of the data matrix in the compressed sensing literature (see, e.g., [15]). Jiang [21] established the Gumbel convergence of the maximum deviation under some polynomial moment condition and under the setup that all entries of X_i are also independent. See [26, 28, 50] and [25] for some refined results. Cai and Jiang [7] showed that $\max_{|j-k|>s_n}|\hat{\gamma}_{jk} - \gamma_{jk}|$ also converges to the Gumbel distribution if $(X_{ij})_{1\leq j\leq p}$ is Gaussian and s_n-dependent for each i. Xiao and Wu [46] considered the extension to the non-Gaussian case and allowed a general dependence structure among entries of X_i. However, the latter two paper both require that the vectors $X_1, \ldots, X_n$ are i.i.d. The problem of further extension to temporally dependent X_i is open. In analyzing fMRI functional connectivity in brain networks in the format of multivariate time series, researchers use the maximum correlation between time series to identify edges that connect the corresponding nodes in a network (cf. [13, 18, 19, 24], among many others). Such applications suggest that an asymptotic theory for maximum deviations of sample covariances is needed.

Our Theorems 3.2 and 3.3 can be applied to the above problem of further extension to temporally dependent processes. Let (X_i) be a mean zero p-dimensional stationary process of form (2.1). To apply Theorems 3.2 and 3.3, one needs to deal with the key issue of computing the functional dependence measure of the p^2-dimensional vector $\mathcal{X}_i = \operatorname{vec}(X_i X_i^\top - \mathbb{E}(X_i X_i^\top))$. Interestingly, our framework allows a natural and elegant treatment. Let $a = (j,k)$, $j,k \le p$ and $\mathcal{X}_{ia} = X_{ij}X_{ik} - \gamma_a$, where $\gamma_a = \mathbb{E}(X_{ij}X_{ik})$. By Hölder's inequality, the functional dependence of the component process $(\mathcal{X}_{ia})_i$:

$$\begin{aligned}\varphi_{i,q/2,a} &:= \|X_{ij}X_{ik} - \mathbb{E}(X_{ij}X_{ik}) - X_{ij,\{0\}}X_{ik,\{0\}} + \mathbb{E}(X_{ij,\{0\}}X_{ik,\{0\}})\|_{q/2} \\ &\le 2\|X_{ij}X_{ik} - X_{ij,\{0\}}X_{ik,\{0\}}\|_{q/2} \\ &\le 2\|X_{ij}(X_{ik} - X_{ik,\{0\}})\|_{q/2} + 2\|(X_{ij} - X_{ij,\{0\}})X_{ik,\{0\}}\|_{q/2} \\ &\le 2\|X_{ij}\|_q \delta_{i,q,k} + 2\|X_{ik}\|_q \delta_{i,q,j}.\end{aligned} \tag{3.13}$$

Hence, we can have an upper bound of the dependence adjusted norm of $(\mathcal{X}_{ia})$

$$\begin{aligned}\|\mathcal{X}_{\cdot a}\|_{q/2,\alpha} &:= \sup_{m\ge 0}(m+1)^\alpha \sum_{i=m}^{\infty} \varphi_{i,q/2,j,k} \\ &\le 2\|X_{\cdot j}\|_{q,0}\|X_{\cdot k}\|_{q,\alpha} + 2\|X_{\cdot k}\|_{q,0}\|X_{\cdot j}\|_{q,\alpha}.\end{aligned} \tag{3.14}$$

Consequently, the uniform and the overall dependence adjusted norms of $\mathcal{X}_i$ are

$$\begin{aligned}&\max_a \|\mathcal{X}_{\cdot a}\|_{q/2,\alpha} \le 4\Psi_{q,0}\Psi_{q,\alpha}, \\ &\Big(\sum_a \|\mathcal{X}_{\cdot a}\|_{q/2,\alpha}^{q/2}\Big)^{2/q} \le 4\Big(\sum_{j=1}^p \|X_{\cdot j}\|_{q,0}^{q/2}\Big)^{2/q}\Big(\sum_{j=1}^p \|X_{\cdot j}\|_{q,\alpha}^{q/2}\Big)^{2/q}.\end{aligned} \tag{3.15}$$

Similarly, the $\mathcal{L}^\infty$ dependence adjusted norm for the process $(\mathcal{X}_i)$ can be calculated by

$$\big\||\mathcal{X}_\cdot|_\infty\big\|_{q/2,\alpha} \le 4\big\||X_\cdot|_\infty\big\|_{q,0}\big\||X_\cdot|_\infty\big\|_{q,\alpha}. \tag{3.16}$$

With (3.13)–(3.16), conditions in Theorems 3.2 and 3.3 can be formulated accordingly, and under those conditions we can have the following Gaussian approximation:

$$\sup_{u\ge 0}\Big|\mathbb{P}\Big(\sqrt{n}\max_a |\hat{\gamma}_a - \gamma_a|/\tau_a \ge u\Big) - \mathbb{P}\Big(\max_a |Z_a/\tau_a| \ge u\Big)\Big| \to 0, \tag{3.17}$$

where $Z = (Z_a)_a \sim N(0, \Sigma_{\mathcal{X}})$, $\Sigma_{\mathcal{X}}$ is the $p^2 \times p^2$ long-run covariance matrix of $(\mathcal{X}_i)_i$ and $(\tau_a^2)_a$ is the diagonal matrix of $\Sigma_{\mathcal{X}}$.

4. A uniform test for distributions of time series. In this section, we shall apply the Gaussian approximation result Theorem 3.2 and test distributions of time series. For the process (X_i) defined in (2.1), let $F_j(u) = \mathbb{P}(X_{ij} \le u)$, $u \in \mathbb{R}$, be the cumulative distribution function (c.d.f.) of X_{ij}, $1 \le j \le p$; let $F_{j,0}(\cdot)$ be the reference c.d.f. We are interested in testing the null hypothesis:

$$H_0 : F_j(\cdot) = F_{j,0}(\cdot) \qquad \text{for all } j = 1, \dots, p. \tag{4.1}$$

In the classical Kolmogorov–Smirnov test with $p = 1$ and i.i.d. data X_{i1}, $i \in \mathbb{Z}$, one uses a test statistic that involves the supremum distance between the empirical and the reference c.d.f.s. Here, we shall apply a smoothing procedure and consider testing an equivalent form of (4.1). In particular, we let $h(u) = H'(u)$ be a probability density function (p.d.f.) such that $h(u) > 0$ for all $u \in \mathbb{R}$, $\sup_u h(u) < \infty$ and let

$$H_j(u) = \int_{\mathbb{R}} F_j(v)h(u - v)\,dv \quad \text{and} \quad H_{j,0}(u) = \int_{\mathbb{R}} F_{j,0}(v)h(u - v)\,dv. \tag{4.2}$$

For example, we can let $h(\cdot)$ be the standard Gaussian p.d.f. In this case, $H_j(\cdot)$ is the c.d.f. of $X_{ij} + \eta$, where $\eta \sim N(0, 1)$ is independent of X_{ij}. Here, we shall consider testing the following equivalent form of (4.1):

$$H_0 : H_j(\cdot) = H_{j,0}(\cdot) \qquad \text{for all } j = 1, \dots, p, \tag{4.3}$$

by using the goodness-of-fit test statistic of the form $\sup_{u\in\mathcal{I}} |\hat{H}_j(u) - H_{j,0}(u)|$, where $\mathcal{I} \subset \mathbb{R}$ is an interval and $\hat{H}_j(u)$ is an unbiased estimate of $H_j(u)$:

$$\hat{H}_j(u) = \frac{1}{n}\sum_{i=1}^{n} H(u - X_{ij}). \tag{4.4}$$

Similar smoothing ideas appeared in the literature. Researchers applied kernel smoothing to overcome the shortcoming of discontinuity of empirical distribution functions; see, for example, [3, 9, 16, 36, 42, 47], among others.

Here, we shall develop a Gaussian approximation theory for

$$\Delta_n := \max_{1\le j\le p} \sup_{u\in\mathcal{I}} \sqrt{n}\big|\hat{H}_j(u) - H_j(u)\big|. \tag{4.5}$$

To this end, we shall carry out a detailed calculation for the functional dependence measures defined in Section 2 of $H(u - X_{ij})$. For presentational clarity here, we only consider marginal distributions and linear processes (X_i) defined in (2.2). We remark that our approach also applies to testing for joint distributions and for nonlinear processes.

ASSUMPTION 4.1. The process (X_i) is of form (2.2) with $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{id})^\top$, where ε_{ij} are i.i.d. with mean 0 and $\|\varepsilon_{ij}\|_\gamma < \infty$, $\gamma > 2$; and coefficient matrices $A_i = (a_{i,jk})_{j\le p,k\le d}$ satisfy $\sum_{i=0}^{\infty} \operatorname{tr}(A_i^\top A_i) < \infty$.

For $j,k=1,\dots,p$ and $u,v\in\mathbb{R}$, define the long-run covariance function

$$\sigma_{j,k}(u,v)=\sum_{l=-\infty}^{\infty}\operatorname{Cov}\big(H(u-X_{0j}),H(v-X_{lk})\big). \tag{4.6}$$

Let $\{Z_j(u), j=1,\dots,p; u\in\mathbb{R}\}$ be a mean 0 Gaussian process such that its covariance function is given by (4.6).

ASSUMPTION 4.2. There exists a constant $c>0$ and a closed finite interval $\mathcal{I}\subset\mathbb{R}$ such that $\min_{1\le j\le p}\min_{u\in\mathcal{I}}\sigma_{j,j}(u,u)\ge c$.

THEOREM 4.3. *Let Assumptions 4.1 and 4.2 be satisfied, and suppose there exists a constant $C_1>0$ such that for all $m\ge 0$,*

$$\sum_{i=m}^{\infty}\left(\sum_{k=1}^{d}\max_j|a_{i,jk}|^2\right)^{\min(\gamma/q,1)/2}\le C_1(1\vee m)^{-\alpha} \tag{4.7}$$

holds for some $q\ge 4$ and $\alpha>0$. Let $\iota=\min(\gamma/q,1)/2$. There exists some constant $\kappa>0$ depending on α and ι such that if p satisfies

$$\log p=o\big(n^{\kappa}\big), \tag{4.8}$$

we have

$$\sup_{u\ge 0}\left|\mathbb{P}(\sqrt{n}\Delta_n\ge u)-\mathbb{P}\Big(\max_{1\le j\le d}\sup_{x\in\mathcal{I}}\big|Z_j(x)\big|\ge u\Big)\right|\to 0. \tag{4.9}$$

REMARK 3. A careful check of the proof of Theorem 4.3 indicates that, for the index κ in (4.8), we can let $\kappa=\kappa_1=[(2\iota+2)/\alpha+8\iota+11]^{-1}$ if $\alpha>1/2-1/q$, and $\kappa=\min(\kappa_1,\alpha/(3+\iota))$ if $0<\alpha<1/2-1/q$.

For i.i.d. random vectors, [22] considered uniform convergence of empirical distribution functions. Theorem 4.3 might be the first result in the literature concerning weak convergence of empirical processes in the high-dimensional setting under dependence.

PROOF OF THEOREM 4.3. We shall divide the proof into 5 steps: discretization of the empirical process; representation of the covariance function; continuity of the approximating Gaussian process; computation of the functional dependence measures; and application of Theorem 3.2.

Step 1: *discretization of the empirical process.* Without loss of generality, let $\mathcal{I}=[0,1]$. Let $\mathcal{L}=n^2$ and $u_\ell=\ell/\mathcal{L}$, $\ell=1,\dots,\mathcal{L}$. For $\mathcal{V}=\{(j,\ell):1\le j\le p, 1\le\ell\le\mathcal{L}\}$, define the $(p\mathcal{L})$-dimensional vector $\mathcal{M}_i=(\mathcal{M}_{iv})_{v\in\mathcal{V}}$ with $\mathcal{M}_{iv}=H(u_\ell-X_{ij})-\mathbb{E}H(u_\ell-X_{ij})$ for $v=(j,\ell)\in\mathcal{V}$. Let $\bar{\mathcal{M}}_n=n^{-1}\sum_{i=1}^n\mathcal{M}_i$. Since

$H(\cdot)$ is increasing and $h_0 = \sup_u h(u) < \infty$, we have by the triangle inequality that

$$\bigl|\Delta_n - \sqrt{n}|\bar{\mathcal{M}}_n|_\infty\bigr| \le \frac{h_0\sqrt{n}}{\mathcal{L}} = \frac{h_0}{n\sqrt{n}}. \tag{4.10}$$

Step 2: *representation of the covariance function*. Define the projection operator $\mathcal{P}^i \cdot = \mathbb{E}(\cdot|\mathcal{F}^i) - \mathbb{E}(\cdot|\mathcal{F}^{i-1})$ and

$$D_j(u) = \sum_{l=0}^{\infty} \mathcal{P}^0 H(u - X_{lj}), \qquad j = 1, \dots, p. \tag{4.11}$$

Recall (4.6) for $\sigma_{j,k}(u, v)$. By the orthogonal decomposition,

$$H(u - X_{0j}) - \mathbb{E}H(u - X_{0j}) = \sum_{m=-\infty}^{\infty} \mathcal{P}^m H(u - X_{0j})$$

and the stationarity of (X_i), we have the representation

$$\begin{aligned} \sigma_{j,k}(u, v) &= \sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} \mathbb{E}\bigl[\mathcal{P}^m H(u - X_{0j})\mathcal{P}^m H(v - X_{lk})\bigr] \\ &= \mathbb{E}\bigl[D_j(u)D_k(v)\bigr]. \end{aligned} \tag{4.12}$$

Since $\mathcal{P}^0 H(u - X_{lj}) = \mathbb{E}[H(u - X_{lj}) - H(u - X_{lj,\{0\}})|\mathcal{F}^0]$, by the first inequality in (4.21) and Jensen's inequality, we have

$$\bigl\|\mathcal{P}^0 H(u - X_{lj})\bigr\| \le \bigl\|H(u - X_{lj}) - H(u - X_{lj,\{0\}})\bigr\| \le 2h_0 b_l \|\varepsilon_{ij}\|, \tag{4.13}$$

where $b_i = (\sum_{k=1}^{d} \max_j |a_{i,jk}|^2)^{1/2}$. By (4.7), $\#\{i : b_i \ge 1\} \le C_1$. If $b_i < 1$, then $b_i \le b_i^{\min(1,\gamma/q)}$. Hence, $\sum_{i=0}^{\infty} b_i \le 2C_1$ and

$$\bigl(\sigma_{jj}(u, u)\bigr)^{1/2} = \bigl\|D_j(u)\bigr\| \le \sum_{l=0}^{\infty}\bigl\|\mathcal{P}^0 H(u - X_{lj})\bigr\| \le 4C_1 h_0 \|\varepsilon_{ij}\|. \tag{4.14}$$

Step 3: *continuity of the approximating Gaussian process.* Let $\zeta = |u - v| \le 1$. Then $|H(u - X_{lj}) - H(v - X_{lj})| \le h_0\zeta$. By (4.11) and (4.13),

$$\bigl\|D_j(v) - D_j(u)\bigr\| \le \sum_{l=0}^{\infty} \min\bigl(4h_0 b_l \|\varepsilon_{ij}\|, h_0\zeta\bigr). \tag{4.15}$$

By (4.7), since $2\iota \le 1$ and $\zeta \le 1$, we have $\sum_{i=m}^{\infty} \min(b_i, \zeta) \le C_1 m^{-\alpha}$ for all $m \ge 1$. Let $J = \lceil \zeta^{-1/(1+\alpha)} \rceil$. Then

$$\begin{aligned} \sum_{i=0}^{J} \min(b_i, \zeta) + \sum_{i=J+1}^{\infty} \min(b_i, \zeta) &\le (J + 1)\zeta + C_1 J^{-\alpha} \\ &\le (C_1 + 3)\zeta^{\alpha/(1+\alpha)}. \end{aligned} \tag{4.16}$$

Hence, by (4.12) and (4.15), for $C_2 = h_0(4\|\varepsilon_{ij}\| + 1)(C_1 + 3)$ we obtain

$$\begin{aligned}\|Z_j(u) - Z_j(v)\|^2 &= \sigma_{j,j}(u,u) + \sigma_{j,j}(v,v) - 2\sigma_{j,j}(u,v) \\ &= \|D_j(v) - D_j(u)\|^2 \le C_2^2|u-v|^{2\alpha/(1+\alpha)}\end{aligned} \tag{4.17}$$

when $|u - v| \le 1$. Let $0 < t \le 1$ and $\lambda = \alpha/(1+\alpha)$. By (4.17) and the Fernique inequality (cf. Section 4.1.3 of [17]), there exists constants $c_1, c_2, c_3 > 0$ only depending λ such that for all $w \ge c_2 C_2 t^\lambda$,

$$\mathbb{P}\Big[\sup_{0\le y\le t}|Z_j(v+y) - Z_j(v)| \ge w\Big] \le c_1\big[1 - \Phi\big(c_3 w/(C_2 t^\lambda)\big)\big], \tag{4.18}$$

where $\Phi(\cdot)$ is the standard normal c.d.f. For $u \in \mathcal{I} = [0,1]$, write $\lfloor u\rfloor_{\mathcal{L}} = \mathcal{L}^{-1}\lfloor \mathcal{L}u\rfloor$, where $\lfloor\cdot\rfloor$ is the floor function. As u changes from 0 to 1, $\lfloor u\rfloor_{\mathcal{L}}$ take values $u_0, u_1, \dots, u_{\mathcal{L}}$. Let

$$w = C_3\mathcal{L}^{-\lambda}\big(\log(pn)\big)^{1/2}, \tag{4.19}$$

where C_3 is a sufficiently large constant. Then by (4.18), we have

$$\begin{aligned}\mathbb{P}\Big[\sup_{u\in\mathcal{I},1\le j\le p}\big|Z_j(u) - Z_j\big(\lfloor u\rfloor_{\mathcal{L}}\big)\big| \ge w\Big] &\le p\mathcal{L}c_1\big[1 - \Phi\big(c_3 w\mathcal{L}^\lambda/C_2\big)\big] \\ &\le \frac{C_4}{pn}.\end{aligned} \tag{4.20}$$

Step 4: *computation of the functional dependence measures.* We shall first bound the functional dependence measures of the vector process $(\mathcal{M}_i)_i$ which is induced by $H(u - X_{ij})$. Let $\varepsilon_{ij}, \varepsilon_{i'j'}, i, i', j, j' \in \mathbb{Z}$, be i.i.d. random variables and $\varepsilon'_i = (\varepsilon'_{i1}, \dots, \varepsilon'_{id})^\top$. Note that $X_{ij} - X_{ij,\{0\}} = a_{i,j\cdot}(\varepsilon_0 - \varepsilon'_0)$, where $a_{i,j\cdot}$ is the jth row of the $A_i = (a_{i,jk})_{j\le p,k\le d}$. Then

$$\begin{aligned}\sup_u\big|H(u - X_{ij}) - H(u - X_{ij,\{0\}})\big| &\le \min\big(1, h_0|X_{ij} - X_{ij,\{0\}}|\big) \\ &= \min\big(1, h_0\big|a_{i,j\cdot}(\varepsilon_0 - \varepsilon'_0)\big|\big) \\ &\le \big(h_0\big|a_{i,j\cdot}(\varepsilon_0 - \varepsilon'_0)\big|\big)^{\min(\gamma/q,1)}.\end{aligned} \tag{4.21}$$

Recall $b_i = (\sum_{k=1}^d \max_j |a_{i,jk}|^2)^{1/2}$. By Lemma C.5, we have

$$\Big\|\max_j\big|a_{i,j\cdot}(\varepsilon_0 - \varepsilon'_0)\big|\Big\|_{\min(\gamma,q)} \le C_5 b_i\sqrt{\log p}, \tag{4.22}$$

where the constant C_5 depends on γ, q and $\|\varepsilon_{ij}\|_\gamma$. Hence,

$$\begin{aligned}\Big\|\sup_{j,u}\big|H(u - X_{ij}) - H(u - X_{ij,\{0\}})\big|\Big\|_q &\le \Big[\mathbb{E}\max_j\big(h_0\big|a_{i,j\cdot}(\varepsilon_0 - \varepsilon'_0)\big|\big)^{\min(\gamma,q)}\Big]^{1/q} \\ &\le C_6(\log p)^\iota b_i^{2\iota},\end{aligned}$$

which by (4.7) implies

$$\begin{aligned}\big\||\mathcal{M}.|_\infty\big\|_{q,\alpha} &:= \sup_{m\geq 0}(m+1)^\alpha \sum_{i=m}^{\infty}\Big\|\max_{j,\ell}\big|H(x_\ell - X_{ij}) - H(x_\ell - X_{ij,\{0\}})\big|\Big\|_q \\ &\leq C_7(\log p)^\iota. \end{aligned} \tag{4.23}$$

Then we can obtain the upper bounds of the dependence adjusted norms by

$$\Theta_{q,\alpha} \leq \big(\log(p\mathcal{L})\big)^{3/2}\big\||\mathcal{M}.|_\infty\big\|_{q,\alpha}, \qquad \Psi_{2,\alpha} \leq \Psi_{q,\alpha} \leq \big\||\mathcal{M}.|_\infty\big\|_{q,\alpha}. \tag{4.24}$$

Step 5: *application of Theorem* 3.2. By Theorem 3.2 [cf. (3.5) in Remark 1, which is applicable here in view of (4.14) and Assumption 4.2], we have

$$\sup_{u\geq 0}\Big|\mathbb{P}\big(\sqrt{n}|\bar{\mathcal{M}}_n|_\infty \geq u\big) - \mathbb{P}\Big(\max_{j\leq p}\max_{\ell\leq\mathcal{L}}\big|Z_j(u_\ell)\big| \geq u\Big)\Big| \to 0, \tag{4.25}$$

if the conditions of Theorem 3.2 are satisfied. Specifically, we have $L_1 = O([(\log p)^{2\iota}(\log pn)^2]^{1/\alpha})$, $\max(W_1, W_2) = O((\log p)^{6\iota}(\log pn)^7)$ as well as $n(\log pn)^{-4}(\log p)^{-2\iota} = O(\min(N_1, N_2))$. For $\alpha > 1/2 - 1/q$, there exists some κ depending on α and ι such that if $\log p = o(n^\kappa)$, (3.2) and (3.3) hold. The other case with $0 < \alpha < 1/2 - 1/q$ can be dealt with similarly. Since

$$\left(\frac{\sqrt{n}}{\mathcal{L}} + \frac{1}{pn}\right)\sqrt{\log(p\mathcal{L})} \to 0, \tag{4.26}$$

by the triangle inequality and Theorem 3 of [12], (4.9) follows in view of (4.10), (4.20), (4.25) and (4.26). □

5. Estimation of long-run covariance matrices. Given the realization $X_1, \dots, X_n$, to apply the Gaussian approximation (3.1), we need to estimate the long-run covariance matrix Σ. Note that $\Sigma/(2\pi)$ is the value of the spectral density matrix of (X_i) at zero frequency. In the one or low-dimensional case, there is a large literature concerning spectral density estimation; see, for example, [2, 27, 30, 34, 39] among others. Assume $\mathbb{E}X_i = 0$. We then consider the batched mean estimate:

$$\hat{\Sigma} = \frac{1}{Mw}\sum_{b=1}^{w} Y_b Y_b^\top = \frac{1}{Mw}\sum_{b=1}^{w}\Big(\sum_{i\in L_b} X_i\Big)\Big(\sum_{i\in L_b} X_i\Big)^\top, \tag{5.1}$$

where the window $L_b = \{1 + (b-1)M, \dots, bM\}$, $b = 1, \dots, w$, the window size $|L_b| = M \to \infty$ and the number of blocks $w = \lfloor n/M \rfloor$. Theorems 5.1 and 5.2 concern the convergence of the above estimate for processes with finite polynomial and finite sub-exponential dependence adjusted norms, respectively. The convergence rate depends in a subtle way on the temporal dependence characterized by α [cf. (2.5)], the uniform and the overall dependence adjusted norms $\Psi_{q,\alpha}$ and $\Upsilon_{q,\alpha}$, respectively, the same size n and the dimension p.

THEOREM 5.1. *Assume $\Psi_{q,\alpha} < \infty$ with $q > 4$, $\alpha > 0$, and $M = O(n^{\varsigma})$ for some $0 < \varsigma < 1$. Let $F_\alpha = wM$ (resp., $wM^{q/2-\alpha q/2}$ or $w^{q/4-\alpha q/2} M^{q/2-\alpha q/2}$) for $\alpha > 1 - 2/q$ (resp., $1/2 - 2/q < \alpha < 1 - 2/q$ or $\alpha < 1/2 - 2/q$). Then for $x \geq \sqrt{w} M \Psi_{q,\alpha}^2$, we have*

$$\mathbb{P}\big(n\big|\mathrm{diag}(\hat{\Sigma}) - \mathbb{E}\,\mathrm{diag}(\hat{\Sigma})\big|_\infty \geq x\big) \lesssim \frac{F_\alpha \Upsilon_{q,\alpha}^q}{x^{q/2}} + p \exp\biggl(-\frac{C_{q,\alpha} x^2}{w M^2 \Psi_{4,\alpha}^4}\biggr), \tag{5.2}$$

$$\mathbb{P}\big(n|\hat{\Sigma} - \mathbb{E}\hat{\Sigma}|_\infty \geq x\big) \lesssim \frac{p F_\alpha \Upsilon_{q,\alpha}^q}{x^{q/2}} + p^2 \exp\biggl(-\frac{C_{q,\alpha} x^2}{w M^2 \Psi_{4,\alpha}^4}\biggr) \tag{5.3}$$

for all large n, where the constants in $\lesssim$ only depend on ς, α and q.

Under stronger moment conditions, we can have an exponential inequality.

THEOREM 5.2. *Assume $\Phi_{\psi_\nu,0} < \infty$ for some $\nu \geq 0$. Then for all $x > 0$, we have*

$$\mathbb{P}\big(n\big|\mathrm{diag}(\hat{\Sigma}) - \mathbb{E}\,\mathrm{diag}(\hat{\Sigma})\big|_\infty \geq x\big) \lesssim p \exp\biggl(-\frac{x^\gamma}{4e\gamma(\sqrt{w} M \Phi_{\psi_\nu,0}^2)^\gamma}\biggr), \tag{5.4}$$

$$\mathbb{P}\big(n|\hat{\Sigma} - \mathbb{E}\hat{\Sigma}|_\infty \geq x\big) \lesssim p^2 \exp\biggl(-\frac{x^\gamma}{4e\gamma(\sqrt{w} M \Phi_{\psi_\nu,0}^2)^\gamma}\biggr), \tag{5.5}$$

where $\gamma = 1/(1 + 2\nu)$ and the constants in $\lesssim$ only depend on ν.

REMARK 4. An alternative estimate of Σ, which also works with unknown mean $\mathbb{E}X_i$, is

$$\tilde{\Sigma} = \frac{1}{wM} \sum_{b=1}^{w} \biggl(\sum_{i \in L_b} X_i - M\bar{X}\biggr)\biggl(\sum_{i \in L_b} X_i - M\bar{X}\biggr)^\top, \tag{5.6}$$

where $\bar{X} = (wM)^{-1} \sum_{i=1}^{wM} X_i$, $w = \lfloor n/M \rfloor$. Then $|\tilde{\Sigma} - \hat{\Sigma}|_\infty = M|\bar{X}|_\infty^2$. Applying Lemma C.2 to $\sum_{i=1}^{wM} X_{ij}$, one can conclude that Theorems 5.1 and 5.2 still hold for $\tilde{\Sigma}$ with $\mathbb{E}\hat{\Sigma}$ therein replaced by $\Sigma_M := \sum_{i=-M}^{M} (1 - |i|/M)\Gamma_i$ (which equals to $\mathbb{E}\hat{\Sigma}$ if $\mathbb{E}X_i = 0$).

COROLLARY 5.3. (i) *Under conditions in Theorem 5.1, we have $|\tilde{\Sigma} - \Sigma|_\infty = O_{\mathbb{P}}(R_n)$, where*

$$\begin{aligned} R_n = n^{-1} \max\{p^{2/q} F_\alpha^{2/q} \Upsilon_{q,\alpha}^2, \sqrt{w} M \Psi_{4,\alpha}^2 \sqrt{\log p}, \sqrt{w} M \Psi_{q,\alpha}^2\} \\ + \Psi_{2,0} \Psi_{2,\alpha} v(M), \end{aligned} \tag{5.7}$$

with $v(M) = 1/M$ *if* $\alpha > 1$, $v(M) = (\log M)/M$ *if* $\alpha = 1$ *and* $v(M) = 1/M^{\alpha}$ *if* $0 < \alpha < 1$. (ii) *Under conditions in Theorem* 5.2, *we have* $|\tilde{\Sigma} - \Sigma|_{\infty} = O_{\mathbb{P}}(R_n^*)$ *with*

$$R_n^* = n^{-1}\sqrt{w}M\Phi^2_{\psi_\nu,0}(\log p)^{1/\gamma} + \Psi_{2,0}\Psi_{2,\alpha}v(M). \tag{5.8}$$

The above corollary easily follows from Theorems 5.1 and 5.2 since the bias $|\Sigma_M - \Sigma|_{\infty} \lesssim \Psi_{2,0}\Psi_{2,\alpha}v(M)$; see the proof of Lemma 7.3.

5.1. *Computing approximated cutoff values.* To apply the Gaussian approximation (3.1) for hypothesis testing or construction of simultaneous confidence intervals, we need to compute χ_θ, the θth quantile of $|D_0^{-1}Z|_{\infty}$, $0 < \theta < 1$. The latter can be computed by simulation if the long-run covariance matrix Σ is known. When it is unknown, we shall use the estimate $\tilde{\Sigma}$ in (5.6). Let $\tilde{D}_0 = [\mathrm{diag}(\tilde{\Sigma})]^{1/2}$. We estimate χ_θ by $\tilde{\chi}_\theta$, the conditional θ-quantile of $|\tilde{D}_0^{-1}\tilde{\Sigma}^{1/2}\eta|_{\infty}$ given $(X_i)_{i=1}^n$, where $\eta \sim N(0, \mathrm{Id}_p)$ is independent of $(X_i)_{i=1}^n$. Note that $\tilde{\chi}_\theta$ can be computed by extensive simulations. This is a Gaussian multiplier resampling method using estimated long-run covariance matrices. Given the level $\alpha \in (0, 1)$, we can reject the null hypothesis $H_0 : \mu = \mu_0$ at level α if $\sqrt{n}|\tilde{D}_0^{-1}(\bar{X}_n - \mu_0)|_{\infty} > \tilde{\chi}_{1-\alpha}$. The $(1-\alpha)$th simultaneous confidence intervals for $\mu = (\mu_1, \dots, \mu_p)^{\top}$ can be constructed as $\hat{\mu}_j \pm \tilde{\chi}_{1-\alpha}\tilde{\sigma}_{jj}^{1/2}/\sqrt{n}$, $1 \le j \le p$. Corollary 5.4 concerns validity of this approach.

COROLLARY 5.4. (i) *Let conditions of Theorem* 3.2 *and Theorem* 5.1 *be satisfied. Further assume* $R_n \log^2 p \to 0$ *with* R_n *given by* (5.7). *Then*

$$\sup_{\theta \in (0,1)} \left|\mathbb{P}\left(\sqrt{n}\left|\tilde{D}_0^{-1}\bar{X}_n\right|_{\infty} \ge \tilde{\chi}_{1-\theta}\right) - \theta\right| \to 0. \tag{5.9}$$

(ii) *Under conditions of Theorem* 3.3 *and Theorem* 5.2, *if* $R_n^* \log^2 p \to 0$ *with* R_n^* *given by* (5.8), *we have* (5.9).

PROOF. (i) Recall (3.1) for ρ_n. Let $\Lambda_n = \sqrt{n}|(\tilde{D}_0^{-1} - D_0^{-1})\bar{X}_n|_{\infty}$. By the triangle inequality and Theorem 3 of [12], for $w > 0$, we have

$$\begin{aligned}
\tilde{\rho}_n &:= \sup_{u \in \mathbb{R}}\left|\mathbb{P}\left(\sqrt{n}\left|\tilde{D}_0^{-1}\bar{X}_n\right|_{\infty} \ge u\right) - \mathbb{P}\left(\left|D_0^{-1}Z\right|_{\infty} \ge u\right)\right| \\
&\le \rho_n + \sup_{u \in \mathbb{R}}\mathbb{P}\left(\left||D_0^{-1}Z|_{\infty} - u\right| \le w\right) + \mathbb{P}(\Lambda_n \ge w) \\
&\lesssim \rho_n + w\sqrt{\log p} + \mathbb{P}(\Lambda_n \ge w).
\end{aligned}$$

Let $V_n = \max_{1 \le j \le p} |(\sigma_{jj}/\tilde{\sigma}_{jj})^{1/2} - 1|$ and $L_n = \max_{1 \le j \le p} |\sigma_{jj} - \tilde{\sigma}_{jj}|$. Then $\Lambda_n \le V_n\sqrt{n}|D_0^{-1}\bar{X}_n|_{\infty}$. Let c be the constant in Assumption 3.1. On the event

$\mathcal{A}_0 = \{L_n \le x\}$ for $x \le c/2$, we have $V_n \le 2L_n/c$. Hence,

$$\begin{aligned}\mathbb{P}(\Lambda_n \ge w) &\le \mathbb{P}(V_n \ge 2x/c) + \mathbb{P}\big(\sqrt{n}|D_0^{-1}\bar{X}_n|_\infty \ge cy/2\big)\\ &\le \mathbb{P}(L_n \ge x) + \rho_n + \mathbb{P}\big(|D_0^{-1}Z|_\infty \ge cy/2\big),\end{aligned}$$

where $w = xy$, $0 < x < c/2$, $y > 0$. It follows that

$$\tilde{\rho}_n \lesssim \rho_n + xy\sqrt{\log p} + \mathbb{P}(L_n \ge x) + \mathbb{P}\big(|D_0^{-1}Z|_\infty \ge cy/2\big).$$

We let $y = C\sqrt{\log p}$, where $C > 0$ is a sufficiently large constant. Note that the marginal variances of $D_0^{-1}Z$ are 1. Let

$$r_n = \frac{1}{n}\max\{F_\alpha^{2/q}\Upsilon_{q,\alpha}^2, \sqrt{w}M\Psi_{4,\alpha}^2\sqrt{\log p}, \sqrt{w}M\Psi_{q,\alpha}^2\} + \Psi_{2,0}\Psi_{2,\alpha}v(M).$$

Let $x = r_n\sqrt{\log p}$. Since $R_n \log^2 p \to 0$ and $r_n \le R_n$, by Corollary 5.3, we have $\mathbb{P}(\mathcal{A}_0) \to 1$. Theorem 3.2 ensures $\rho_n \to 0$. Hence, $\tilde{\rho}_n \to 0$.

Let $T_n = |\tilde{\Sigma} - \Sigma|_\infty$ and $W_n = \max_{1\le j\le p}|\tilde{\sigma}_{jj}/\sigma_{jj} - 1|$. By the elementary inequality $|1 - \sqrt{ab}| \le |1-a| + (1-a)^2 + |1-b| + (1-b)^2$, we have

$$\begin{aligned}|\tilde{D}_0^{-1}\tilde{\Sigma}\tilde{D}_0^{-1} - D_0^{-1}\Sigma D_0^{-1}|_\infty &\le \max_{1\le j,k\le p}\left(\left|\frac{\tilde{\sigma}_{jk} - \sigma_{jk}}{\sqrt{\sigma_{jj}\sigma_{kk}}}\right| + \left|1 - \frac{\sqrt{\tilde{\sigma}_{jj}\tilde{\sigma}_{kk}}}{\sqrt{\sigma_{jj}\sigma_{kk}}}\right|\right)\\ &\le \frac{T_n}{c} + 2W_n + 2W_n^2 \le \frac{3T_n}{c} + \frac{2T_n^2}{c^2}.\end{aligned} \tag{5.10}$$

Let event $\mathcal{A} = \{T_n \le z_n\}$ where $z_n = R_n^{1/2}/\log p$. Since $R_n \log^2 p \to 0$, we have $z_n/R_n \to \infty$. By Corollary 5.3, $\mathbb{P}(\mathcal{A}) \to 1$. Since $z_n \to 0$, by (5.10) and following the arguments of Theorem 3.1 in [10], we have

$$\sup_{\theta\in(0,1)}\big|\mathbb{P}(\sqrt{n}|\tilde{D}_0^{-1}\bar{X}_n|_\infty \ge \tilde{\chi}_{1-\theta}) - \theta\big| \lesssim \tilde{\rho}_n + \pi\left(\frac{3z_n}{c} + \frac{2z_n^2}{c^2}\right) + \mathbb{P}(T_n \ge z_n),$$

where $\pi(z) = z^{1/3}(1 \vee \log(p/z))^{2/3}$. Since $R_n \log^2 p \to 0$, (5.9) follows.

(ii) The proof is similar to (i), and thus is omitted. □

6. Inequalities for high-dimensional time series with finite polynomial moments. Tail probability inequalities play an important role in simultaneous inference. In this section, we shall derive powerful tail probability inequalities for high-dimensional stationary vectors; cf. Theorems 6.1 and 6.2. They are of independent interest. The proofs require Theorem 4.1 of [31], a deep Rosenthal–Burkholder-type bound on moments of Banach-spaced martingales, and Lemma C.6, a Fuk–Nagaev-type inequality for the sum of independent random vectors. We refer the readers to Appendix C for tail probability inequalities in the one-dimensional case under finite polynomial or exponential moment conditions.

Let X_i be a mean zero p-dimensional stationary process and $T_n = \sum_{i=1}^n X_i$, $T_{n,m} = \sum_{i=1}^n X_{i,m}$ where $X_{i,m} = \mathbb{E}(X_i|\varepsilon_{i-m}, \dots, \varepsilon_i)$. We are interested in bounding the tail probabilities of $\mathbb{P}(|T_n - T_{n,m}|_\infty \ge x)$ and $\mathbb{P}(|T_n|_\infty \ge x)$ for large x. Write $\ell = \ell(p) = 1 \vee \log p$.

THEOREM 6.1. *Assume* $\||X.|_\infty\|_{q,\alpha} < \infty$, *where* $q > 2$ *and* $\alpha \ge 0$, *and* $\Psi_{2,\alpha} < \infty$:

(i) *If* $\alpha > 1/2 - 1/q$, *for* $x \gtrsim \sqrt{n\ell}\Psi_{2,\alpha} m^{-\alpha} + n^{1/q}\ell^{3/2}\||X.|_\infty\|_{q,\alpha} m^{1/2-1/q-\alpha}$,

$$\mathbb{P}\big(|T_n - T_{n,m}|_\infty \ge x\big) \lesssim \frac{n\ell^{q/2}\||X.|_\infty\|_{q,\alpha}^q}{m^{\alpha q+1-q/2}x^q} + \exp\biggl(-\frac{C_{q,\alpha}x^2 m^{2\alpha}}{n\Psi_{2,\alpha}^2}\biggr)$$

holds for all $1 \le m \le n$, *where the constant in* $\lesssim$ *only depends on* q *and* α.

(ii) *If* $0 < \alpha < 1/2 - 1/q$, *then for* $x \gtrsim \sqrt{n\ell}\Psi_{2,\alpha} m^{-\alpha} + n^{1/2-\alpha}\ell^{3/2}\||X.|_\infty\|_{q,\alpha}$,

$$\mathbb{P}\big(|T_n - T_{n,m}|_\infty \ge x\big) \lesssim \frac{n^{q/2-\alpha q}\ell^{q/2}\||X.|_\infty\|_{q,\alpha}^q}{x^q} + \exp\biggl(-\frac{C_{q,\alpha}x^2 m^{2\alpha}}{n\Psi_{2,\alpha}^2}\biggr).$$

PROOF. Let $s = \ell = 1 \vee \log p$. Then $\mathbb{P}(|T_n - T_{n,m}|_\infty \ge x)$ is equivalent to $\mathbb{P}(|T_n - T_{n,m}|_s \ge x)$, since for any vector $v = (v_1, \dots, v_p)^\top$, $|v|_\infty \le |v|_s \le p^{1/s}|v|_\infty$. Let $L = \lfloor(\log n - \log m)/(\log 2)\rfloor$, $\varpi_l = 2^l$ if $1 \le l < L$, $\varpi_L = \lfloor n/m \rfloor$ and $\tau_l = m\varpi_l$ for $1 \le l < L$, $\tau_0 = m$, $\tau_L = n$. Define $M_{n,l} = T_{n,\tau_l} - T_{n,\tau_{l-1}}$ for $1 \le l \le L$ and write

$$T_n - T_{n,m} = T_n - T_{n,n} + \sum_{l=1}^L M_{n,l}. \tag{6.1}$$

Notice that $T_n - T_{n,n} = \sum_{j=n}^\infty T_{n,j+1} - T_{n,j}$. By Lemma C.5,

$$\big\||T_n - T_{n,n}|_s\big\|_q \le \sum_{j=n}^\infty \big\||T_{n,j+1} - T_{n,j}|_s\big\|_q \le \sum_{j=n}^\infty C_q(ns)^{1/2}\omega_{j+1,q},$$

where the constant C_q only depends on q. By Markov's inequality, we have

$$\mathbb{P}\big(|T_n - T_{n,n}|_s \ge x\big) \le \frac{\||T_n - T_{n,n}|_s\|_q^q}{x^q} \le \frac{C_q(ns)^{q/2}\Omega_{n+1,q}^q}{x^q}. \tag{6.2}$$

For each $1 \le l \le L$, define

$$Y_{i,l} = \sum_{k=(i-1)\tau_l+1}^{(i\tau_l)\wedge n} (X_{k,\tau_l} - X_{k,\tau_{l-1}}), \qquad \text{for } 1 \le i \le \lfloor n/\tau_l \rfloor;$$

$$R_{n,l}^e = \sum_{i \text{ is even}} Y_{i,l} \quad \text{and} \quad R_{n,l}^o = \sum_{i \text{ is odd}} Y_{i,l}.$$

Let $c = q/2 - 1 - \alpha q$; let $\lambda_l = l^{-2}/(\pi^2/3)$ if $1 \le l \le L/2$ and $\lambda_l = (L+1-l)^{-2}/(\pi^2/3)$ if $L/2 < l \le L$. Since $Y_{i,l}$ and $Y_{i',l}$ are independent for $|i - i'| > 1$, by Lemma C.6, for any $x > 0$,

$$\begin{aligned}&\mathbb{P}(|R^e_{n,l}|_s - 2\mathbb{E}|R^e_{n,l}|_s \ge \lambda_l x)\\&\quad\le \frac{C_q \sum_{i \text{ is even}} \mathbb{E}|Y_{i,l}|_s^q}{(\lambda_l x)^q} + \exp\left(-\frac{(\lambda_l x)^2}{3\sum_{i \text{ is even}} |\sigma_{Y_i,l}|_s^2}\right),\end{aligned}$$

where $\sigma_{Y_i,l} = (\|Y_{i1,l}\|_2, \dots, \|Y_{ip,l}\|_2)^\top$. By Lemma C.5,

$$\big\||Y_{i,l}|_s\big\|_q \le C_q(\tau_l s)^{1/2}\tilde{\omega}_{l,q}, \qquad \text{where } \tilde{\omega}_{l,q} = \sum_{k=\tau_{l-1}+1}^{\tau_l} \omega_{k,q} \le \frac{\||X_\cdot|_\infty\|_{q,\alpha}}{\tau_{l-1}^\alpha}.$$

For $1 \le j \le p$, by Theorem 3.2 of [6],

$$\|Y_{ij,l}\|_2 \le \sqrt{\tau_l}\tilde{\delta}_{l,2,j}, \qquad \text{where } \tilde{\delta}_{l,2,j} = \sum_{k=\tau_{l-1}+1}^{\tau_l} \delta_{k,2,j} \le \frac{\|X_{\cdot j}\|_{2,\alpha}}{\tau_{l-1}^\alpha},$$

which implies $|\sigma_{Y_i,l}|_s \lesssim \tau^{1/2}\tau_{l-1}^{-\alpha}\Psi_{2,\alpha}$. So, we obtain

$$\begin{aligned}&\mathbb{P}(|R^e_{n,l}|_s - 2\mathbb{E}|R^e_{n,l}|_s \ge \lambda_l x)\\&\quad\le \frac{C_1 n s^{q/2}}{x^q} \cdot \frac{\tau_l^{q/2-1}\tilde{\omega}_{l,q}^q}{\lambda_l^q} + \exp\left(-\frac{C_2(\lambda_l x)^2\tau_{l-1}^{2\alpha}}{n\Psi_{2,\alpha}^2}\right).\end{aligned} \tag{6.3}$$

By Lemma 8 in [12],

$$\begin{aligned}\mathbb{E}|R^e_{n,l}|_s &\lesssim \sqrt{ns}\,\tau_{l-1}^{-\alpha}\Psi_{2,\alpha} + n^{1/q}s^{3/2}\tau_l^{1/2-1/q}\tilde{\omega}_{l,q}\\&\lesssim \frac{\sqrt{ns}\Psi_{2,\alpha}}{(m\varpi_l)^\alpha} + \frac{n^{1/q}s^{3/2}\||X_\cdot|_\infty\|_{q,\alpha}}{(m\varpi_l)^{-c/q}}.\end{aligned}$$

Notice that $\lambda_l^{-1}(m\varpi_l)^{c/q} \lesssim n^{c/q}$ for $c > 0$ and $\min_{l\ge 0}\lambda_l\varpi_l^{-c/q} > 0$ for $c < 0$, and $\min_{l\ge 0}\lambda_l\varpi_l^\alpha > 0$. Hence, $\mathbb{E}|R^e_{n,l}|_s \lesssim \lambda_l x$ always holds and (6.3) implies

$$\mathbb{P}(|R^e_{n,l}|_s \ge \lambda_l x) \le \frac{C_1 n s^{q/2}}{x^q} \cdot \frac{\tau_l^{q/2-1}\tilde{\omega}_{l,q}^q}{\lambda_l^q} + \exp\left(-\frac{C_2(\lambda_l x)^2\tau_{l-1}^{2\alpha}}{n\Psi_{2,\alpha}^2}\right). \tag{6.4}$$

A similar inequality holds for $R^o_{n,l}$. Let

$$A = \sum_{l=1}^L \frac{\varpi_l^c}{\lambda_l^q} \quad \text{and} \quad B = \sum_{l=1}^L \exp\left(-\frac{C_5 x^2\lambda_l^2\varpi_l^{2\alpha}}{nm^{-2\alpha}\Psi_{2,\alpha}^2}\right).$$

Since $\sum_{l=1}^{L} \lambda_l \le 1$ and $|M_{n,l}|_s \le |R_{n,l}^e|_s + |R_{n,l}^o|_s$, by (6.4),

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{l=1}^{L} M_{n,l}\right|_s \ge 2x\right) &\le \sum_{l=1}^{L} \mathbb{P}(|M_{n,l}|_s \ge 2\lambda_l x) \\ &\le \sum_{l=1}^{L} [\mathbb{P}(|R_{n,l}^e|_s \ge \lambda_l x) + \mathbb{P}(|R_{n,l}^o|_s \ge \lambda_l x)] \\ &\le \frac{C_3 n m^c s^{q/2} \||X_\cdot|_\infty\|_{q,\alpha}^q}{x^q} A + C_4 B.\end{aligned} \tag{6.5}$$

Let $\nu := \min_{l \ge 1} \lambda_l^2 \varpi_l^{2\alpha} > 0$. By the definition of ϖ_l and λ_l and by elementary calculations, there exists a constant $C_6 > 1$ such that for all $t \ge 1$,

$$\sum_{l=1}^{L} \exp(-C_5 t \lambda_l^2 \varpi_l^{2\alpha}) \le C_6 \exp(-C_5 t \nu). \tag{6.6}$$

If $c > 0$, it can be obtained that $A \le C_7 \varpi_L^c \le C_7 n^c / m^c$. If $c < 0$, then $A \le C_8$. Hence, combining (6.1), (6.2), (6.5), (6.6), Theorem 6.1 follows. □

THEOREM 6.2. *Assume* $\||X_\cdot|_\infty\|_{q,\alpha} < \infty$, *where* $q > 2$ *and* $\alpha \ge 0$, *and* $\Psi_{2,\alpha} < \infty$: (i) *If* $\alpha > 1/2 - 1/q$, *then for* $x \gtrsim \sqrt{n\ell}\Psi_{2,\alpha} + n^{1/q}\ell^{3/2}\||X_\cdot|_\infty\|_{q,\alpha}$,

$$\mathbb{P}(|T_n|_\infty \ge x) \le \frac{C_{q,\alpha} n \ell^{q/2} \||X_\cdot|_\infty\|_{q,\alpha}^q}{x^q} + C_{q,\alpha} \exp\left(-\frac{C_{q,\alpha} x^2}{n \Psi_{2,\alpha}^2}\right). \tag{6.7}$$

(ii) *If* $0 < \alpha < 1/2 - 1/q$, *then for* $x \gtrsim \sqrt{n\ell}\Psi_{2,\alpha} + n^{1/2-\alpha}\ell^{3/2}\||X_\cdot|_\infty\|_{q,\alpha}$,

$$\mathbb{P}(|T_n|_\infty \ge x) \le \frac{C_{q,\alpha} n^{q/2-\alpha q} \ell^{q/2} \||X_\cdot|_\infty\|_{q,\alpha}^q}{x^q} + C_{q,\alpha} \exp\left(-\frac{C_{q,\alpha} x^2}{n \Psi_{2,\alpha}^2}\right). \tag{6.8}$$

PROOF. The proof is similar to that of Theorem 6.1, and thus is omitted. □

7. Proofs of Theorem 3.2 and Theorem 3.3. The main result in this section is Theorem 7.4, which provides an error bound of the Gaussian approximation. Theorems 3.2 and 3.3 follow from Theorem 7.4.

7.1. *An error bound of the Gaussian approximation.* We shall apply the m-dependence approximation approach. For $m \ge 0$, define

$$X_{i,m} = (X_{i1,m}, \ldots, X_{ip,m})^\top = \mathbb{E}(X_i | \varepsilon_{i-m}, \varepsilon_{i-m+1}, \ldots, \varepsilon_i). \tag{7.1}$$

Write $T_X = \sum_{i=1}^n X_i$ and $T_{X,m} = \sum_{i=1}^n X_{i,m}$. For simplicity, suppose $n = (M + m)w$, where $M \gg m$ and $M, m, w \to \infty$ (to be determined) as $n \to \infty$. We apply the block technique and split the interval $[1, n]$ into alternating large blocks $L_b =$

$[(b-1)(M+m)+1, bM+(b-1)m]$ and small blocks $S_b = [bM+(b-1)m+1, b(M+m)]$, $1 \le b \le w$. Let

$$Y_b = \sum_{i \in L_b} X_i, \qquad Y_{b,m} = \sum_{i \in L_b} X_{i,m}, \qquad T_Y = \sum_{b=1}^{w} Y_b, \qquad T_{Y,m} = \sum_{b=1}^{w} Y_{b,m}.$$

Let Z_b, $1 \le b \le w$, be i.i.d. $N(0, MB)$ and $Z_{b,m}$ be i.i.d. $N(0, M\tilde{B})$, where the covariance matrices B and $\tilde{B}$ are respectively given by

$$B = (b_{ij})_{i,j=1}^{p} = \operatorname{Cov}(Y_b/\sqrt{M}) \quad \text{and} \quad \tilde{B} = (\tilde{b}_{ij})_{i,j=1}^{p} = \operatorname{Cov}(Y_{b,m}/\sqrt{M}). \tag{7.2}$$

Write $T_{Z,m} = \sum_{b=1}^{w} Z_{b,m}$ and let $Z \sim N(0, \Sigma)$.

LEMMA 7.1. (i) *Assume $\Theta_{q,\alpha} < \infty$ for some $q > 2$ and $\alpha > 0$. Then there exists some constant $C_{q,\alpha}$ such that for $y > 0$*

$$\mathbb{P}(|T_X - T_{Y,m}|_\infty \ge y) \lesssim f_1^*(y) + f_2^*(y) =: f^*(y) \tag{7.3}$$

where the constant in $\lesssim$ only depends on q and α,

$$f_1^*(y) = \begin{cases} y^{-q} n m^{q/2-1-\alpha q} \Theta_{q,\alpha}^q + p \exp\Big(-\dfrac{C_{q,\alpha} y^2 m^{2\alpha}}{n \Psi_{2,\alpha}^2}\Big), \\ \qquad \alpha > 1/2 - 1/q, \\ y^{-q} n^{q/2-\alpha q} \Theta_{q,\alpha}^q + p \exp\Big(-\dfrac{C_{q,\alpha} y^2 m^{2\alpha}}{n \Psi_{2,\alpha}^2}\Big), \\ \qquad \alpha < 1/2 - 1/q, \end{cases} \tag{7.4}$$

and

$$f_2^*(y) = \begin{cases} y^{-q} w m \Theta_{q,\alpha}^q + p \exp\Big(-\dfrac{C_{q,\alpha} y^2}{m w \Psi_{2,\alpha}^2}\Big), \\ \qquad \alpha > 1/2 - 1/q, \\ y^{-q} (wm)^{q/2-\alpha q} \Theta_{q,\alpha}^q + p \exp\Big(-\dfrac{C_{q,\alpha} y^2}{w m \Psi_{2,\alpha}^2}\Big), \\ \qquad \alpha < 1/2 - 1/q. \end{cases} \tag{7.5}$$

(ii) *Assume $\Phi_{\psi_\nu,\alpha} < \infty$ for some $\nu \ge 0$ and $\alpha > 0$. Let $\beta = 2/(1+2\nu)$. Then there exists a constant $C_\beta > 0$ such that for $y > 0$,*

$$\mathbb{P}(|T_X - T_{Y,m}|_\infty \ge y) \lesssim f_1^\diamond(y) + f_2^\diamond(y) =: f^\diamond(y), \tag{7.6}$$

where the constant in $\lesssim$ only depends on β and α,

$$f_1^\diamond(y) = p \exp\Big\{-C_\beta \Big(\frac{y m^\alpha}{\sqrt{n}\Phi_{\psi_\nu,\alpha}}\Big)^\beta\Big\},$$

$$f_2^\diamond(y) = p \exp\Big\{-C_\beta \Big(\frac{y}{\sqrt{mw}\Phi_{\psi_\nu,0}}\Big)^\beta\Big\}.$$

LEMMA 7.2. *Let $D = (d_{ij})_{i,j=1}^p$ be a diagonal matrix. Assume that there exist constants $c > 0$, $c_2 > c_1 > 0$ such that $c < \min_{1 \le j \le p} d_{jj}$ and $c_1 \le \tilde{b}_{jj}/d_{jj} \le c_2$ for all $1 \le j \le p$. Assume $\Psi_{q,0} < \infty$ for some $q \ge 4$. Then for all $\lambda \in (0, 1)$,*

$$\begin{aligned}
&\sup_{t\in\mathbb{R}} \bigl|\mathbb{P}\bigl(\bigl|D^{-1/2}T_{Y,m}/\sqrt{n}\bigr|_\infty \le t\bigr) - \mathbb{P}\bigl(\bigl|D^{-1/2}T_{Z,m}/\sqrt{n}\bigr|_\infty \le t\bigr)\bigr| \\
&\quad \lesssim w^{-1/8}\bigl(\Psi_{3,0}^{3/4} \vee \Psi_{4,0}^{1/2}\bigr)\bigl(\log(pw/\lambda)\bigr)^{7/8} + w^{-1/2}\bigl(\log(pw/\lambda)\bigr)^{3/2} u_m(\lambda) + \lambda \\
&\quad =: h\bigl(\lambda, u_m(\lambda)\bigr),
\end{aligned}$$

where the constant in $\lesssim$ depends on c, c_1, c_2 and q and α for (i), *and β for* (ii) *below, and $u_m(\lambda) \le u_m^*(\lambda)$ in* (i), *and $u_m(\lambda) \le u_m^\diamond(\lambda)$ in* (ii).

(i) *Assume $\Theta_{q,\alpha} < \infty$ for some $q \ge 4$ and $\alpha > 0$, then*

$$u_m^*(\lambda) = \begin{cases} \max\bigl\{\Theta_{q,\alpha}\bigl(\lambda^{-1}w\bigr)^{1/q} M^{1/q-1/2}, \Psi_{2,\alpha}\sqrt{\log(pw/\lambda)}\bigr\}, \\ \qquad \alpha > 1/2 - 1/q, \\ \max\bigl\{\Theta_{q,\alpha}\bigl(\lambda^{-1}w\bigr)^{1/q} M^{-\alpha}, \Psi_{2,\alpha}\sqrt{\log(pw/\lambda)}\bigr\}, \\ \qquad \alpha < 1/2 - 1/q. \end{cases} \tag{7.7}$$

(ii) *Assume $\Phi_{\psi_\nu,0} < \infty$ for some $\nu \ge 0$. Then*

$$u_m^\diamond(\lambda) = \max\bigl\{\Phi_{\psi_\nu,0}\bigl(\log(pw/\lambda)\bigr)^{1/\beta}, \sqrt{\log(pw/\lambda)}\bigr\}. \tag{7.8}$$

LEMMA 7.3. *Assume $\Psi_{2,\alpha} < \infty$ for some $\alpha > 0$. Let $D = (d_{ij})_{i,j=1}^p$ be a diagonal matrix such that there exist some constants $0 < C_1 < C_2$ such that $C_1 \le \sigma_{jj}/d_{jj} \le C_2$ for all $1 \le j \le p$. Then we have*

$$\begin{aligned}
&\sup_{t\in\mathbb{R}} \bigl|\mathbb{P}\bigl(\bigl|D^{-1/2}T_{Z,m}/\sqrt{n}\bigr|_\infty \le t\bigr) - \mathbb{P}\bigl(\bigl|D^{-1/2}Z\bigr|_\infty \le t\bigr)\bigr| \\
&\quad \lesssim \pi\Bigl(\max_{1\le j\le p} d_{jj}^{-1}\Psi_{2,\alpha}\Psi_{2,0}\bigl(m^{-\alpha} + v(M)\bigr) + wm/n\Bigr),
\end{aligned}$$

where $\pi(x) = x^{1/3}(1 \vee \log(p/x))^{2/3}$ for $x > 0$ and $v(M)$ is the same as defined in Corollary 5.3.

THEOREM 7.4. *Let $\Sigma_0 = \operatorname{diag}(\Sigma)$ and $D_0 = \Sigma_0^{1/2}$. Let Assumption* 3.1 *be satisfied.* (i) *Assume $\Theta_{q,\alpha} < \infty$, where $q \ge 4$ and $\alpha > 0$. Let $\chi(m, M) = \Psi_{2,\alpha}\Psi_{2,0}(m^{-\alpha} + v(M)) + wm/n$, where $v(M)$ is given in Corollary* 5.3. *Recall* (3.1) *for ρ_n. Then for every $\lambda \in (0, 1)$ and $\eta > 0$,*

$$\rho_n \lesssim f^*(\sqrt{n}\eta) + \eta\sqrt{\log p} + h\bigl(\lambda, u_m^*(\lambda)\bigr) + \pi\bigl(\chi(m, M)\bigr). \tag{7.9}$$

(ii) *Assume $\Phi_{\psi_\nu,\alpha} < \infty$, where $\nu \ge 0$ and $\alpha > 0$. Then for every $\lambda \in (0, 1)$ and $\eta > 0$,*

$$\rho_n \lesssim f^\diamond(\sqrt{n}\eta) + \eta\sqrt{\log p} + h\bigl(\lambda, u_m^\diamond(\lambda)\bigr) + \pi\bigl(\chi(m, M)\bigr). \tag{7.10}$$

PROOF. (i) By Lemma 7.2(i) and Lemma 7.3, we have for every $\lambda \in (0, 1)$,

$$\begin{aligned} &\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left(\left|D_0^{-1} T_{Y,m}/\sqrt{n}\right|_\infty \le t\right) - \mathbb{P}\left(\left|D_0^{-1} Z\right|_\infty \le t\right)\right| \\ &\quad \lesssim h\left(\lambda, u_m^*(\lambda)\right) + \pi\left(\Psi_{2,\alpha}\Psi_{2,0}\left(m^{-\alpha} + v(M)\right) + wm/n\right). \end{aligned} \tag{7.11}$$

Observe that the Gaussian vector $D_0^{-1}Z$ has marginal variance 1. By Theorem 3 of [12], for every $\eta > 0$,

$$\sup_{t\in\mathbb{R}} \mathbb{P}\left(\left|\left|D_0^{-1} Z\right|_\infty - t\right| \le \eta\right) \lesssim \eta\sqrt{\log p}. \tag{7.12}$$

By the triangle inequality, for every $\eta > 0$, we have

$$\begin{aligned} &\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left(\left|D_0^{-1} T_X/\sqrt{n}\right|_\infty > t\right) - \mathbb{P}\left(\left|D_0^{-1} T_{Y,m}/\sqrt{n}\right|_\infty > t\right)\right| \\ &\quad \le \mathbb{P}\left(\left|D_0^{-1}(T_X - T_{Y,m})/\sqrt{n}\right|_\infty > \eta\right) + \sup_{t\in\mathbb{R}} \mathbb{P}\left(\left|\left|D_0^{-1} T_{Y,m}/\sqrt{n}\right|_\infty - t\right| \le \eta\right), \end{aligned}$$

which implies Theorem 7.4(i) in view of Lemma 7.1(i), (7.11) and (7.12).

(ii) Inequality (7.10) can be obtained by replacing f^* and u_m^* with $f^\diamond$ and $u_m^\diamond$ in the above proof. □

7.2. *Proofs of Theorem* 3.2 *and Theorem* 3.3.

PROOF. Recall (7.3) for $f^*(\cdot)$. By Theorem 7.4, for $\alpha > 1/2 - 1/q$, to have (3.1), we need

$$\pi\left(\Psi_{2,\alpha}\Psi_{2,0}\left(m^{-\alpha} + v(M)\right) + wm/n\right) \to 0 \tag{7.13}$$

and for some $\eta > 0$ and $\lambda \in (0, 1)$,

$$f^*(\sqrt{n}\eta) + \eta\sqrt{\log p} \to 0, \tag{7.14}$$

$$h\left(\lambda, u_m^*(\lambda)\right) \to 0. \tag{7.15}$$

First, (7.13) requires $m \gg L_1$, $wm \ll n(\log p)^{-2}$, $w \ll n(\log p)^{-2}(\Psi_{2,\alpha}\Psi_{2,0})^{-1}$ if $\alpha > 1$ and $w \ll n/L_1$ if $0 < \alpha < 1$. Moreover, (7.14) requires $m \gg \max(L_0, (\Psi_{2,\alpha} \log p)^{1/\alpha})$ with $L_0 = (n^{1/q-1/2}(\log p)^{1/2}\Theta_{q,\alpha})^{1/(\alpha-1/2+1/q)}$ and $wm \ll \min(N_1, N_2)$. And (7.15) needs (3.2) and $w \gg \max(W_1, W_2)$. We also need $M \asymp n/w \gg m$. Notice that $(\Psi_{2,\alpha} \log p)^{1/\alpha} \lesssim L_1$, $N_2 \lesssim n(\log p)^{-2}$, $N_2 \le n(\log p)^{-2}(\Psi_{2,\alpha}\Psi_{2,0})^{-1}$ and under (3.2), $L_0 \to 0$. If

$$L_1 \max(W_1, W_2) = o(1) \min(n, N_1, N_2), \tag{7.16}$$

then we can always choose m and w such that (3.1) holds. Observe that $N_2 \lesssim n$, then (7.16) is reduced to (3.3).

For $0 < \alpha < 1/2 - 1/q$, the function f^* in (7.14) is replaced by $f^\diamond$ [cf. (7.6)], which implies $\Theta_{q,\alpha}(\log p)^{1/2} = o(n^\alpha)$, $m \gg (\Psi_{2,\alpha} \log p)^{1/\alpha}$ and

$wm \ll \min(N_2, N_3)$. And u_m^* in (7.15) is replaced by $u_m^\diamond$, implying $w \gg \max(W_1, W_2, W_3)$. By the similar argument, if (3.4) is further assumed, then (3.1) also holds for the case $0 < \alpha < 1/2 - 1/q$.

The proof of Theorem 3.3 is similar to that of Theorem 3.2, and thus is omitted. □

REMARK 5. In the proof of Theorem 3.2, we exclude the case $\alpha = 1$ when $\alpha > 1/2 - 1/q$. If $\alpha = 1$, we need to impose the additional assumption

$$\max(W_1, W_2) = o\big(n/(L_1 \log n)\big) \tag{7.17}$$

to ensure (7.13). The above condition is very mild since (3.3) implies that $\max(W_1, W_2) = o(n/L_1)$. If $\log n \lesssim (\log p)^2 \Psi_{2,\alpha}^2$, which trivially holds in the high-dimensional case $p \asymp n^\kappa$ with some $\kappa > 0$, we have $N_2 = O(n/\log n)$, and hence (3.3) implies (7.17). Similarly, in Theorem 3.3 we shall further assume $\max(W_1, W_4) = o(n/(L_1 \log n))$ if $\alpha = 1$.

Acknowledgments. We thank two anonymous referees, an Associate Editor and the Editor for their helpful comments that have improved the paper.

SUPPLEMENTARY MATERIAL

Supplement to "Gaussian approximation for high dimensional time series" (DOI: 10.1214/16-AOS1512SUPP; .pdf). This supplemental file contains the additional technical proofs and a simulation study.

REFERENCES

[1] ALEXOPOULOS, C. and GOLDSMAN, D. (2004). To batch or not to batch? *ACM Trans. Model. Comput. Simul.* **14** 76–114.

[2] ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York. MR0283939

[3] AZZALINI, A. (1981). A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika* **68** 326–328. MR0614972

[4] BRADLEY, R. C. (2007). *Introduction to Strong Mixing Conditions*. Kendrick Press.

[5] BÜHLMANN, P. (2002). Bootstraps for time series. *Statist. Sci.* **17** 52–72. MR1910074

[6] BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42. MR0365692

[7] CAI, T. T. and JIANG, T. (2011). Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *Ann. Statist.* **39** 1496–1525. MR2850210

[8] CHEN, X., SHAO, Q. M., WU, W. B. and XU, L. (2015). Self-normalized Cramér type moderate deviations under dependence. Preprint. Available at arXiv:1409.3642.

[9] CHENG, M.-Y. and PENG, L. (2002). Regression modeling for nonparametric estimation of distribution and quantile functions. *Statist. Sinica* **12** 1043–1060. MR1947061

[10] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.* **41** 2786–2819. MR3161448

[11] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2013). Testing many moment inequalities. Preprint. Available at arXiv:1312.7614.
[12] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probab. Theory Related Fields* **162** 47–70. MR3350040
[13] DECO, G., JIRSA, V. K. and MCINTOSH, A. R. (2013). Resting brains never rest: Computational insights into potential cognitive architectures. *Trends Neurosci*. **36** 268–274.
[14] DEDECKER, J., DOUKHAN, P., LANG, G., LEÓN, J. R., LOUHICHI, S. and PRIEUR, C. (2007). *Weak Dependence*: *With Examples and Applications*. *Lecture Notes in Statistics* **190**. Springer, New York. MR2338725
[15] DONOHO, D. L., ELAD, M. and TEMLYAKOV, V. N. (2006). Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Trans*. *Inform*. *Theory* **52** 6–18. MR2237332
[16] FALK, M. (1985). Asymptotic normality of the kernel quantile estimator. *Ann*. *Statist*. **13** 428–433. MR0773180
[17] FERNIQUE, X. (1975). Regularité des trajectoires des fonctions aléatoires gaussiennes. In *École d'Été de Probabilités de Saint-Flour, IV*-1974. *Lecture Notes in Math*. **480** 1–96. Springer, Berlin. MR0413238
[18] HIPP, J. F., HAWELLEK, D. J., CORBETTA, M., SIEGEL, M. and ENGEL, A. K. (2012). Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nat*. *Neurosci*. **15** 884–890.
[19] HUTCHISON, R. M., WOMELSDORF, T., ALLEN, E. A., BANDETTINI, P. A., CALHOUN, V. D., CORBETTA, M., DELLA PENNA, S., DUYN, J. H., GLOVER, G. H., GONZALEZ-CASTILLO, J., HANDWERKER, D. A., KEILHOLZ, S., KIVINIEMI, V., LEOPOLD, D. A., DE PASQUALE, F., SPORNS, O., WALTER, M. and CHANG, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage* **80** 360–378.
[20] IBRAGIMOV, I. A. and LINNIK, YU. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff Publishing, Groningen. MR0322926
[21] JIANG, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *Ann*. *Appl*. *Probab*. **14** 865–880. MR2052906
[22] KOSOROK, M. R. and MA, S. (2007). Marginal asymptotics for the "large p, small n" paradigm: With applications to microarray data. *Ann*. *Statist*. **35** 1456–1486. MR2351093
[23] LAHIRI, S. N. (2003). *Resampling Methods for Dependent Data*. Springer, Berlin.
[24] LARSON-PRIOR, L. J., OOSTENVELD, R., PENNA, S. D., MICHALAREAS, G., PRIOR, F., BABAJANI-FEREMI, A., SCHOFFELEN, J. M., MARZETTI, L., DE PASQUALE, F., POMPEO, F. D., STOUT, J., WOOLRICH, M., LUO, Q., BUCHOLZ, R., FRIES, P., PIZZELLA, V., ROMANI, G. L., CORBETTA, M. and SNYDER, A. Z. (2013). Adding dynamics to the human connectome project with {MEG}. *NeuroImage* **80** 190–201.
[25] LI, D., LIU, W.-D. and ROSALSKY, A. (2010). Necessary and sufficient conditions for the asymptotic distribution of the largest entry of a sample correlation matrix. *Probab. Theory Related Fields* **148** 5–35. MR2653220
[26] LI, D. and ROSALSKY, A. (2006). Some strong limit theorems for the largest entries of sample correlation matrices. *Ann*. *Appl*. *Probab*. **16** 423–447. MR2209348
[27] LIU, W. and WU, W. B. (2010). Asymptotics of spectral density estimates. *Econometric Theory* **26** 1218–1245. MR2660298
[28] LIU, W.-D., LIN, Z. and SHAO, Q.-M. (2008). The asymptotic distribution and Berry–Esseen bound of a new test for independence in high dimension with an application to stochastic optimization. *Ann*. *Appl*. *Probab*. **18** 2337–2366. MR2474539
[29] NAGAEV, S. V. (1979). Large deviations of sums of independent random variables. *Ann*. *Probab*. **7** 745–789. MR0542129

[30] NEWEY, W. K. and WEST, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55** 703–708. MR0890864

[31] PINELIS, I. (1994). Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.* **22** 1679–1706. MR1331198

[32] POLITIS, D. N., ROMANO, J. P. and WOLF, M. (1999). *Subsampling*. Springer, New York. MR1707286

[33] PORTNOY, S. (1986). On the central limit theorem in $\mathbf{R}^p$ when $p \to \infty$. *Probab. Theory Related Fields* **73** 571–583. MR0863546

[34] PRIESTLEY, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press, San Diego.

[35] PRIESTLEY, M. B. (1988). *Non-linear and Non-stationary Time Series Analysis*. Academic Press, San Diego.

[36] REISS, R.-D. (1981). Nonparametric estimation of smooth distribution functions. *Scand. J. Stat.* **8** 116–119. MR0623587

[37] ROSENBLATT, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **42** 43–47. MR0074711

[38] ROSENBLATT, M. (1971). *Markov Processes. Structure and Asymptotic Behavior*. Springer, New York. MR0329037

[39] ROSENBLATT, M. (1985). *Stationary Sequences and Random Fields*. Springer, Berlin.

[40] TONG, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. Oxford Univ. Press, New York. MR1079320

[41] TSAY, R. S. (2005). *Analysis of Financial Time Series*, 2nd ed. Wiley, Hoboken, NJ. MR2162112

[42] WANG, J., CHENG, F. and YANG, L. (2013). Smooth simultaneous confidence bands for cumulative distribution functions. *J. Nonparametr. Stat.* **25** 395–407. MR3056092

[43] WIENER, N. (1958). *Nonlinear Problems in Random Theory*. Wiley, New York. MR0100912

[44] WU, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154 (electronic). MR2172215

[45] WU, W. B. (2011). Asymptotic theory for stationary processes. *Stat. Interface* **0** 1–20.

[46] XIAO, H. and WU, W. B. (2013). Asymptotic theory for maximum deviations of sample covariance matrix estimates. *Stochastic Process. Appl.* **123** 2899–2920. MR3054550

[47] YAMATO, H. (1972/73). Uniform convergence of an estimator of a distribution function. *Bull. Math. Statist.* **15** 69–78. MR0329113

[48] ZHANG, D. and WU, W. B (2016). Supplement to "Gaussian approximation for high dimensional time series." DOI:10.1214/16-AOS1512SUPP.

[49] ZHANG, X. and CHENG, G. (2014). Bootstrapping high dimensional time series. Preprint. Available at arXiv:1406.1037.

[50] ZHOU, W. (2007). Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Trans. Amer. Math. Soc.* **359** 5345–5363. MR2327033

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, SAN DIEGO
9500 GILMAN GRIVE # 0112
LA JOLLA, CALIFORNIA 92093-0112
USA
E-MAIL: daz076@ucsd.edu

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
GEORGE HERBERT JONES LABORATORY
5747 S. ELLIS AVENUE
CHICAGO, ILLINOIS 60637
USA
E-MAIL: wbwu@galton.uchicago.edu