The Annals of Statistics
1994, Vol. 22, No. 4, 2062–2088

ASYMPTOTIC DISTRIBUTION OF STATISTICS IN TIME SERIES

BY F. GÖTZE AND C. HIPP

University of Bielefeld and University of Karlsruhe

Verifiable conditions are given for the validity of formal Edgeworth expansions for the distribution of sums $X_1 + \cdots + X_n$, where $X_i = F(Z_i, \dots, Z_{i+p-1})$ and $Z_1, Z_2, \dots$ is a strict sense stationary sequence that can be written as $Z_j = g(\varepsilon_{j-k}: k \ge 0)$ with an iid sequence (ε_i) of innovations. These models include nonlinear functions of ARMA processes (Z_i) as well as certain nonlinear AR processes. The results apply to many statistics in (nonlinear) time series models.

1. Introduction and summary. Consider a sequence ε_j, $j \in \mathbb{Z}$, of iid random variables and a measurable function $g: \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ as well as a function $h: \mathbb{R}^p \to \mathbb{R}^k$ with a *uniformly* continuous derivative. Define for $j \in \mathbb{Z}$,

$$(1)\qquad \begin{aligned} Z_j &:= g(\varepsilon_{j-i}: i \ge 0)\\ X_j &:= h(Z_j, \dots, Z_{j+p-1}) \quad \text{where } \mathbb{E}\|Dh(Z_i, \dots, Z_{i+p-1})\| \le K \end{aligned}$$

for some constant $K > 0$. This defines strictly stationary sequences of random variables. A representation of this type is possible, for example, for stationary ARMA processes and for certain stationary nonlinear AR processes that will be discussed in the examples below. We are interested in the validity of formal Edgeworth expansions of order $s - 2$ for the distribution of

$$S_n := n^{-1/2}(X_1 + \cdots + X_n - n\mathbb{E}X_1).$$

There are many statistics T_n in time series models such that the following holds: valid Edgeworth expansions for S_n imply valid Edgeworth expansions for T_n.

These statistics include least squares estimators for parameters in the spectral density of a Gaussian ARMA process (see [18–20]), as well as least squares estimators for ARMA parameters in general ARMA processes ([3]) and functions of autocovariances in general ARMA processes ([4]).

The general question of validity of Edgeworth expansions for statistics of the above form is answered in an earlier paper ([5]). For extensions of these results see [9, 10]. However, the verification of the conditions in [5] is not always straightforward. In this paper we state sufficient conditions for the validity of Edgeworth expansions for these problems which can be checked easily and which are even necessary.

The conditions given in [5], pages 216 and 217, specialized to the case (1), read as follows:

$$(2)\qquad \mathbb{E}\|X_1\|^{s+1} < \infty.$$

Received September 1992; revised December 1993.
AMS 1991 *subject classifications.* Primary 62E20; secondary 60F05.
Key words and phrases. Edgeworth expansions, statistics in time series models.

There exist constants $K < \infty$ and $\alpha > 0$ such that for $m \geq 1$,

$$\mathbb{E}\|g(\varepsilon_j: j \geq 0) - g(\varepsilon_0, \dots, \varepsilon_m, 0, \dots)\| \leq K e^{-\alpha m}. \tag{3}$$

Finally, there exist $r > 0$ and $\alpha > 0$ such that for arbitrary large fixed $\kappa > 1$ and all $n > m > \alpha^{-1}$ and $t \in \mathbb{R}^k$ with $n^\kappa > \|t\| > \alpha$,

$$\mathbb{E}\left|\mathbb{E}\left(\exp\left(\sqrt{-1}t^T(X_0 + \dots + X_{2m})\right) \middle| \varepsilon_j: |j - m| \geq r\right)\right| \leq e^{-\alpha}. \tag{4}$$

In fact, we did require in [5] that (3) holds for X_j and furthermore that $\liminf_n \inf_{\|t\| = \alpha} \operatorname{var}(t^T S_n) > 0$. Both conditions are satisfied here as a consequence of the special structure of the r.v.'s Z_j and X_j as defined in (1) and conditions (1)–(4). [See Lemma 2.1, (14) and (15).] The *uniform* continuity of h is a condition which is used solely to prove an approximation like (3) for X_j. To this end one might use other conditions as well. (See the proof of Lemma 2.1.)

Only recently, the moment condition (2) has been relaxed in [12].

The above conditions appear if we choose the σ-fields $\mathcal{D}_j, j \in \mathbb{Z}$, on page 216 in [5] to be generated by $\varepsilon_{j-r}, \varepsilon_{j-r+1}, \dots, \varepsilon_{j+r}$. Notice that conditions (2.4) and (2.6) in [5] are obviously satisfied. The slightly weaker statements using upper bounds on t and m in (4) are an immediate consequence of the actual use of condition (4) in Lemma (3.33) and Lemma (3.43) of [5]. There we did *not* assume that the Z_j are strictly stationary. More generally, in applications for finite time series, the Z_j are often nonstationary in the sense that for some $d \geq 1$,

$$Z_j' := g(\varepsilon_j, \varepsilon_{j-1}, \dots, \varepsilon_{-d}, 0, 0, \dots) \tag{5}$$

and $X_j' := h(Z_i', \dots, Z_{i-p+1}')$ are defined similarly as above. Let

$$S_n' := n^{-1/2}\left[(X_1' - \mathbb{E}X_1') + \dots + (X_n' - \mathbb{E}X_n')\right].$$

For $r = 0, \dots, s$ and $t \in \mathbb{R}^d$ and letting $\chi_{r,n}(t)$ be the cumulant of $\sqrt{-1}t^T S_n$ of order r,

$$\chi_{r,n}(t) = \frac{d^r}{dx^r} \log \mathbb{E} \exp\left[x\sqrt{-1}t^T S_n\right]\Big|_{x=0}.$$

Define the formal Edgeworth expansion $\Psi_{n,s}$ of S_n by its characteristic function

$$\widehat{\Psi}_{n,s}(t) = \exp\left[\chi_{2,n}(t)\right] \sum_{r=0}^{s-2} n^{-r/2} \widetilde{P}_{r,n}(t),$$

where the functions $\widetilde{P}_{r,n}(t)$ are defined by the formal identity

$$\exp\left[\chi_{2,n}(t) + \sum_{r=3}^{\infty} \tau^{r-2} n^{(r-2)/2} \chi_{r,n}(t)\right] = \exp\left[\chi_{2,n}(t)\right] \sum_{r=0}^{\infty} \tau^r \widetilde{P}_{r,n}(t).$$

Similarly, let $\Psi'_{n,s}$ denote the Edgeworth expansion of S'_n. Corollary (2.9) in [5] together with our Remark 2.2 yield that Edgeworth expansions for S_n and for S'_n are valid in the following sense.

THEOREM 1.1. *Under conditions* (2)–(4) *we have uniformly for convex measurable subsets* $C \subset \mathbb{R}^d$,

$$\mathbb{P}\{S_n \in C\} = \Psi_{n,s}(C) + o\big(n^{-(s-2)/2}\big)$$

and

$$\mathbb{P}\{S'_n \in C\} = \Psi'_{n,s}(C) + o\big(n^{-(s-2)/2}\big).$$

The crucial condition to be checked is condition (4), the *conditional Cramér condition*.

Since it is less technical to check (4) in the stationary case, we stick to that case in the proofs. We shall now verify condition (3) for the case of ARMA processes as well as nonlinear AR(1) processes.

EXAMPLE 1.1. Consider the ARMA recursion

$$\sum_{\nu=0}^{d} a_\nu Z_{j-\nu} = \sum_{\nu=0}^{q} b_\nu \varepsilon_{j-\nu}, \qquad j \in \mathbb{Z},$$

where $a_0 = b_0 = 1, d \geq 1$ and the polynomial $a(z) = \Sigma_{\nu=0}^{p} a_\nu z^\nu$ has only zeros outside the unit disk $|z| \leq 1$. Then a stationary sequence $Z_j, j \in \mathbb{Z}$, satisfying the above recursion is given by

$$Z_j = \sum_{i=0}^{\infty} c_i \varepsilon_{j-i}, \qquad j \in \mathbb{Z},$$

where the coefficients c_i are defined via $\Sigma_{\nu=0}^{\infty} c_\nu z^\nu = \Sigma_{\nu=0}^{q} b_\nu z^\nu / a(z)$. Furthermore, the coefficients c_i are exponentially decreasing such that $|c_m| \leq K e^{-\alpha m}$ and therefore condition (3) holds for

$$g(\varepsilon) = \sum_{i=0}^{\infty} c_i \varepsilon_i$$

provided $\mathbb{E}|\varepsilon_1| < \infty$.

Nonlinear AR(p) models of the type

$$Z_{j+1} = f(Z_j, Z_{j-1}, \ldots, Z_{j-d+1}) + \varepsilon_{j+1}, \qquad j \in \mathbb{Z},$$

like the threshold model of Tong [22], the exponential AR of Ozaki [14] and others, have been an active area of research lately. Compare the monographs of Priestley [17] and Tong [23]. The question of stationarity and geometric ergodicity, that is, in our setting conditions (1) and (3), has been studied using the Foster–Tweedie condition in [24] and [13], page 90; see also [21]. These authors assume for $d = 1$ that f is bounded on compact sets and is strictly contracting outside a compact set. Furthermore, the distribution of ε_j is supposed to have a positive density with respect to the Lebesgue measure.

To apply the result of Theorem 1.1 and to simplify the discussion, let us use in the following some stronger conditions for stationarity.

EXAMPLE 1.2. Let $f\colon \mathbb{R}^d \to \mathbb{R}$ be strongly contracting in the sense that for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$,

$$|f(\mathbf{x}) - f(\mathbf{y})| \le \sum_{i=1}^{d} \rho_i |x_i - y_i|,$$

where $\rho_i > 0$ and $\Sigma_{i=1}^{d} \rho_i \le \rho < 1$. To construct a stationary sequence $Z_j, j \in \mathbb{Z}$, satisfying the autoregressive recursion above, define the function $g\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ by

$$g(\varepsilon) := \lim_n g_n(\varepsilon), \tag{6}$$

where $g_0(\varepsilon) = \varepsilon_0$ and

$$g_{n+1}(\varepsilon) := f\Big(g_n(\theta\varepsilon), g_n\big(\theta^2\varepsilon\big), \dots, g_n\big(\theta^d\varepsilon\big)\Big) + \varepsilon_0,$$

with $(\theta\varepsilon)_j = \varepsilon_{j+1}, j \in \mathbb{Z}$. The limit in (6) exists in the L_1-sense if we assume in addition that

$$\mathbb{E}|g_1(\varepsilon) - g_0(\varepsilon)| = \mathbb{E}|f(\varepsilon_1, \dots, \varepsilon_d)| < \infty,$$

since we have for $n \ge 1$,

$$\begin{aligned} \mathbb{E}|g_{n+1}(\varepsilon) - g_n(\varepsilon)| &\le \mathbb{E}\sum_{i=1}^{d} \rho_i |g_n(\theta^i\varepsilon) - g_{n-1}(\theta^i\varepsilon)| \\ &\le \rho\mathbb{E}|g_n(\varepsilon) - g_{n-1}(\varepsilon)|. \end{aligned}$$

The process defined by $Z_j := g(\varepsilon_{j-i}, i \ge 0)$ with g defined in (6) satisfies the autoregressive recursion

$$Z_{j+1} = f(Z_j, Z_{j-1}, \dots, Z_{j-d+1}) + \varepsilon_{j+1}, \qquad j \in \mathbb{Z}.$$

Furthermore, it is strictly stationary and satisfies (3) since

$$\begin{aligned}\mathbb{E}|g(\varepsilon_j: j \geq 0) - g_m(\varepsilon_j: j \geq 0)| &\leq \sum_{i=m}^{\infty} \mathbb{E}|g_i(\varepsilon_j: j \geq 0) - g_{i+1}(\varepsilon_j: j \geq 0)| \\ &\leq \sum_{i=m}^{\infty} \rho^{i-1}\mathbb{E}|f(\varepsilon_1, \dots, \varepsilon_m)| \\ &= \rho^{m-1}\mathbb{E}|f(\varepsilon_1, \dots, \varepsilon_m)|/(1-\rho).\end{aligned}$$

Notice that whenever the random variables Z_j admit a representation (1) satisfying (3) and $h\colon \mathbb{R}^p \to \mathbb{R}^k$ is smooth, then the sequence $X_j := h(Z_j, \dots, Z_{j+p-1})$ also admits a representation (1) satisfying (3); see Lemma 2.1, (14).

For a function $h\colon \mathbb{R}^p \to \mathbb{R}^k$ define for $m \geq p$ and $z \in \mathbb{R}^m$ with $z_i := z_{i-m}$, $m < i \leq m + p - 1$,

$$H_m(z_1, \dots, z_m) := \sum_{i=1}^{m} h(z_i, \dots, z_{i+p-1}). \tag{7}$$

The following result provides conditions such that the smoothness condition (4) holds in cases corresponding to Examples 1.1 and 1.2. The proof of this result is deferred to Section 2.

THEOREM 1.2. *Let $c_j, j \geq 0$, be a sequence of real numbers such that for $K < \infty, \alpha > 0$ we have*

$$|c_m| \leq K e^{-\alpha m}, \qquad m \geq 0.$$

Let $Z_j = \Sigma_{i=0}^{\infty} c_i \varepsilon_{j-i}, j \in \mathbb{Z}$, where $\varepsilon_j, j \in \mathbb{Z}$, denotes an iid sequence such that $\mathbb{E}|\varepsilon_1| < \infty$. Let $h\colon \mathbb{R}^p \to \mathbb{R}^k$ denote a uniform continuously differentiable function satisfying $\mathbb{E}\|h^{(i)}(Z_1, \dots, Z_p)\| < \infty, i = 1, \dots, k$. Assume that ε_1 has a positive continuous density and that the moment condition (2) holds. Then the distribution of $S_n := n^{-1/2}(X_1 + \cdots + X_n - n\mathbb{E}X_1)$ does not *admit a multivariate Edgeworth expansion of order $s - 2$ if and only if there exists $0 \neq \mathbf{a} \in \mathbb{R}^k, \alpha \in \mathbb{R}$ such that for $m = 2p - 1$,*

$$\mathbf{a}^T H_m(z_1, \dots, z_m) \equiv \alpha \sum_{1}^{m} z_i + \beta \text{ and } \alpha \sum_{1}^{\infty} c_i = 0. \tag{8}$$

The same result holds for S_n'.

Surprisingly, the torus sum H_m which was of importance for the validity of Edgeworth expansions with error $o(n^{-1/2})$ in [7] and for local limit theorems in [6] shows up again in this problem.

Our second result covers the case considered in Example 1.2.

THEOREM 1.3. *Let $f: \mathbb{R}^d \to \mathbb{R}$ denote a strongly contracting and continuously differentiable function, and let $\varepsilon_j, j \in \mathbb{Z}$, denote a sequence of iid random variables with positive and continuous density, satisfying*

$$\mathbb{E}|f(\varepsilon_1, \dots, \varepsilon_d)| < \infty.$$

Let $Z_j := g(\varepsilon_{j-i}, i \geq 0), j \in \mathbb{Z}$, denote the strongly stationary sequence of Example 1.2, *resp. let Z'_j denote the nonstationary sequence* (5) *satisfying the recursion*

$$Z_{j+1} = f(Z_j, Z_{j-1}, \dots, Z_{j-d+1}) + \varepsilon_{j+1}, \qquad j \in \mathbb{Z}.$$

Let $h: \mathbb{R}^p \to \mathbb{R}^k$ denote a continuously differentiable function satisfying

$$\mathbb{E}\|h^{(i)}(Z_1, \dots, Z_p)\| < \infty, \qquad i = 1, \dots, k.$$

Define $X_j := h(Z_j, \dots, Z_{j-p+1}), j \geq 0$, and assume that the moment condition (2) *holds. Then the distribution of $S_n := n^{-1/2}(X_1 + \cdots + X_n - n\mathbb{E}X_1)$ does* not *admit a multivariate Edgeworth expansion of order $s-2$ if and only if there exists $0 \neq \mathbf{a} \in \mathbb{R}^k, \beta \in \mathbb{R}$ such that for $m = 2p-1$ the torus sum of* (7) *satisfies*

$$\mathbf{a}^T H_m(z_1, \dots, z_m) \equiv \beta. \tag{9}$$

Note that condition (9) *for $m = 2p-1$ entails its validity for all $m \geq p$, too. The same result holds for the nonstationary sequence Z'_j.*

The next result is concerned with threshold AR(1) processes introduced in Tong [22]. Here the autoregressive recursion is given by

$$Z_{j+1} = \begin{cases} \alpha Z_j + \varepsilon_{j+1}, & \text{if } Z_j > 0, \\ \beta Z_j + \varepsilon_{j+1}, & \text{if } Z_j \leq 0. \end{cases} \tag{10}$$

We assume here that

$$|\alpha| < 1 \quad \text{and} \quad |\beta| < 1. \tag{11}$$

This is only a small part of the region $\alpha < 1, \beta < 1, \alpha\beta < 1$ for which Petrucelli and Woolford [16] proved the existence of a stationary solution. Assuming condition (11), we obtain a representation of the process Z_j as

$$Z_j = g(\varepsilon_{j-1} \colon i \geq 0)$$

in the same way as in Example 1.2, resp. a representation (5). However, g no longer has partial derivatives *everywhere*.

THEOREM 1.4. *Let α, β satisfy* (11) *and let $Z_j, j \geq 0$, be a stationary sequence satisfying the recursion* (10) *where $\varepsilon_j, j \in \mathbb{Z}$, is a sequence of iid random variables with positive continuous density, and where ε_{j+1} is independent of Z_i, $i \leq j$. Define $X_j := h(Z_j, \dots, Z_{j-p+1}), j \geq 0$, where h is the function introduced in*

Theorem 1.3. *Then the distribution of* $S_n := n^{-1/2}(X_1 + \cdots + X_n - n\mathbb{E}X_1)$ *does* not *admit a multivariate Edgeworth expansion of order* $s - 2$ *if and only if there exists* $0 \neq \mathbf{a} \in \mathbb{R}^k, \beta \in \mathbb{R}$ *such that for* $m = 2p - 1$ *the torus sum of* (7) *satisfies*

$$\mathbf{a}^T H_m(Z_1, \dots, Z_m) \equiv \beta \quad a.s. \tag{12}$$

A similar result holds in the nonstationary case.

Based on Theorem 1.2, the validity of formal Edgeworth expansions for the distribution of maximum likelihood estimators for the parameters $\vartheta := (\vartheta_1, \dots, \vartheta_p)$ in a general AR(d) process $Z_n = f(Z_{n-1}, \dots, Z_{n-d} \mid \vartheta) + \varepsilon_n$ can be derived. Let L denote the log density of ε_j and let $s(Z_0, \dots, Z_{-d+1} \mid \theta)$ denote the log likelihood of the density of the distribution of $Z_0, \dots, Z_{-d+1}$. The log likelihood function of such a generalized AR(d) process is given by (see, e.g., [11])

$$s\big(Z_0, \dots, Z_{-d+1} \mid \vartheta\big) + \sum_{i=1}^{n} L\Big(Z_i - f\big(Z_{i-1}, \dots, Z_{i-d} \mid \vartheta\big)\Big).$$

Under certain regularity conditions the maximum likelihood estimator $\widehat{\vartheta}_n$ admits a stochastic expansion of length l. Thus the statistic

$$T_n\big(\mathbf{Z} \mid \vartheta\big) := (\widehat{\vartheta}_n - \vartheta)\sqrt{n}$$

or its Studentized version can be written as

$$T_n\big(\mathbf{Z} \mid \vartheta\big) = U_{n,0} \sum_{\nu=0}^{l} n^{-\nu/2} Q_\nu(U_{n,0}, \dots, U_{n,\nu}) + o_P\big(n^{-l/2}\big), \tag{13}$$

where

$$U_{n,\nu} := \sum_{i=1}^{n} \big[h_\nu(Z_i, \dots, Z_{i-d}) - \mathbb{E}h_\nu(Z_i, \dots, Z_{i-d})\big] n^{-1/2}$$

for $\nu = 0, \dots, l$ and Q_ν denotes a vector of polynomials. In the particular case of the mle, we have

$$h_0(Z_i, \dots, Z_{i-d}) = L'(\varepsilon_i) f_\vartheta\big(Z_{i-1}, \dots, Z_{i-d} \mid \vartheta\big),$$
$$h_1(Z_i, \dots, Z_{i-d}) = L''(\varepsilon_i) f_\vartheta\big(Z_{i-1}, \dots, Z_{i-d} \mid \vartheta\big)^2 + L'(\varepsilon_i) f_{\vartheta\vartheta}(Z_{i-1}, \dots, Z_{i-d} \mid \vartheta\big),$$

where f_ϑ and $f_{\vartheta\vartheta}$ denote the vector of first and the matrix of second derivatives of f with respect to ϑ and $h_2(Z_i, \dots, Z_{i-d})$ is defined similarly in terms of third likelihood derivatives. In other cases (e.g., for M-estimators) the stochastic expansion has a similar structure.

Up to an error of order $O(n^{-1+\varepsilon})$ an Edgeworth approximation for the distribution of $T_n(\mathbf{Z} \mid \vartheta)$ holds for the general class of dependent sequences considered in Götze and Hipp [5] provided that *only* the $U_{n,0}$ component satisfies the smoothness condition (4). This result is a consequence of Theorem 3.1 of Götze and Künsch [8].

For stochastic expansions (13) of statistics involving a k-dimensional parameter ϑ of the nonlinear time series introduced above, the following result holds.

THEOREM 1.5. *Let ε_j, $j \in \mathbb{Z}$, be iid with smooth positive density. Let $Z_1, Z_2, \dots$ be a nonlinear AR(d) process satisfying the recursion*

$$Z_n = f\left(Z_{n-1}, \dots, Z_{n-d} \mid \vartheta\right) + \varepsilon_n.$$

Here ϑ is an unknown k-dimensional parameter vector. The function $f(\mathbf{z} \mid \vartheta)$ is a differentiable and strongly contracting function (for fixed ϑ) such that

$$\mathbb{E}_\vartheta \left| f\left(\varepsilon_1, \dots, \varepsilon_d \mid \vartheta\right) \right| < \infty.$$

The sequence Z_j may be given either by (1) *or in the nonstationary case by* (5) *and is denoted by Z'_j as before.*

Assume that the statistic $T_n(\mathbf{Z} \mid \vartheta)$ admits a stochastic expansion of type (13) *such that*

$$\mathbb{E}\|h_\nu(Z_i, \dots, Z_{i-d})\|^5 < \infty$$

for $\nu = 0, 1, 2$ where $\|x\|$ denotes the Euclidean norm. Furthermore, assume that h_0 has the following structure:

$$h_0(Z_j, \dots, Z_{j-d}) := \varphi(\varepsilon_j) F\left(Z_{j-1}, \dots, Z_{j-d} \mid \varphi\right),$$

where φ is a real-valued function with $\mathbb{E}_\vartheta \varphi(\varepsilon_j) = 0$ and $F\colon \mathbb{R}^d \to \mathbb{R}$. Assume that h_ν, $\nu = 0, 1$, are continuously differentiable functions of z_j satisfying $\mathbb{E}\|h^{(i)}(Z_1, \dots, Z_p)\| < \infty$, $i = 1, \dots, d+1$. Finally, assume that $h_\nu(\mathbf{0}) = 0$ for $\nu = 0, 1, 2$. [In the maximum likelihood case this follows from $f(\mathbf{0} \mid \vartheta) = 0$ in a neighborhood of ϑ.] Then $T_n(\mathbf{Z} \mid \vartheta)$ and $T_n(\mathbf{Z}' \mid \vartheta)$ admit valid k-dimensional Edgeworth expansions up to an error of order $o(n^{-1})$ if and only if

$$I := \mathbb{E}_\vartheta\left(h_0^T h_0\right)(Z_i, \dots, Z_{i-d}) \text{ is positive definite.}$$

Given moment conditions on $L'(\varepsilon_j)$ and $f(Z_{i-1}, \dots, Z_{i-d} \mid \vartheta)$, Edgeworth expansions hold for $(\widehat{\vartheta}_n - \vartheta)\sqrt{n}$ iff the Fisher information I is positive definite.

2. Lemmas and proofs.

In the following two lemmas we shall use the notation $i := \sqrt{-1}$.

LEMMA 2.1. *Conditions* (1)–(3) *imply that there is a constant K_1 such that there exists a $\mathcal{D}_{j+p-1-m, j+p-1}$-measurable r.v. X_j^* such that*

$$\mathbb{E}\|X_j - X_j^*\| \le K_1 \exp\left[-\alpha \frac{s}{3(s+1)} m\right]. \tag{14}$$

Conditions (1)–(4) *together imply*

$$\liminf_n \inf_{\|t\|=\alpha} \operatorname{var}(t^T S_n) > \frac{1}{2q}\left(1 - \exp[-2\alpha]\right), \tag{15}$$

where q is defined in (16).

PROOF OF LEMMA 2.1. Assuming (14), we shall prove (15) first. With the notation of conditions (2)–(4), let $\gamma := (s-1)/(2s)$ and $C_{s,\alpha} := 2^{-1/2}(1-\exp[-2\alpha])^{1/2}(1-\exp[-\alpha\gamma])^{-1}$:

$$q := r + 1 + \left\lfloor (\gamma\alpha)^{-1}\log\left(K_1\alpha^{s+3}C_{s,\alpha}\mathbb{E}^{1/s}\|X_1\|^{s+1}\right)\right\rfloor, \tag{16}$$

where $\lfloor x \rfloor$ denotes the integral part of x. For $n \geq 2q$ and $\mathcal{F}_j$ denoting the σ-field generated by $\varepsilon_l, l \leq jq + p - 1$, we define

$$\Delta_j := n^{1/2}t^T\left[\mathbb{E}(S_n \mid \mathcal{F}_j) - \mathbb{E}(S_n \mid \mathcal{F}_{j-1})\right].$$

Writing $n = Lq + N$ with $N < q$, the following variance decomposition:

$$\operatorname{var}(t^T S_n) = n^{-1}\sum_{j=1}^{L}\mathbb{E}\Delta_j^2 + \mathbb{E}\left[t^T\left(S_n - \mathbb{E}(S_n \mid \varepsilon_l, l \leq n + p - 1 - N)\right)\right]^2 \tag{17}$$

holds. Since $X_\nu, \nu \leq jq$, are $\mathcal{F}_j$-measurable we obtain

$$\begin{aligned}\Delta_j &= t^T\left[\sum_{\nu=jq-q+1}^{jq-1}\left(X_\nu - \mathbb{E}(X_\nu \mid \mathcal{F}_{j-1})\right) + \sum_{\nu=jq}^{n}\mathbb{E}(X_\nu \mid \mathcal{F}_j) - \mathbb{E}(X_\nu \mid \mathcal{F}_{j-1})\right]\\ &:= V_j + R_j, \quad \text{say.}\end{aligned} \tag{18}$$

Define $\varepsilon_{\nu,M} := (\varepsilon_{\nu+p-1}, \varepsilon_{\nu+p-2}, \ldots, \varepsilon_{Mq+p-1}, 0, \ldots)$ for $\nu \geq Mq$ and $\varepsilon_\nu := (\varepsilon_{\nu+p-l}: l \geq 1)$. We have

$$\mathbb{E}^{1/2}\Delta_j^2 \geq \mathbb{E}^{1/2}V_j^2 - \mathbb{E}^{1/2}R_j^2$$

and

$$\begin{aligned}\mathbb{E}^{1/2}R_j^2 &\leq \sum_{\nu=jq}^{n}\mathbb{E}^{1/2}\left[\mathbb{E}(t^T X_\nu \mid \mathcal{F}_j) - \mathbb{E}(t^T X_\nu \mid \mathcal{F}_{j-1})\right]^2\\ &\leq \sum_{\nu=jq}^{n}\mathbb{E}^{1/2}\left[t^T X_\nu - \mathbb{E}(t^T X_\nu \mid \mathcal{F}_{j-1})\right]^2\\ &\leq \sum_{\nu=jq}^{n}\mathbb{E}^{1/2}\left(t^T g_0(\varepsilon_\nu) - t^T g_0(\varepsilon_{\nu,j-1})\right)^2\\ &\leq C\alpha^2\sum_{\nu=jq}^{n}\mathbb{E}^\gamma\|g_0(\varepsilon_\nu) - g_0(\varepsilon_{\nu,j-1})\|\\ &\leq C\alpha^2\sum_{\nu=jq}^{n}K\exp\left[-\alpha\gamma(\nu - jq + q)\right]\\ &\leq C\alpha^2 K\left(1 - \exp[-\alpha\gamma]\right)^{-1}\exp[-\alpha\gamma q],\end{aligned} \tag{19}$$

where $C := \mathbb{E}^{1/s}|t^T X_1|^{s+1}$. The inequality (19) is a consequence of Hölder's inequality $\mathbb{E}^{1/2}X^2 \le \mathbb{E}^{\gamma}|X|\mathbb{E}^{1/2-\gamma}|X|^{s+1}$. By definition of q we arrive at

$$\mathbb{E}R_j^2 \le \frac{1}{2}\big(1 - \exp[-2\alpha]\big).$$

The inequality $x^2/2 \ge 2\sin^2(x/2) = 1 - \cos(x)$ together with

$$\mathrm{Var}(Z) = \frac{1}{2}\mathbb{E}(Z - \overline{Z})^2 \ge 1 - |\mathbb{E}\exp[iZ]|^2$$

for any r.v. Z and an independent copy, say $\overline{Z}$, as well as condition (4) with $\|t\| = \alpha$, entails

$$\mathbb{E}V_j^2 \ge 1 - \exp[-2\alpha]. \tag{20}$$

Relations (17)–(20) together yield for $n > 2q$,

$$\mathrm{var}\big(t^T S_n\big) \ge \frac{1}{2q}\big(1 - \exp[-2\alpha]\big),$$

thus proving (15).

By definition (1) there is a measurable function g_0 (the composition of h and g) such that $X_j = g_0(\varepsilon_{j+p-l}: l \ge 1)$. For $m > p - 1$ let $X_j' := g_0(\varepsilon_{j+p-1}, \dots, \varepsilon_{j-m+p-1}, 0, \dots)$ and $a_m := \exp[\alpha m/3]$. Define a $\mathcal{D}_{j+p-1-m,\, j+p-1}$-measurable r.v.

$$X_j^* := X_j' I\big(\|X_j'\| \le a_m^{1/(s+1)}\big).$$

Furthermore, let B_m denote the set of sequences ε_j such that

$$\|Dh(X_j, \dots, X_{j+p-1})\| \le K a_m$$

and

$$|Z_{j+\nu} - g(\varepsilon_{j+\nu}, \dots, \varepsilon_{j+\nu-m}, 0, \dots)| \le K\exp\big[-\alpha(m-\nu)\big]a_m.$$

By the assumption of uniform continuity (1), there is an m_0 such that for $m \ge m_0$ and $\varepsilon_j, j \in \mathbb{Z}$, in B_m, $\|Dh\| \le 2Ka_m$ on the segment connecting the Z vectors of the infinite and truncated (at m) sequence of ε_j. Thus by Chebyshev's inequality we have

$$\begin{aligned}\mathbb{E}\|X_j - X_j^*\| &\le \mathbb{E}\|X_j\|^{s+1}a_m^{-s/(s+1)} + \mathbb{E}\|X_j - X_j'\| I_{\{\|X_j\| \le a_m^{1/(s+1)}\}}\big(I_{B_m} + I_{B_m^c}\big)\\ &\le \mathbb{E}\|X_j\|^{s+1}a_m^{-s/(s+1)} + 2pK^2\exp\big[p\alpha - \alpha m/3\big] + 2a_m^{1-1/(s+1)}.\end{aligned}$$

This proves (14). □

LEMMA 2.2. *Let $O \subset \mathbb{R}^k$ denote an open ball with radius r and let F: $\mathbb{R}^k \to \mathbb{R}^k$ denote a measurable, injective and continuously differentiable function on O such that for some constants $\eta > 0$ and $M < \infty$ and all $x \in O$,*

$$\eta \le |\det F'(x)| \le M \quad \textit{and} \quad \|F'(x)\| \le M.$$

Let h denote a density on $\mathbb{R}^k$ satisfying $h(x) \geq \eta$, $x \in O$, and fix $\delta > 0$. Then there exists $\rho < 1$ depending only on k, η, δ, M and r such that for $t \in \mathbb{R}^k$ with $\|t\| \geq \delta$,

$$\left|\int e^{it^T F(x)} h(x)\, dx\right| \leq \rho.$$

PROOF. By change of variables we obtain

$$\int_O e^{it^T F(x)} h(x)\, dx = \int_{FO} e^{it^T u} \frac{h\left(F^{-1}(u)\right)}{\left|\det F'\left(F^{-1}(u)\right)\right|}\, du.$$

For $x \in O$ we have

$$\frac{h(x)}{|\det F'(x)|} \geq \frac{\eta}{M}$$

and therefore

$$\left|\int_{FO} e^{it^T u}\left(\frac{h\left(F^{-1}(u)\right)}{\left|\det F'\left(F^{-1}(u)\right)\right|} - \frac{\eta}{M}\right) du\right| \leq \int_O h(x)\, dx - \frac{\eta}{M}\int_{FO} du.$$

Fix $1 \leq j \leq k$ and $u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_k$. Then $\{u_j\colon (u_1, \dots, u_k) \in FO\}$ is an interval with endpoints $a < b$, say, and

$$\left|\int_a^b e^{it^T u}\, du_j\right| = \left|\frac{1}{it_j}\left(e^{it_j b} - e^{it_j a}\right)\right| \leq \frac{2}{|t_j|}.$$

Let $A := \{(u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_k)\colon \exists\, u_j \in \mathbb{R}\colon (u_1, \dots, u_k) \in FO\}$. We have

$$\left|\int_{FO} e^{it^T u}\, du\right| \leq \frac{2}{|t_j|}\int_A du_1 \dots du_{j-1}\, du_{j+1} \dots du_k.$$

Since A is the projection of FO onto $\mathbb{R}^{k-1}$ and FO is contained in a ball with radius Mr, we obtain

$$\left|\int_{FO} e^{it^T u}\, du\right| \leq \frac{2}{|t_j|}(2Mr)^{k-1}.$$

Since $1 \leq j \leq k$ was arbitrary, we can find $\xi > 0$ depending on k, η, M and r only such that for $t \in \mathbb{R}^k$ and $\|t\| \geq \xi$,

$$\left|\int_{FO} e^{it^T u}\, du\right| \leq \frac{1}{2}\int_{FO} dx.$$

Hence for these t,

$$\begin{aligned}\left|\int e^{it^T F(x)} h(x)\, dx\right| &\leq \int_{O^c} h(x)\, dx + \int_O h(x)\, dx - \frac{\eta}{M}\int_{FO} dx + \frac{\eta}{2M}\int_{FO} dx\\ &= 1 - \frac{\eta}{2M}\int_{FO} dx \leq 1 - \frac{\eta^2}{2M}\int_O dx.\end{aligned}$$

The k-variate version of Theorem 1 in Petrov [15], page 10, yields for any c.f. $\phi(t), t \in \mathbb{R}^k$, with $|\phi(t)| \le c < 1$ for $\|t\| > \xi$,

$$|\phi(t)| \le 1 - \frac{1-c^2}{8\xi^2}\|t\|^2 \quad \text{for } \|t\| < \xi.$$

This immediately proves the assertion. □

We shall replace the smoothness condition (4) by a condition that is easier to check for the case that $X_j, j \in \mathbb{Z}$, admits a representation

$$X_j = g(\varepsilon_{j-i}: i \ge 0),$$

with a function g having all its partial derivatives exponentially bounded and continuous on a set with large probability. Again, $\varepsilon_j, j \in \mathbb{Z}$, is a sequence of independent random variables with distribution P. We write $\mathbb{P}$ for the joint distribution of $\varepsilon_j, j \in \mathbb{Z}$. For $j \in \mathbb{Z}$ and $\mathbf{y} \in \mathbb{R}^{\mathbb{Z}}, x \in \mathbb{R}$ let $(\mathbf{y}, x)^j$ be the sequence with coordinates

$$\varepsilon_i = \begin{cases} y_i, & i < j, \\ x, & i = j, \\ y_{i-1}, & i > j. \end{cases}$$

It will be convenient to write g as a function on $\mathbb{R}^{\mathbb{Z}}$; that is, instead of (1) we start from a representation

$$X_j := g(\varepsilon_{j-i}: i \in \mathbb{Z}), \qquad j \in \mathbb{Z}.$$

LEMMA 2.3. *Assume that* $g: \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^k$ *satisfies the following conditions*:

(i) *(Exponentially small Lipschitz.) There exist* $K < \infty$ *and* $\alpha > 0$ *such that for* $j \in \mathbb{Z}$ *and* $x_1, x_2 \in \mathbb{R}$,

$$\mathbb{E}\left\|g\left((\varepsilon, x_1)^j\right) - g\left((\varepsilon, x_2)^j\right)\right\| \le K e^{-\alpha|j|}|x_1 - x_2|.$$

(ii) *(Almost sure continuity of partial derivatives.) For* $j \in \mathbb{Z}$ *there exists* $G_j \subset \mathbb{R}, P(G_j) = 1$, *such that for all* $x_0 \in G_j, \eta, \delta > 0$ *there exists* $\tau > 0$ *satisfying*

$$\mathbb{P}\left\{\mathbf{y} \in \mathbb{R}^{\mathbb{Z}}: \forall x \in \mathbb{R}, |x - x_0| < \tau, \frac{\partial}{\partial\varepsilon_0}X_j \text{ exists at the point } (\mathbf{y}, x)^j \text{ and } \left|\frac{\partial}{\partial\varepsilon_0}X_j\left((\mathbf{y}, x)^j\right) - \frac{\partial}{\partial\varepsilon_0}X_j\left((\mathbf{y}, x_0)^j\right)\right| \le \delta\right\} \ge 1 - \eta.$$

(iii) *(Nondegenerate derivative on a set of positive probability.) For some distinct* $l_1, \dots, l_k \geq 0$,

$$\det\left(\sum_{j=0}^{\infty} \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j \colon \nu = 1, \dots, k\right) \neq 0$$

on a set of positive $\mathbb{P}$*-probability.*

If, in addition, ε_j *admits a positive continuous density, then the smoothness condition* (4) *holds for the sequence* $X_j, j \in \mathbb{Z}$.

PROOF. By assumption (iii) we can find a number $\eta > 0$ and a set A of sequences $\mathbf{y} \in \mathbb{R}^{\mathbb{Z}}$ with $\mathbb{P}(A) > 0$ such that for $\mathbf{y} \in A$ the partial derivatives

$$\frac{\partial}{\partial \varepsilon_0} X_j, \qquad j \in \mathbb{Z},$$

are defined at $\mathbf{y}$, and

$$\left|\det\left(\sum_{j=0}^{\infty} \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j(\mathbf{y}) \colon \nu = 1, \dots, k\right)\right| \geq \eta.$$

For any $k \times k$ matrix B let $\|B\| := \sup\{\|Bx\| \colon \|x\| \leq 1\}$ and let $C(k)$ denote a universal constant satisfying

$$|\det B_1 - \det B_2| \leq C(k)\big(\|B_1\|^{k-1} + \|B_2\|^{k-1}\big)\|B_1 - B_2\|$$

for arbitrary $k \times k$ matrices B_1 and B_2.

From property (i) we obtain

$$\sum_{j=0}^{\infty} \mathbb{E}\left\|\frac{\partial}{\partial \varepsilon_0} X_j(\mathbf{y})\right\| 1_A(\mathbf{y}) < \frac{2K}{1 - e^{-\alpha}}.$$

Hence there exists a subset $A' \subset A$ with $\mathbb{P}(A') > 0$ and $m_0 > 0$ such that for $\mathbf{y} \in A'$,

$$\left\|\left(\sum_{j=m_0+1}^{\infty} \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j(\mathbf{y}) \colon \nu = 1, \dots, k\right)\right\| \leq \frac{\eta}{4C}, \tag{21}$$

where $C := 2C(k)(2K/(1 - \exp[-\alpha]))^k$.

Estimating the difference of determinants yields for all $\mathbf{y} \in A'$,

$$\left|\det\left(\sum_{j=0}^{m_0} \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j \colon \nu = 1, \dots, k\right)\right| \geq \frac{3}{4}\eta. \tag{22}$$

For $\mathbf{y} \in \mathbb{R}^{\mathbb{Z}}$ and $\boldsymbol{x} = (x_1, \dots, x_k)$ let $(\mathbf{y}, \mathbf{x})$ be the sequence, such that for $\nu = 1, \dots, k, x_\nu$ is inserted at place l_ν, and all other places are filled with the components of $\mathbf{y}$:

$$(\mathbf{y}, \mathbf{x})_j = \begin{cases} x_j, & \text{if } j \in \{l_1, \dots, l_k\}, \\ y_{j-i}, & \text{if } i = 1, \dots, k \text{ and } l_{i-1} < j < l_{i+1}. \end{cases}$$

Here $l_0 = -\infty$ and $l_{k+1} = +\infty$. For $\mathbf{x} \in \mathbb{R}^k$ let

$$A(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^{\mathbb{Z}} : (\mathbf{y}, \mathbf{x}) \in A\}.$$

Since $\mathbb{P}(A) > 0$, we can find $\mathbf{x}^{(0)} \in G_{l_1} \times \cdots \times G_{l_k}$ such that $\mathbb{P}(A(\mathbf{x}^{(0)})) > 0$. The condition (ii) implies that for $\delta > 0$ there exists a small ball $B \subset \mathbb{R}^k$ containing $\mathbf{x}^{(0)}$ and a set $A'' \subset A'$ with $\mathbb{P}(A'') > 0$ such that for $\mathbf{y} \in A''$ and $\mathbf{x} \in B$ and for $\nu = 1, \dots, k$ and $j = 0, \dots, m_0$,

$$\frac{\partial}{\partial \varepsilon_{l_\nu}} X_j \text{ exists at } (\mathbf{y}, \mathbf{x}) \text{ and } \left\| \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j(\mathbf{y}, \mathbf{x}^{(0)}) - \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j(\mathbf{y}, \mathbf{x}) \right\| \le \delta.$$

Choosing δ sufficiently small, we obtain (estimating the change of determinants) from this and (22), $\mathbf{x} \in B$ and $\mathbf{y} \in A''$,

$$\left| \det \left(\sum_{j=0}^{m_0} \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j((\mathbf{y}, \mathbf{x})) : \nu = 1, \dots, k \right) \right| \ge \frac{\eta}{2}.$$

This implies that for $m \ge m_0$ and $\mathbf{x} \in B, \mathbf{y} \in A''$ we have by (21) [which holds for $(\mathbf{y}, \mathbf{x})$, too] again estimating the change of determinants

$$\left| \det \left(\sum_{j=0}^{m_0} \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j((\mathbf{y}, \mathbf{x})) : \nu = 1, \dots, k \right) \right| \ge \frac{\eta}{4}.$$

With Lemma 2.2 we obtain that for $\delta > 0$ there exists $\rho < 1$ depending only on δ, η and B (and not on m) such that for $m \ge m_0$ and all $\mathbf{y} \in A''$ and $\|t\| \ge \delta$,

$$\left| \int \exp\left[it^T(X_0 + \cdots + X_m)\right] \prod_{\nu=1}^{k} h(\varepsilon_{l_\nu}) \, d\varepsilon_{l_1} \cdots d\varepsilon_{l_k} \right| \le \rho.$$

The left-hand side is an upper bound for

$$\left| \mathbb{E}\Big(\exp\left[it^T(X_0 + \cdots + X_{2m})\right] \,\big|\, \varepsilon_j : |j - m| \ge r \Big) \right|,$$

with $r = \max(l_1, \dots, l_k)$. This implies

$$\mathbb{E}\left| \mathbb{E}\Big(\exp\left[it^T(X_0 + \cdots + X_{2m})\right] \,\big|\, \varepsilon_j : |j - m| \ge r \Big) \right| \le 1 - \mathbb{P}(A'') + \rho \mathbb{P}(A''),$$

thus proving the lemma. □

In the following we shall prove conditions (i)–(iii) of Lemma 2.3 for the sequences of dependent r.v.'s X_j of Theorems 1.2–1.4, thus proving these results.

Let $h\colon \mathbb{R}^p \to \mathbb{R}^k$ denote a differentiable function. We denote by $\mathbf{x}_{j;m}^p \in \mathbb{R}^p$ the vector $(x_j, \dots, x_{j+p-1})$, $j \geq 0$, using the identification $j + m \equiv j$, $j \geq 0$. When there are no ambiguities we denote it by $\mathbf{x}_{j;m}$ and write $\mathbf{x}_j$ for $(x_j, \dots, x_{j+p-1})$ for short. Write

$$h^{(j)}(\mathbf{x}_1) := \frac{\partial}{\partial x_j} h(x_1, \dots, x_p).$$

The following condition for torus sums like (7) will be frequently used:

$$H_m := \sum_{\nu=1}^{m} h(\mathbf{x}_{\nu;m}) = \alpha \sum_{\nu=1}^{m} x_\nu + \beta_m \tag{23}$$

for all x's. This condition can be described by the following lemma.

LEMMA 2.4. *The following statements* (i) *and* (ii) *are equivalent and imply property* (iii).

(i) $$\sum_{j=1}^{p} \frac{\partial}{\partial x_p} h(\mathbf{x}_{j;\,2p-1}) \equiv \alpha.$$

(ii) *Condition* (23) *holds for* $m = 2p - 1$.

(iii) *Condition* (23) *holds for all* $m \geq p$.

REMARK 2.1. Condition (23) for $m = n + p - 1$ yields the representation

$$n^{1/2} S_n = \sum_{j=1}^{n} h(\mathbf{x}_j) = \alpha \sum_{j=1}^{m} x_j + \beta_m - \sum \{h(\mathbf{x}_{j;m})\colon n < j \leq m\}. \tag{24}$$

PROOF OF LEMMA 2.4. Condition (i) may be written as

$$\sum_{j=1}^{2p-1} h^{(2p-j)}(\mathbf{x}_{j;\,2p-1}) \equiv \alpha. \tag{25}$$

Taking derivatives with respect to x_p in (23) results in the left-hand side of (25), thus proving that (ii) implies (i). Furthermore, by symmetry, partial derivatives of H_m with respect to any x_j, $1 \leq j \leq m$, are equal. Assuming (i) implies that for $m \geq 2p - 1$ the sum H_m has all partial derivatives equal to α. Therefore (23) holds for any $m \geq 2p - 1$. Finally, choosing *special* arguments x_j, such that $x_{j+m} \equiv x_j$ for $p \leq m \leq 2p - 1$ in (25), the left-hand side of this equality equals the derivative of H_m with respect to x_p. Similarly as above this shows that (23) also holds for $p \leq m \leq 2p - 1$, thus proving the lemma. □

EXAMPLE 2.1. Functions h which lead to the degenerate case in Theorems 1.2–1.4 are, for example,

$$h(x_1, x_2) = g(x_1) - g(x_2) + ax_1,$$

where g and a are arbitrary. Notice that for $h(x_1, x_2) = g_1(x_1)g_2(x_2) - g_1(x_2)g_2(x_1)$ with general functions g_1, g_2, condition (23) holds for $m = 2$ but not for $m = 3$.

REMARK 2.2. To verify the expansion for S_n', note that (3) implies a similar condition for Z_j' as well. Define

$$\tau_{N,m}(t) := \mathbb{E}\Big|\mathbb{E}\Big(\exp\big[it^T(X_{N-m} + \cdots + X_{N+m})\big]\big|\varepsilon_j\colon |j - N| \geq r\Big)\Big|.$$

Let $\tau_{N,m}'(t)$ be defined similarly for Z_j'. Then (3) entails

$$|\tau_{N,m}(t) - \tau_{N,m}'(t)| = O\big(|t|Kme^{-\alpha|N-m|}\big) = O\big(n^{-1}\big) \tag{26}$$

uniformly for any $l_n + m < N < n - l_n - m$ and $\|t\| < n^{\kappa}$, where $l_n := n^{\varepsilon}$ for some $\varepsilon > 0$. Thus condition (4) entails a similar inequality for the nonstationary case and yields the expansion for S_n' again by Corollary (2.9) in [5]. By Remark 3.45 on page 235 in [5], it follows that condition (4) needs to be checked for sums of length $2m$ involving the random vectors $Z_{l_n}', \ldots, Z_{n-l_n}'$ only. [This allows us to show, e.g., that with $Z_{-l_n}', \ldots, Z_{d-1}'$ and $Z_{n+1-l_n}', \ldots, Z_n'$ defined differently from (5) or (1) the expansion result still holds.]

PROOF OF THEOREM 1.2. Recall that here we have $Z_j := \sum_{i=0}^{\infty} c_i\varepsilon_{j-i}, j \in \mathbb{Z}$, and $X_j := h(\mathbf{Z}_j)$ using the notation $\mathbf{Z}_j := (Z_j, \ldots, Z_{j+p-1})$ of Lemma 2.4. Thus

$$\frac{\partial}{\partial\varepsilon_0}X_j = \sum_{i=1}^{p} c_{j+i-1}h^{(i)}(\mathbf{Z}_j)$$

so that, for continuous $h^{(1)}, \ldots, h^{(k)}, (\partial/\partial\varepsilon_0)X_j$ is a continuous function of ε_0. Condition (i) of Lemma 2.3 follows under the assumption that $c_j \to 0$ exponentially, and $\mathbb{E}\|h^{(i)}(\mathbf{Z}_1)\| < \infty, i = 1, \ldots, k$ holds. Condition (iii) follows from uniform continuity of $h^{(i)}$ on compact subsets of $\mathbb{R}^k, i = 1, \ldots, k$.

To show condition (ii), note that the power series

$$c(z) := \sum_{\nu=0}^{\infty} c_{\nu}z^{\nu}$$

has a radius of convergence $\beta > 1$. Hence there exists $1 < \beta' < \beta$ such that $c(z)$ has a finite number of zeros in $|z| \leq \beta'$. Let $p(z)$ denote a polynomial of degree q defined as $p(z) := \sum_{\nu=0}^{q} p_{\nu}z^{\nu}$ which has the same zeros with the same multiplicity as $c(z)$. Furthermore, let $A(z)$ denote the holomorphic function

$$A(z) := \frac{p(z)}{c(z)}$$

defined for $|z| \leq \beta'$. Expanding $A(z)$ in a power series around $z = 0$ yields $A(z) := \Sigma_{\nu=0}^{\infty} A_\nu z^\nu$. By Cauchy's inequalities we have

$$|A_\nu| \leq \max_{|z|=\beta'} \left|\frac{p(z)}{c(z)}\right| \left(\frac{1}{\beta'}\right)^\nu \quad \text{for } \nu \geq 0. \tag{27}$$

Hence the radius of convergence for $A(z)$ is at least β'. For all $j \geq 0$ we have

$$\sum_{l=0}^{j} A_l c_{j-l} = p_j$$

which is 0 if $j > q$.

Furthermore, we write (using the notation of Lemma 2.4) $C(l_1, \dots, l_k)$ for the matrix with columns

$$\sum_{j=0}^{\infty} \frac{\partial}{\partial \varepsilon_{l_\nu}} X_j = \sum_{j=0}^{\infty} \sum_{i=1}^{p} c_{j+i-1-l_\nu} h^{(i)}(\mathbf{Z}_j), \qquad \nu = 1, \dots, k.$$

According to Lemma 2.3 we may assume that for arbitrary $l_1, \dots, l_k$,

$$\det C(l_1, \dots, l_k) = 0 \quad \mathbb{P}\text{-almost surely.} \tag{28}$$

This implies with $\mu := q + p + 1$,

$$\sum_{l_\nu \geq 0, \nu=1,\dots,k} A_{l_1} \dots A_{l_k} \det C\big(l_1, l_2 + \mu, \dots, l_k + \mu(d-1)\big) = 0 \tag{29}$$

$\mathbb{P}$-almost surely, where this sum is defined in the L_1-sense; notice that the coefficients A_j are exponentially small by (27).

The sum (29) is the determinant of the matrix with columns

$$\sum_{l_\nu=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=1}^{p} A_{l_\nu} c_{j+i-1-l_\nu-(\nu-1)\mu} h^{(i)}(\mathbf{Z}_j)$$
$$= \sum_{j=0}^{\infty} \sum_{i=1}^{p} p_{j+i-1-(\nu-1)\mu} h^{(i)}(\mathbf{Z}_j).$$

Notice that the summation over j on the right-hand side is a finite sum because of $p_j = 0, j > q$.

The case $k = 1$. In the expression

$$\sum_{j=0}^{\infty} \sum_{i=1}^{p} p_{j+i-1} h^{(i)}(\mathbf{Z}_j),$$

the only term depending on ε_{q+p-1} is $p_q h^{(1)}(\mathbf{Z}_q)$. Since $p_q \neq 0$ the function $h^{(1)}(\mathbf{z}_1)$ must be independent of z_p. The only sum of terms depending on ε_{q+p-2} is given by

$$p_q\big(h^{(1)}(\mathbf{Z}_q) + h^{(2)}(\mathbf{Z}_{q-1})\big) + p_{q-1} h^{(1)}(\mathbf{Z}_{q-1}). \tag{30}$$

Since $h^{(1)}$ does not depend on its last entry, the term

$$h^{(1)}(\mathbf{x}_1) + h^{(2)}(\mathbf{x}_0)$$

must be independent of x_{p-1}. It was independent of x_p, hence it does not depend on x_{p-1}, x_p. Repeating this procedure, we arrive at the conclusion that for $j = 1, \ldots, p$ the terms

$$D_j := \sum_{i=1}^{j} h^{(i)}(\mathbf{x}_{q-i+1})$$

do not depend on $x_{q+p-j}, \ldots, x_{q+p-1}$. We shall use these observations to derive that the torus sum H_m defined in (7) must be linear. Since $D_1 = h^{(1)}(\mathbf{x}_q)$ does not depend on x_{q+p-1}, we obtain that $D_1^* := h^{(q)}(\mathbf{x}_{q-p+1})$ does not depend on x_{q-p+1}; in fact, when $\delta := x_1' - x_1 > 0$ we have

$$\begin{aligned}
&\frac{1}{\delta}\big(h(x_1, \ldots, x_{p-1}, x_p + \delta) - h(x_1, \ldots, x_p)\big)\\
&\quad - \frac{1}{\delta}\big(h(x_1', x_2, \ldots, x_{p-1}, x_p + \delta) - h(x_1', x_2, \ldots, x_p)\big)\\
&\quad = \frac{1}{\delta}\int_{x_1'}^{x_1} \big(h^{(1)}(u, x_2, \ldots, x_{p-1}, x_p + \delta) - h^{(1)}(u, x_2, \ldots, x_p)\big)\, du = 0.
\end{aligned}$$

(This is immediate if f is two times differentiable by interchanging partial derivatives.)

Similarly, the fact that

$$D_2 = h^{(1)}(\mathbf{x}_q) + h^{(2)}(\mathbf{x}_{q-1})$$

does not depend on x_{q+p-2} yields that

$$D_2^* := h^{(p-1)}(\mathbf{x}_{q-p+2}) + h^{(p)}(\mathbf{x}_{q-p+1})$$

does not depend on x_{q-p+2} and x_{q-p+1}. In general, for $j = 1, \ldots, p$,

$$D_j^* := \sum_{i=j}^{p} h^{(i)}(\mathbf{x}_{q-i+1})$$

does not depend on $x_{q-j+1}, \ldots, x_{q-p+1}$. We conclude that

$$D_p = D_1^*$$

does not depend on any of its arguments $x_{q-p+1}, \ldots, x_{q+p-1}$; that is, D_p is a constant equal to α, say. By Lemma 2.4 this is equivalent to the torus sum G_{2p-1} being affine linear; that is, the torus sum condition of Theorem 1.2 holds.

We have to show that $\alpha \neq 0$ and condition (28) together imply $\Sigma_0^\infty c_j = 0$. This follows choosing $l \geq p$ in view of

$$0 = \sum_{j=0}^{\infty}\sum_{i=1}^{p} c_{j+i-1-l} h^{(i)}(\mathbf{Z}_j) = \sum_{\nu=0}^{\infty} c_j \sum_{i=1}^{p} h^{(i)}(\mathbf{Z}_{l-i+\nu+1})$$
$$= \alpha \sum_{\nu=0}^{\infty} c_\nu.$$

The case $k > 1$. Assume that the determinant of

$$\sum_{j=0}^{\infty}\sum_{i=1}^{p} p_{j+i-1-(\nu-1)\mu} h^{(i)}(\mathbf{Z}_j), \qquad \nu = 1, \dots, k, \tag{31}$$

vanishes $\mathbb{P}$-almost everywhere. We have to show that there exists $0 \neq \mathbf{a} \in \mathbb{R}^k$ such that the torus sum $\mathbf{a}^T G$ is constant or affine linear, and if $\mathbf{a}^T G$ is not constant that $\Sigma_0^\infty c_j = 0$. We shall use induction in k. The only column depending on $\varepsilon' := (\varepsilon_{(k-1)\mu}, \dots, \varepsilon_{q+(k-1)\mu+p-1})$ is the column with index k since we had chosen $\mu = q + p + 1$. If $f_1, \dots, f_k$ denote the components of f, we can find an $\mathbf{a} = (a_1, \dots, a_k)$, such that

$$0 = \sum_{\nu=1}^{k} a_\nu \sum_{j=0}^{\infty}\sum_{i=1}^{p} p_{j+i-1-(k-1)\mu} h_\nu^{(i)}(\mathbf{Z}_j) \tag{32}$$

$\mathbb{P}$-almost surely. In fact, the entries to $\mathbf{a}$ may be chosen as $k-1$-variate minorants of the given $k \times k$ determinant. Thus $\mathbf{a}$ depends on $\varepsilon_j, j < (k-1)\mu$, but not on ε'. If $\mathbf{a} = 0$ $\mathbb{P}$-almost surely, then, in particular, $a_k = 0$ $\mathbb{P}$-almost surely, which indicates that the determinant of (32) vanishes $\mathbb{P}$-almost surely for dimension $k-1$. In this case we know by induction that for some $0 \neq \mathbf{a}' \in \mathbb{R}^{k-1}$ the torus sum $\Sigma_{\nu=1}^{k-1} a_\nu G_\nu$ must be constant or nonconstant and affine linear and—in the nonconstant case—we have $\Sigma_0^\infty c_j = 0$. Now $\mathbf{a} = (\mathbf{a}', 0) \in \mathbb{R}^k$ is the vector that does all we need.

If relation (32) holds for some $\mathbf{a} \neq 0$, consider

$$\varphi(x_1, \dots, x_p) := \sum_{\nu=1}^{k} a_\nu f_\nu(x_1, \dots, x_p)$$

so that (32) reads

$$0 = \sum_{j=0}^{\infty}\sum_{i=1}^{p} p_{j+i-1-(k-1)\mu} \varphi^{(i)}(\mathbf{Z}_j).$$

As in the case $k = 1$ we can conclude that the torus sum $\mathbf{a}^T G$ must be constant or affine linear, and, in the nonconstant case, $\Sigma_0^\infty c_j = 0$. This proves one part of

Theorem 1.2 for $k > 1$. When the torus sum is constant or affine linear, Lemma 2.4 yields

$$\begin{aligned} n^{1/2}S_n &= \alpha \sum_{j=0}^{n-1} Z_j + \beta_{n+p-1} + O_P(n^{-1/2}) \\ &= \alpha \sum_{\nu=0}^{\infty} \varepsilon_\nu \sum_{j=0}^{n-1-\nu} c_j + \beta_{n+p-1} + O_P(n^{-1/2}). \end{aligned}$$

Hence $\alpha \Sigma_{\nu=0}^{\infty} c_\nu = 0$ together with the exponential decrease of c_ν implies $\mathrm{Var}(S_n) = O(n^{-1})$. Thus the asymptotic distribution of S_n is degenerate and Edgeworth expansions do not hold. This completes the proof of Theorem 1.2. □

PROOF OF THEOREM 1.3. Using the notation $\mathbf{Z}_j^d := (Z_j, \dots, Z_{j-d+1})$ and $\mathbf{Z}_j^p := (Z_j, \dots, Z_{j-p+1})$ as in Lemma 2.4 (with the order changed), we obtain for an autoregressive recursion with a smooth and strongly contracting function $f \colon \mathbb{R}^d \to \mathbb{R}$ the following relation for $(\partial/\partial\varepsilon_0)Z_j$:

$$\frac{\partial}{\partial \varepsilon_0} Z_j = \delta_{j,0} + \sum_{i=1}^{d} f^{(i)}(\mathbf{Z}_{j-1}^d) \frac{\partial}{\partial \varepsilon_0} Z_{j-i}, \qquad j \in \mathbb{Z}. \tag{33}$$

If f is strongly contracting, we have $\|f^{(i)}(\mathbf{z})\| \le \rho_i, \Sigma \rho_i \le \rho < 1$. Hence

$$\left\| \frac{\partial}{\partial \varepsilon_0} Z_j \right\| \le \rho^j, \qquad j \ge 0.$$

Furthermore, we have

$$\frac{\partial}{\partial \varepsilon_0} X_j = \sum_{i=1}^{p} h^{(i)}(\mathbf{Z}_j^p) \frac{\partial}{\partial \varepsilon_0} Z_{j-i+1} \tag{34}$$

and therefore

$$\left\| \frac{\partial}{\partial \varepsilon_0} X_j \right\| \le \sum_{i=1}^{q} \|h^{(i)}(\mathbf{Z}_j^p)\| \rho^{j-i+1}$$

for $j \ge 1$ where $q := \min(p, j-1)$.

Thus, by the assumptions on h, condition (i) of Lemma 2.3 is satisfied in this case. Again, condition (ii) follows from uniform continuity of $f^{(i)}, i = 1, \dots, d$, and $h^{(j)}, j = 1, \dots, p$, on compact subsets of $\mathbb{R}^k$.

To simplify the argument, let us consider the case $d = 1$ first. Here the recursion yields a simple expression for the partial derivatives. Using the convention $\Pi_{i=0}^{-1} := 1$, we have

$$\frac{\partial}{\partial \varepsilon_0} Z_j = \prod_{i=0}^{j-1} f'(Z_i), \qquad j \ge 0. \tag{35}$$

This implies that, especially for nonsmooth f, the partial derivative exists for all points $\mathbf{z} \in \mathbb{R}^{\mathbb{Z}}$ such that the right-hand side exists.

It remains to verify condition (iii) of Lemma 2.3; that is,

$$S := \sum_{j=0}^{\infty} \frac{\partial}{\partial \varepsilon_0} X_j = \sum_{i=1}^{p} \sum_{j=0}^{\infty} h^{(i)}(\mathbf{Z}_j^p) \frac{\partial}{\partial \varepsilon_0} Z_{j-i+1} \tag{36}$$

cannot be 0 $\mathbb{P}$-almost everywhere. Let θ_R denote the right shift

$$\theta_R(\varepsilon)_j := \varepsilon_{j+1}, \qquad \varepsilon \in \mathbb{R}^{\mathbb{Z}}, \qquad j \in \mathbb{Z}.$$

By the representation of Z_j in terms of ε_j of Example 1.2, we have $Z_{j+1} = Z_j \circ \theta_R$ and therefore $X_{j+1} = X_j \circ \theta_R$. Thus relation (35) yields

$$\frac{\partial}{\partial \varepsilon_0} Z_{j+1} = f'(Z_0) \left(\frac{\partial}{\partial \varepsilon_0} Z_j \right) \circ \theta_R. \tag{37}$$

Hence we obtain

$$f'(Z_0) \left(\sum_{j=i-1}^{\infty} h^{(i)}(\mathbf{Z}_j^p) \frac{\partial}{\partial \varepsilon_0} Z_{j-i+1} \right) \circ \theta_R = \sum_{j=i}^{\infty} h^{(i)}(\mathbf{Z}_j^p) \frac{\partial}{\partial \varepsilon_0} Z_{j-i+1}. \tag{38}$$

This immediately implies, by (34) and (36),

$$S = \sum_{i=1}^{p} h^{(i)}(\mathbf{Z}_{i-1}^p) + f'(Z_0) S \circ \theta_R. \tag{39}$$

In the general case $d > 1$ we will show by induction in $j \geq 1$,

$$\frac{\partial}{\partial \varepsilon_0} Z_j = \sum_{i=1}^{d} f^{(i)}(\mathbf{Z}_{i-1}^d) \left(\frac{\partial}{\partial \varepsilon_0} Z_{j-i} \circ \theta_R^i \right), \tag{40}$$

which holds for $j = 1$ by (33). For $j \geq 2$ we have by induction, interchanging summations and (33),

$$\begin{aligned}
\frac{\partial}{\partial \varepsilon_0} Z_j &= \sum_{i=1}^{d} f^{(i)}(\mathbf{Z}_{j-1}^d) \left\{ \delta_{j-i,0} + \sum_{l=1}^{d} f^{(l)}(\mathbf{Z}_{l-1}^d) \left(\frac{\partial}{\partial \varepsilon_0} Z_{j-i-l} \right) \circ \theta_R^l \right\} \\
&= \sum_{l=1}^{d} f^{(l)}(\mathbf{Z}_{l-1}^d) \left\{ \delta_{j-l,0} + \sum_{i=1}^{d} \left(f^{(i)}(\mathbf{Z}_{j-l-1}^d) \left(\frac{\partial}{\partial \varepsilon_0} Z_{j-i-l} \right) \right) \circ \theta_R^l \right\} \\
&= \sum_{l=1}^{d} f^{(l)}(\mathbf{Z}_{l-1}^d) \left(\frac{\partial}{\partial \varepsilon_0} Z_{j-l} \right) \circ \theta_R^l,
\end{aligned}$$

where we used $\mathbf{Z}_j^d = \mathbf{Z}_{j-l}^d \circ \theta_R^l$ in the second equality and (33) again in the last equality. This proves (40).

As in (36)–(39) this result implies

$$S = \sum_{i=1}^{p} \left\{ h^{(i)}(\mathbf{Z}_{i-1}^{p}) + \sum_{j=i}^{\infty} h^{(i)}(\mathbf{Z}_{j}^{p}) \sum_{l=1}^{d} f^{(l)}(\mathbf{Z}_{j-i-l}^{d}) \left(\frac{\partial}{\partial \varepsilon_0} Z_{j-i-l+1} \right) \circ \theta_R^l \right\}$$
$$= \sum_{i=1}^{p} \left\{ h^{(i)}(\mathbf{Z}_{i-1}^{p}) + \sum_{l=1}^{d} f^{(l)}(\mathbf{Z}_{l-1}^{d}) \left(\sum_{j=i+l-1}^{\infty} h^{(i)}(\mathbf{Z}_{j-l}^{p}) \frac{\partial}{\partial \varepsilon_0} Z_{j-i-l+1} \right) \circ \theta_R^l \right\}.$$

Hence we obtain the following generalization of (39):

$$S = \sum_{i=1}^{p} h^{(i)}(\mathbf{Z}_{i-1}^{p}) + \sum_{i=1}^{d} f^{(i)}(\mathbf{Z}_{i-1}^{d}) S \circ \theta_R^i. \tag{41}$$

Assuming that S is 0 $\mathbb{P}$-a.e. implies by stationarity that $S \circ \theta^i$ is 0 $\mathbb{P}$-a.e. and therefore

$$\sum_{i=1}^{p} h^{(i)}(\mathbf{Z}_{i-1}^{p}) = 0 \quad \mathbb{P}\text{-a.e.}$$

This condition implies via Lemma 2.4 that the torus sum H_{2p-1} satisfies

$$H_{2p-1} \equiv \beta_{2p-1}. \tag{42}$$

This contradicts the assumption in Theorem 1.3 and shows (iii) of Lemma 2.3. The lemma is turn guarantees the smoothness condition (4). Finally, by Theorem 1.1 Edgeworth expansions hold up to the order $s-2$. For the *only if* part of Theorem 1.3 notice that (42) implies by Lemma 2.4 that S_n = constant + $O_P(n^{-1/2})$; that is, the distribution of S_n is degenerate in the sense that it is a sum of $O(p)$ r.v.'s only and therefore Edgeworth expansions are not valid. This completes the proof of Theorem 1.3. □

PROOF OF THEOREM 1.4. In the threshold model of Theorem 1.4, if $|\alpha|$, $|\beta| < 1$, condition (i) of Lemma 2.3 follows as above since f is strongly contracting. Condition (ii) is more delicate. From the product structure we see that it is sufficient to show that for $x_0 \in \mathbb{R}$ and $\delta, \eta > 0$ there exists $\tau > 0$ such that

$$\mathbb{P}\left\{ \varepsilon \in \mathbb{R}^{\mathbb{Z}} : \forall |x - x_0| \le \tau, Z_i\big((\varepsilon, x)^i\big) \ne 0 \right\} \ge 1 - \eta. \tag{43}$$

[Recall that in the representation $Z_j = g(\varepsilon_{j-i}: i \ge 0)$ the argument ε_0 appears as the ith argument.] For $i = 1$ we have

$$Z_1 := h\big(f(X_{-1}) + \varepsilon_0\big) + \varepsilon_1$$

and the probability (43) is given by

$$\mathbb{P}\left\{ \varepsilon \in \mathbb{R}^{\mathbb{Z}} : \forall |x - x_0| < \tau, f\big(f(Z_{-1}) + x\big) + \varepsilon_1 \ne 0 \right\}.$$

This probability, by the choice of τ, can be made arbitrarily close to 1 if ε_1 admits a continuous density. Notice that, for Z_{-1} fixed, the set

$$\Big\{f\big(f(X_{-1})+x\big)\colon |x-x_0|<\tau\Big\}$$

lies in an interval of length at most 2τ. For $i=2$ we similarly obtain the probability

$$\mathbb{P}\Big\{\varepsilon\in\mathbb{R}^{\mathbb{Z}}\colon \forall\,|x-x_0|<\tau, f\Big(f\big(f(Z_{-1})+x\big)+\varepsilon_1\Big)+\varepsilon_2\neq 0\Big\}.$$

This is close to 1 for small τ, since for fixed ε_1 and Z_{-1} the set

$$\Big\{f\Big(f\big(f(X_{-1})+x\big)+\varepsilon_1\Big)+\varepsilon_2\colon |x-x_0|<\tau\Big\}$$

is again contained in an interval of length 2τ.

Let M be an upper bound for the density of ε_1. We conclude for arbitrary i:

$$\mathbb{P}\Big\{\varepsilon\in\mathbb{R}^{\mathbb{Z}}\colon \forall\,|x-x_0|<\tau, Z_i\big((\varepsilon,x)^i\big)\neq 0\Big\}\geq 1-2M\tau.$$

Hence condition (ii) of Lemma 2.3 holds in the threshold model of Theorem 1.4 for the variables Z_j.

Condition (iii) of Lemma 2.3 in this case follows similarly as in (35)–(39). Notice that whenever the right-hand side of (35) is defined this equality holds and therefore (37) holds $\mathbb{P}$-a.s., too. Assuming that S defined in (36) is 0 $\mathbb{P}$-a.s., we obtain by stationarity that $S\circ\theta_R$ must be 0 $\mathbb{P}$-a.s., contradicting (39) that holds $\mathbb{P}$-a.s. □

PROOF OF THEOREM 1.5. Write as before $\mathbf{Z}_j := (Z_j,\dots,Z_{j-d+1})$. Furthermore, define $h(\mathbf{Z}_j) := (h_0,h_1,h_2)(\mathbf{Z}_j)\in\mathbb{R}^L$ and the torus sum $H_m := (H_{m,0},H_{m,1},H_{m,2})$ as in (9) with $m := 2d-1$. Let M denote the dimension of the vector space of functions with real coefficients spanned by the component functions of H_m in L^2. Let $\mathbf{0}$ denote the null vector. Note that the assumption $h(\mathbf{0})=0$ implies that a linear relation of the type

$$\mathbf{a}^T H_m(\mathbf{z})\equiv\text{const} \tag{44}$$

for some $\mathbf{a}\in\mathbb{R}^L$ is equivalent by continuity to const $=0$, that is, a linear relation for the components of H_m. We may now distinguish several cases.

The case $M=L$. Here a linear relation like (44) does not hold. Thus the moment conditions on h and (9) imply via Theorem 1.3 that a *multivariate Edgeworth expansion holds* for $U := (U_{n,0},U_{n,1},U_{n,2})$. Applying the usual techniques like the "delta method"; that is, integrating the multivariate Edgeworth expansion over the region defined by the asymptotically linear stochastic expansion polynomial taking values in a convex set and expanding the result in powers of $n^{-1/2}$, we arrive at the formal expansion based on the cumulants of the statistic $T_n(\mathbf{Z} \mid \vartheta)$. As in [1], this method works provided the ε-boundary of the region defined by the stochastic expansion has Gaussian probability proportional to ε. This in turn follows from the polynomial structure of the stochastic expansion.

Furthermore, we claim that the k components of $H_{m,0}$ are independent functions which implies $M \geq k$. To demonstrate this fact, assume that there exists an $\mathbf{a} \in \mathbb{R}^k$, $\mathbf{a} \neq \mathbf{0}$ such that

$$\mathbf{a}^T \sum_{j=d}^{m+d} h_0(\mathbf{Z}_j) \equiv \mathbf{0}, \tag{45}$$

where $m := 2d - 1$ and $Z_{m+j} \equiv Z_j$. By Lemma 2.4 we conclude that (45) for $m = 2d - 1$ implies (45) for all $m \geq 2d - 1$. Since $f_\theta(\mathbf{0} \mid \theta) = f(\mathbf{0} \mid \theta) = 0$ we may choose $X_j = 0$ for all $j = 1, \dots, m$ consistently with the cyclic boundary conditions. Thus we conclude $\beta_m = 0$ for any $m \geq 2d - 1$. Furthermore, notice that by assumption we have $\mathbb{E}\varphi(\varepsilon_j) = 0$. Recall that the r.v.'s $Z_1, \dots, Z_{d-1}, \varepsilon_d, \varepsilon_{d+1}, \dots, \varepsilon_m$ are independent and let the remaining variables be determined by

$$\begin{aligned} Z_j &:= f(\mathbf{Z}_{j-1} \mid \theta) + \varepsilon_j, \\ \varepsilon_i &:= Z_i - f(\mathbf{Z}_{i-1} \mid \theta) \\ &:= Z_i - f(Z_{m-d+i}, \dots, Z_m, Z_1, \dots, Z_{i-1}), \end{aligned} \tag{46}$$

where $d \leq j \leq m$ and $1 \leq i < d$.

Let $\mathcal{A}_d$ denote the σ-field $\sigma(Z_1, \dots, Z_{d-1}, \varepsilon_d)$. Taking the conditional expectation of (45) with respect to $\mathcal{A}_d$ yields

$$\varphi(\varepsilon_d)\mathbf{a}^T F_\vartheta(\mathbf{Z}_{d-1} \mid \vartheta) + H_m(\mathcal{A}_d) \equiv 0, \tag{47}$$

where

$$H_m(\mathcal{A}_d) := \sum_{i=1}^{d-1} \mathbb{E}\varphi(\varepsilon_i)\mathbf{a}^T F_\vartheta(\mathbf{Z}_i \mid \vartheta)$$

for any $m \geq 2d - 1$. Since $Z_j = g(\varepsilon_{j-i}\colon i \geq 0)$ satisfies the ergodicity condition (3), we obtain

$$\lim_{m \to \infty} H_m(\mathbf{Z}_{d-1}, \varepsilon_d) = H(\mathbf{Z}_{d-1}) \quad \text{a.s.} \tag{48}$$

Conditional expectation of (47) given $Z_1, \dots, Z_{d-1}$ yields in the limit as $m \to \infty$ $H(Z_1, \dots, Z_{d-1}) \equiv \mathbf{0}$. Thus

$$\varphi(\varepsilon_d)\mathbf{a}^T F_\vartheta(X_{d-1} \mid \vartheta) \equiv 0, \tag{49}$$

contradicting the assumption of a positive-definite Fisher information matrix I. This proves the assertion.

The case $M < L$. Here we may choose a basis of length $M \geq k$, say H_m^*, of components of H_m that contains the k components of $H_{m,0}$. Thus the remaining $L - M$ components, say H_m^{**}, of H_m for $m = 2d - 1$ may be expressed as

$$H_m^{**} = AH_m^*, \tag{50}$$

where A denotes a $(L - M) \times M$ matrix that does depend on the distribution of $Z_1, \ldots, Z_{2d-1}$ and the function h only. By Lemma 2.4 the linear relations $H_{2d-1}^{**} - AH_{2d-1}^* = \mathbf{0}$ imply the relations $H_n^{**} - AH_n^* = \mathbf{0}$. Thus we may express the corresponding components of U_n, say U_n^{**}, by

$$\begin{aligned} U_n^{**} &= AU_n^* + \left[A(H^* - U_n^*) - (H_n^{**} - U_n^{**})\right] \\ &:= AU_n^* + \Delta_n. \end{aligned}$$

The h_2 components of Δ_n appear in the $O(n^{-1})$ term of $T_n(\mathbf{Z} \mid \vartheta)$ only; that is, their contribution to $T_n(\mathbf{Z} \mid \vartheta)$ is of order $O_P(n^{-3/2})$. We replace the r.v.'s Z_j (resp. the nonstationary versions Z_j') for $n - d < j \leq n$ by

$$\widetilde{Z}_j := g(\varepsilon_j, \ldots, \varepsilon_{j-l_n}, 0, 0, \ldots) 1_{\{|g| < n^\beta\}} \tag{51}$$

for some $1/4 < \beta < 1/2$ and $l_n := n^\varepsilon, \varepsilon > 0$ small. This amounts to a change in $T_n(\mathbf{Z} \mid \vartheta)$ of order $O_P(n^{-1}e^{-\delta l_n})$ by the smoothness assumption on h_1 and a change of order $O(1)$ with probability $O(n^{-5/\beta}) = o(n^{-1})$ by the moment assumptions on h_1. Let us denote the corresponding version of U_n^{**} by $\widetilde{U}_n$. Then the cumulants up to the order 4 of U_n^{**} and $\widetilde{U}_n$ differ by $o(n^{-1})$ only. (Compare Lemma 3.30 of [5].) Thus by standard smoothing inequalities (see Lemma 11.4 and Corollary 11.5 in [2]), we obtain that the Edgeworth expansions of $T_n(\mathbf{Z} \mid \vartheta)$ and its modified version, say $\widetilde{T}_n(\mathbf{Z} \mid \vartheta)$, differ by $o(n^{-1})$ only. Let $\mathcal{E}$ denote the σ-field generated by $\varepsilon_j, j \leq d - 1$, and $\varepsilon_{n-l_n}, \ldots, \varepsilon_n$. Then

$$\mathbb{E}(\widetilde{U}_n \mid \mathcal{E}) = \mathbb{E}(AU_n^* \mid \mathcal{E}) + \Delta(\mathcal{E}),$$

where $\Delta(\mathcal{E})$ is constant given $\mathcal{E}$. Thus, conditionally on $\mathcal{E}, \widetilde{T}_n(\mathbf{Z} \mid \vartheta)$ is a polynomial in U_n^* only. The distribution of $Z_j, 0 < j \leq n$, given $\mathcal{E}$ is not stationary but (3) together with the smoothness of h_0 and h_1 implies that for $l_n < j < n - l_n$ the L^1-difference between Z_j and the conditional version with some ε_j fixed is exponentially small. Similar to (5), it follows that U_n^* given $\mathcal{E}$ satisfies the conditional Cramér condition (4) since the linear independence of H_m^* implies as in the proof of Theorem 1.3 via Lemma 2.3 that (4) holds for U_n^* (based on the stationary unconditioned version of Z_j). Furthermore, (3) holds conditional on $\mathcal{E}$ and therefore we obtain an Edgeworth expansion for the conditional distribution of $\widetilde{T}_n(\mathbf{Z} \mid \vartheta)$ given $\mathcal{E}$ with error of order $o_p(n^{-1})$. Taking expectations over these conditional expansions finally yields an Edgeworth expansion [up to an error $o(n^{-1})$] that is equal to the formal expansion since the conditioned Edgeworth expansions are standard (i.e., depend on the cumulants of the distribution only).

On the other hand, when I is not positive definite the stochastic expansion is asymptotically degenerate and a multivariate Edgeworth expansion for $T_n(\mathbf{Z} \mid \vartheta)$ does not hold. This concludes the proof of Theorem 1.5. □

Acknowledgment. We would like to thank a referee for his careful reading of the first version and for helpful remarks which improved the paper.

REFERENCES

[1] BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451. [Correction (1980) **8** 1399.]

[2] BHATTACHARYA, R. N. and RANGA RAO, R. (1986). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.

[3] BOSE, A. (1988). Higher order approximations for auto-covariances from linear processes with applications. *Statistics* **19** 259–269.

[4] BOSE, A. (1988). Edgeworth correction by bootstrap in autoregression. *Ann. Statist.* **16** 1709–1722.

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

[6] GÖTZE, F. and HIPP, C. (1989). Asymptotic expansions for potential functions of iid random fields. *Probab. Theory Related Fields* **82** 349–370.

[7] GÖTZE, F. and HIPP, C. (1990). Local limit theorems for sums of finite range potentials of a Gibbsian random field. *Ann. Probab.* **18** 810–828.

[8] GÖTZE, F. and KÜNSCH, H. (1993). Second order correctness of the blockwise bootstrap for stationary observations. SFB 343 Bielefeld Preprint 93-061.

[9] JENSEN, J. L. (1986). A note on the work of Götze and Hipp concerning asymptotic expansions for sums of weakly dependent random vectors. Memoirs 10. Dept. Theoretical Statistics, Aarhus Univ. [Summary in *Proceedings of the Fourth Prague Symposium on Asymptotic Statistics* (P. Mandl and M. Hušková, eds.) (1989) 295–303.]

[10] JENSEN, J. L. (1991). A note on asymptotic expansions for sums over a weakly dependent random field. Research Report 232. Dept. Theoretical Statistics, Aarhus Univ.

[11] KREISS, J. P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15** 112–133.

[12] LAHIRI, S. N. (1993). Refinements in asymptotic expansions for sums of weakly dependent random vectors. *Ann. Probab.* **21** 791–799.

[13] NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-negative Operators.* Cambridge Univ. Press.

[14] OZAKI, T. (1980). Non-linear time series models for non-linear vibrations. *J. Appl. Probab.* **17** 84–93.

[15] PETROV, YU. V. (1975). *Sums of Independent Random Variables.* Springer, Berlin.

[16] PETRUCELLI, J. D. and WOOLFORD, S. W. (1984). A threshold AR(1) model. *J. Appl. Probab.* **21** 270–286.

[17] PRIESTLEY, M. B. (1988). *Non-linear and Non-stationary Time Series Analysis.* Academic, London.

[18] TANIGUCHI, M. (1983). On the second order asymptotic efficiency of estimators of Gaussian ARMA processes. *Ann. Statist.* **11** 157–169.

[19] TANIGUCHI, M. (1984). Validity of Edgeworth expansions for statistics of time series. *J. Time Ser. Anal.* **5** 37–51.

[20] TANIGUCHI, M. (1986). Third order asymptotic properties of maximum likelihood estimators for Gaussian ARMA processes. *J. Multivariate Anal.* **18** 1–31.

[21] TJØSTHEIM, D. (1990). Nonlinear time series and Markov chains. *Adv. in Appl. Probab.* **22** 587–611.

[22] TONG, H. (1978). On a threshold model. In *Pattern Recognition and Signal Processing* (C. H. Chen, ed.). Sijthoff and Nordhoff, Alphen aan den Rijn, The Netherlands.

[23] TONG, H. (1990). *Non-linear Time Series: A Dynamical System Approach.* Oxford Univ. Press.

[24] TWEEDIE, R. L. (1983). Criteria for rates of convergence of Markov chains, with applications to queueing theory. In *Papers in Probability, Statistics and Analysis* (J. F. C. Kingman and G. E. H. Reuter, eds.) 260-276. Cambridge Univ. Press.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BIELEFELD
33501 BIELEFELD
GERMANY

DEPARTMENT OF ECONOMICS
UNIVERSITY OF KARLSRUHE
76128 KARLSRUHE
GERMANY