Electronic Journal of Statistics
Vol. 10 (2016) 3723–3751
ISSN: 1935-7524
DOI: 10.1214/16-EJS1213

Stability and asymptotics for autoregressive processes

Likai Chen and Wei Biao Wu*

Department of Statistics, University of Chicago
Chicago, IL 60637, USA
e-mail: lkchen@galton.uchicago.edu; wbwu@galton.uchicago.edu

Abstract: The paper studies infinite order autoregressive models for both temporal and spatial processes. We present sufficient conditions for the existence of stationary distributions. To understand the underlying dynamics and to capture the dependence structure, we introduce functional dependence measures and relate them with Lipschitz coefficients of the data-generating mechanisms. Our stability result allows both short- and long-range dependence. With functional dependence measures, we can establish an asymptotic theory for the underlying processes.

MSC 2010 subject classifications: 62M10, 62E20.
Keywords and phrases: Autoregressive models, Markov process, nonlinear time series, stationarity, functional dependence measure, invariance principle.

Received July 2016.

Contents

*Supported in part by NSF Grant DMS-14-05410.

1. Introduction

Nonlinear autoregressive (AR) processes have been extensively studied in the literature; see Priestley [25], Tong [31], Tjøstheim [30], Fan and Yao [12] and Wu and Shao [35], among others. To study such processes, one needs to deal with two fundamental issues: stability and asymptotic theory. For the former, one should develop sufficient conditions on the mechanism of the underlying process so that it can have a stationary solution. The asymptotic theory is useful for the related statistical inference. In this paper we shall consider both issues for a general class of nonlinear AR(∞) models.

To fix the idea, we adopt the following setting. Let $\epsilon_t, t \in \mathbb{Z}$, be independent and identically distributed (i.i.d.) random elements on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Consider

$$X_t = F_{\epsilon_t}(X_{t-1}, X_{t-2}, X_{t-3}, \ldots), \quad t \in \mathbb{Z}, \tag{1}$$

where $F_{\epsilon_t}(x_1, x_2, \ldots) = F(x_1, x_2, \ldots; \epsilon_t)$ is a real-valued function and can be viewed as the data generating mechanism of the process (X_t). We can view (1) as an AR(∞) process. By Wu and Shao [35], for the special AR(1) process

$$X_t = F_{\epsilon_t}(X_{t-1}), \tag{2}$$

assuming that there exists x_* such that $\mathbb{E}|F_{\epsilon_0}(x_*)|^p < \infty, p \geq 1$, and the contraction condition

$$L := \sup_{x \neq x'} \frac{\|F_{\epsilon_0}(x) - F_{\epsilon_0}(x')\|_p}{|x' - x|} < 1, \tag{3}$$

where $\|Z\|_p = (\mathbb{E}|Z|^p)^{1/p}$, then there exists a stationary solution of the form

$$X_t = g(\epsilon_t, \epsilon_{t-1}, \ldots) = g(\xi_t), \tag{4}$$

where g is a measurable function and $\xi_t = (\epsilon_t, \epsilon_{t-1}, \ldots)$ is the shift process. See Diaconis and Freedman [7] and Jarner and Tweedie [18] for related contributions. If condition (3) fails with $L = 1$, then (X_t) may not have a stationary solution. A prominent example is the random walk $X_t = X_{t-1} + \epsilon_t$ which has $L = 1$. Shao and Wu [29] considered the AR(d) processes with finite lag d:

$$X_t = F_{\epsilon_t}(X_{t-1}, \ldots, X_{t-d}), \quad t \in \mathbb{Z}, \tag{5}$$

and obtained a similar result: (5) has a stationary solution if

$$\sum_{i=1}^{d} a_i < 1, \tag{6}$$

where $a_i \geq 0$ are Lipschitz constants: for all $s_1, \ldots, s_d; w_1, \ldots, w_d$:

$$\|F_{\epsilon_t}(s_1, \ldots, s_d) - F_{\epsilon_t}(w_1, \ldots, w_d)\|_p \leq \sum_{i=1}^{d} a_i |s_i - w_i|.$$

For the AR(∞) process (1), it turns out that, interestingly, contraction condition like (3) may not be needed for stationarity. For example, consider the fractional integration model

$$(1-B)^d X_t = \epsilon_t, \quad 0 < d < 1/2, \tag{7}$$

where $BX_t = X_{t-1}$ is the backshift operator. We can rewrite (7) in the form of (1) with

$$X_t = \sum_{k \geq 1} a_k X_{t-k} + \epsilon_t, \text{ where } a_k = -\frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)}.$$

Then the Lipschitz constants (a_k) for the corresponding linear function F have sum $\sum_{k \geq 1} a_k = 1$, while (7) does have a stationary solution since $0 < d < 1/2$. In Section 2 we shall study the stability problem for AR(∞) long memory processes. Hence differently from (2), AR(∞) models can allow both short- and long-range dependence.

For extension to spatial processes, we consider the simultaneous autoregressive scheme

$$X_t = G_{\epsilon_t}(\{X_{t+v},\ v \neq \mathbf{0}\}), \qquad t, v \in \mathbb{Z}^d, \tag{8}$$

where $G_{\epsilon_t}(\cdot) = G(\cdot; \epsilon_t)$ is a real-valued function. Let $\eta_t = (\epsilon_{t-v}, v \in \mathbb{Z}^d)$. We say that

$$X_t = g(\{\epsilon_{t-v}, v \in \mathbb{Z}^d\}) = g(\eta_t), \tag{9}$$

where $g(\cdot)$ is a measurable function, is a stationary solution if it satisfies the relation (8). When $d = 1$, then (8) reduces to the two-sided AR(∞) process

$$X_t = G_{\epsilon_t}(\ldots, X_{t-2}, X_{t-1}, X_{t+1}, X_{t+2}, \ldots). \tag{10}$$

Differently from (1), the autoregressive scheme (10) allows non-causality. Properties for spatial processes are studied by Whittle [33] and Besag [3] among others. Gaussian and linear spatial processes have been widely studied in the literature. For linear processes Whittle [33] proposed ways to transform bilateral models to unilateral ones so that results in time series can be applied. The case with nonlinear processes is more challenging. Here, we will study stationary distributions for bilateral models directly using an idea which is similar to loopy propagation under a short-range dependence condition.

To perform statistical inference for the process (X_t) such as hypothesis testing and construction of confidence intervals, we need to establish an asymptotic theory. In particular we will present a central limit theorem and a Gaussian approximation result. To this end, we need to measure the decay speed of dependence. In this paper we adopt the framework of functional dependence measures introduced by Wu [34] which is easy to work with for a broad class of functions and enables us to obtain sharp approximation rate. The main task lies in building a convolution relationship between Lipschitz coefficients of $F_{\epsilon_t}(\cdot)$ and functional dependence measures of the underlying processes. See Section 3.1 for details.

Using the functional dependence decay results we can derive a CLT and a quenched CLT for the stationary process, and a Gaussian approximation which can be used in various applications such as change-point analysis. For the augmented GARCH(1,1) model, Aue, Berkes and Horváth [1] obtained a Gaussian approximation rate $o(n^{5/12+\epsilon})$, $\epsilon > 0$. Using our result, we can derive a sharper Gaussian approximation for both Model (1) and augmented GARCH(∞), and our rate is optimal in view of the classical Gaussian approximation for i.i.d. random variables by Komlós, Major and Tusnády [22].

The paper is organized as follows. In Section 2 we present sufficient conditions for the existence of stationary distributions for both temporal and spatial models. It turns out that, interestingly, our result can also be applied to random coefficient models which are not in the form of (1); see Section 2.2 where the augmented GARCH(∞) processes are discussed. In Section 3, we introduce functional dependence measure and apply it to our models. Based on that we derive the relationship of decay rate between coefficients and functional dependence measure, and develop various asymptotic results. Proofs are given in Section 4.

2. Stationary distribution

In Section 2.1 we shall present sufficient conditions for the existence of stationary distributions of model (1) with short-range dependence. The theorem is also applicable to random coefficient models; see Section 2.2. Section 2.3 (resp. 2.4) concerns long-range dependent processes (resp. simultaneous autoregressive schemes), while Section 2.5 deals with extensions to non-stationary processes.

2.1. Short-range dependent AR(∞) processes

To state our main stability result for the process (1), we shall introduce a Lipschitz type condition. For a random variable Z, we say $Z \in \mathcal{L}^p$, if $\|Z\|_p := [\mathbb{E}|Z|^p]^{1/p} < \infty$.

Condition 1. *(Stochastic Lipschitz Continuity). Assume that there exist constants $a_k \geq 0$ $k \in \mathbb{N}$, such that for all $\mathbf{w} = (w_1, w_2, \dots)$ and $\mathbf{s} = (s_1, s_2, \dots)$, $F_{\epsilon_0}(\boldsymbol{w}) \in \mathcal{L}^p$, $p \geq 1$ and*

$$\|F_{\epsilon_0}(\mathbf{w}) - F_{\epsilon_0}(\mathbf{s})\|_p \leq \sum_{k=1}^{\infty} a_k |w_k - s_k|. \tag{11}$$

Definition 1. *On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{Z}}, \mathbb{P})$, a process (X_t) is said to be adapted, if for each t, X_t is $\mathcal{F}_t$ measurable.*

From now on, always let $\mathcal{F}_t$ be the σ-field generated by $\xi_t = (\epsilon_t, \epsilon_{t-1}, ...)$.

Theorem 1. *For* (1), *assume Condition 1 holds with* $p \geq 1$ *and* $(a_k)_{k \geq 1}$ *satisfy*

$$\sum_{k=1}^{\infty} a_k < 1. \tag{12}$$

Then there exists a unique strictly stationary solution in $\mathcal{L}^p$ *adapted to* $(\mathcal{F}_t)_{t \in \mathbb{Z}}$.

To incorporate the case with $0 < p < 1$, we need to slightly modify Condition 1.

Corollary 1. *Let* $0 < p < 1$. *Assume Condition 1 with* (11) *replaced by*

$$\mathbb{E}|F_{\epsilon_0}(w_1, w_2, \ldots) - F_{\epsilon_0}(s_1, s_2, \ldots)|^p \leq \sum_{k=1}^{\infty} a_k |w_k - s_k|^p. \tag{13}$$

Further assume (12). Then (1) *has a unique strictly stationary distribution in* $\mathcal{L}^p$ *adapted to* $(\mathcal{F}_t)_{t \in \mathbb{Z}}$.

Example 1. Let $F_{\epsilon_0}(s_1, s_2, \ldots)$ be Lipschitz continuous with constants $h_k(\epsilon_0)$, namely,

$$|F_{\epsilon_0}(s_1, s_2, \ldots) - F_{\epsilon_0}(w_1, w_2, \ldots)| \leq \sum_{k=1}^{\infty} h_k(\epsilon_0) |w_k - s_k|. \tag{14}$$

If $h_k(\epsilon_0) \in \mathcal{L}^p$, then we have (11) and (13) with $a_k = \|h_k(\epsilon_0)\|_p^{p'}$, where $p' = \min(p, 1)$. For example, consider the infinite order bilinear process with $F_{\epsilon_0}(s_1, s_2, \ldots) = \epsilon_0 + \sum_{i=1}^{\infty}(u_i + v_i \epsilon_0) s_i$, where u_i, v_i are real parameters, then $h_i(\epsilon_0) = |u_i + v_i \epsilon_0|$. Finite order bilinear processes are considered in Granger and Andersen [16] and Rao and Gabr [27]. □

To study the existence of stationary solutions, we use the idea of backward iteration, enlightened by the "coupling from the past" algorithm in Propp and Wilson [26]. Traditionally forward iterations are considered. For the simple Markov chain example (2), for convenience assume that 0 is in the state space, one checks whether the forward iteration

$$\tilde{X}_0 = 0,\ \tilde{X}_t = F_{\epsilon_t} \circ \ldots \circ F_{\epsilon_1}(\tilde{X}_0) = F_{\epsilon_t}(\tilde{X}_{t-1}),\ t = 1, 2, \ldots,$$

converges weakly as $t \to \infty$. If it converges weakly to a distribution π (say), then π is a stationary distribution. For the backward iteration, we let

$$X_t^{(t-n)} = F_{\epsilon_t} \circ \ldots \circ F_{\epsilon_{t-n+1}}(0) = F_{\epsilon_t}(X_{t-1}^{(t-n)}),\ n = 1, 2, \ldots;\ X_t^{(t)} = 0,\ t \in \mathbb{Z}.$$

Under suitable conditions on $F_{\epsilon_t}(\cdot)$, $X_t^{(t-n)}$ converges almost surely as $n \to \infty$ and the limit, denoted by X_t, satisfies (2). For the AR(∞) process (1), we follow a similar idea to generate the sequence. Let

$$X_t^{(t)} = 0,\ X_t^{(t-n)} = F_{\epsilon_t}(X_{t-1}^{(t-n)}, \ldots, X_{t-n}^{(t-n)}, 0, 0, \ldots),\ n = 1, 2, \ldots. \tag{15}$$

Note that $X_n^{(m)}$ has the same distribution as $X_{n+k}^{(m+k)}$. Under suitable conditions, the sum $\sum_{n=1}^{\infty} |X_t^{(t-n+1)} - X_t^{(t-n)}|$ exists and thus $X_t^{(t-n)}$ converges almost surely to X_t (say) as $n \to \infty$. In the proof of Theorem 1 we shall make the latter idea rigorous.

2.2. Random coefficient models

Interestingly, our stationarity theory also applies to the following process

$$X_t = f_0(\epsilon_{t-1}, \epsilon_{t-2}, \dots) + \sum_{k=1}^{\infty} f_k(\epsilon_{t-1}, ..., \epsilon_{t-k}) X_{t-k}, \quad t \in \mathbb{Z}, \tag{16}$$

where ϵ_t are i.i.d. and $\{f_k\}_{k\geq 0}$ are real-valued functions.

Let $\Lambda : \mathbb{R}^+ \to \mathbb{R}$ be an invertible function. Based on (16), we can define the augmented GARCH process Y_t by restricting $f_k(\epsilon_{t-1}, ..., \epsilon_{t-k})$ to $f_k(\epsilon_{t-k})$ and letting

$$\sigma_t^2 = \Lambda^{-1}(X_t), \quad Y_t = \sigma_t \epsilon_t, \tag{17}$$

provided that $\Lambda(x)$ is invertible; see Duan [10]. Augmented GARCH contains many commonly used ARCH models.

Example 2. [28] considered the general form ARCH(∞) process: let $\xi_t,\ t \in \mathbb{Z}$, be i.i.d. nonnegative random variables and

$$\rho_t = \beta_0 + \sum_{k=1}^{\infty} \beta_k \xi_{t-k} \rho_{t-k}, \quad Y_t = \rho_t \xi_t, \tag{18}$$

where the parameters $\beta_k \geq 0, k \geq 0$. □

Example 3. Motivated by a Box-Cox transformation of the conditional variances, the power GARCH has the form $Y_t = \sigma_t \epsilon_t$ with

$$\sigma_t^\delta = \alpha_0 + \sum_{i\geq 1} \alpha_i \sigma_{t-i}^\delta + \sum_{i\geq 1} \beta_i |X_{t-i}|^\delta = \alpha_0 + \sum_{i\geq 1} (\alpha_i + \beta_i |\epsilon_{t-i}|^\delta) \sigma_{t-i}^\delta.$$

The lag one version was considered by Carrasco and Chen [5] and they studied its moment and mixing properties. □

Example 4. [23] studied the Exponential GARCH $Y_t = \sigma_t \epsilon_t$ with

$$\log \sigma_t^2 = w + \sum_{k\geq 1} \beta_k (\theta \epsilon_{t-k} + \lambda(|\epsilon_{t-k}| - \mathbb{E}|\epsilon_{t-k}|)) + \sum_{k\geq 1} \alpha_k \log \sigma_{t-k}^2. \quad □$$

Example 5. Ding, Granger and Engle [8] introduced the asymmetric power ARCH (PARCH) $Y_t = \sigma_t \epsilon_t$ with

$$\begin{aligned} \sigma_t^\delta &= \alpha_0 + \sum_{i\geq 1} \alpha_i (|X_{t-i}| - \gamma_i Y_{t-i})^\delta + \sum_{j\geq 1} \beta_j \sigma_{t-i}^\delta \\ &= \alpha_0 + \sum_{i\geq 1} [\alpha_i (|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^\delta + \beta_i] \sigma_{t-i}^\delta. \end{aligned} \quad □$$

See [10] and Aue, Berkes and Horváth [1] for other GARCH models. Aue, Berkes and Horváth [1] derived Gaussian approximation for partial sum processes. [17] used blocking technique to derive various asymptotic properties. The stability and asymptotic properties for infinite lag Augmented GARCH have not been discussed in the literature. Here we can tackle the latter problem by using the similar method as for process (1).

Corollary 2. *Assume that* $f_0(\epsilon_0, \epsilon_{-1}, \ldots)$, $f_k(\epsilon_1, ..., \epsilon_k)$, $k \geq 1$, *are in* $\mathcal{L}^p$ *with* $p \geq 1$,

$$\sum_{k\geq 1} \|f_k(\epsilon_1, ..., \epsilon_k)\|_p < 1. \tag{19}$$

Then process (16) *has a unique* $\mathcal{L}^p$ *strictly stationary solution with form* $X_t = h(\epsilon_{t-1}, \epsilon_{t-2}, \ldots)$ *where* h *is a measurable function. Moreover, the expectation*

$$\mathbb{E}X_t = \Big(1 - \sum_{k\geq 1} \mathbb{E}f_k(\epsilon_1, ..., \epsilon_k)\Big)^{-1} \mathbb{E}(f_0(\epsilon_0, \epsilon_{-1}, \ldots)).$$

Remark 1. Inequalities (12) and (19) can be viewed as contracting conditions. In situations that they equal 1, extra assumptions are needed to guarantee a stationary solution; see Section 2.3. Such a process can be long-memory. □

For $0 < p < 1$, we need a slight modification of condition (19) and consider an approach similar to Corollary 1. Douc, Roueff and Soulier [9] used Volterra expansion to establish the stationary distribution. Here we can deal with a more general situation.

Corollary 3. *If* $\mathbb{E}|f_0(\epsilon_0, \epsilon_{-1}, \ldots)|^p < \infty$ *and* $\sum_{k\geq 1} \mathbb{E}|f_k(\epsilon_1, ..., \epsilon_k)|^p < 1$, $0 < p < 1$, *then there exists a stationary solution for (16) which has a finite pth norm.*

Remark 2. Notice that in Model (16), if all coefficients $f_k \geq 0$ for $k \geq 0$, then the stationary distribution for (16) is nonnegative. This is useful in checking the existence of $\Lambda^{-1}(X_t)$, for example if $\Lambda(x) = x$, then for $\sigma_t^2 = \Lambda^{-1}(X_t)$, we require X_t to be nonnegative. □

Example 6. For the ARCH(∞) model (18), we let in (17) $X_t = \rho_t, \epsilon_k = \xi_k$ and $f_k(x) = \beta_k x$. Corollary 2 gives the sufficient condition for stationarity of (Y_t):

$$\left(\sum_{k=1}^{\infty} \beta_k\right) \mathbb{E}\xi_0 < 1.$$

The above condition is also proposed in Giraitis, Kokoszka and Leipus [13]. With the special structure (18), they apply the Volterra series expansion, which is also used in subsequent works; see Kazakevicius and Leipus [21], Giraitis, Leipus and Surgailis [14, 15] among others. In comparison, our treatment does not rely on this special structure. Instead we use a convolution relation and backward generation which can be applied to a broader class of nonlinear models. □

2.3. Long-range dependent AR(∞) processes

If condition (12) in Theorem 1 is violated and $\sum_{k=1}^{\infty} a_k = 1$, as the fractional integration process (7) shows, a stationary solution can still possibly exist. For convenience we still assume that 0 is in the state space. For process (1) define $X_{t+k}^{(t)}$ recursively:

$$X_t^{(t)} = 0, \qquad X_{t+k}^{(t)} = F_{\epsilon_{t+k}}(X_{t+k-1}^{(t)}, \ldots, X_t^{(t)}, 0, \ldots), \quad k \geq 1. \tag{20}$$

Condition 2. *There exists some $p > 0$ such that*

$$\sum_{k=0}^{\infty} \|\mathbb{E}\big(X_0^{(-k-1)}|\mathcal{G}_{-k+1}\big) - X_0^{(-k)}\|_p < \infty, \ \textit{where } \mathcal{G}_t = (\epsilon_t, \epsilon_{t+1}, \ldots).$$

We can view $\mathcal{G}_t$ as a backward shift process. This condition can be interpreted as that the cumulative influence of the initial state is finite. It holds for many processes. Below we shall consider the example of random coefficient AR models.

Example 7. Let $f_k,\ k \geq 0$, be real valued functions with $\mathbb{E}f_0(\epsilon_t) = 0$. Consider the process

$$X_t = \sum_{k\geq 1} f_k(\epsilon_t) X_{t-k} + f_0(\epsilon_t). \tag{21}$$

A special case of (21) is the bilinear model with $f_k(\epsilon_t) = u_k + v_k\epsilon_t$, where u_k, v_k are real parameters. Construct sequence $(X_t^{(t-n)})$ by (20) based on (21):

$$X_t^{(t)} = 0, \qquad X_t^{(t-n)} = \sum_{k=1}^{n} f_k(\epsilon_t) X_{t-k}^{(t-n)} + f_0(\epsilon_t).$$

Thus the difference follows

$$\mathbb{E}(X_t^{(t-n-1)}|\mathcal{G}_{t-n+1}) - X_t^{(t-n)} = \sum_{k=1}^{n} f_k(\epsilon_t)\Big(\mathbb{E}(X_{t-k}^{(t-n-1)}|\mathcal{G}_{t-n+1}) - X_{t-k}^{(t-n)}\Big). \tag{22}$$

Notice the initial value for above iteration is

$$\mathbb{E}(X_{t-n}^{(t-n-1)}|\mathcal{G}_{t-n+1}) - X_{t-n}^{(t-n)} = \mathbb{E}f_0(\epsilon_{t-n}) = 0.$$

Consequently by induction and (22) we get $\mathbb{E}(X_t^{(t-n-1)}|\mathcal{G}_{t-n+1}) - X_t^{(t-n)} = 0$, hence Condition 2 naturally holds. □

For a sequence $l = (l_1, l_2, \ldots)$, denote the generating function $Q_l(s) = \sum_{n=1}^{\infty} l_n s^n$.

Theorem 2. *For process (1), let (11) in Condition 1 hold with coefficients $a = (a_1, a_2, \ldots)$, $p \geq 2$. Assume $F_{\epsilon_0}(0, 0, \ldots) \in \mathcal{L}^p$ and Condition 2. Assume $Q_a(1) = 1$ and*

$$\int_0^{2\pi} \frac{1}{|1 - Q_a(e^{i\theta})|^2} d\theta < \infty. \tag{23}$$

Then (1) exists a stationary $\mathcal{L}^p$ solution.

We now apply Theorem 2 to (21). Assume that $a_k = \|f_k(\epsilon_0)\|_p < \infty$, $p \geq 2$ and (23) holds, then there exists a stationary $\mathcal{L}^p$ solution.

REMARK 3. Note that $Q_a(1) = 1$. If $1-Q_a(s) = (1-s)^d G(s)$ where $0 < d < 1/2$ and $|G(s)|$ is bounded from below by a constant $c > 0$, then (23) holds. Also we can replace (23) by some corresponding conditions on the tail sum,

$$\int_0^{2\pi} \frac{1}{|Q_A(e^{i\theta})(e^{-i\theta}-1)|^2} d\theta < \infty, \quad A_1 = 1, \tag{24}$$

where $A = (A_1, A_2, ...)$, $A_k = \sum_{s \geq k} a_s$ and $Q_A(s) = \sum_{k\geq 1} A_k s^k$, then we obtain (23) since

$$Q_a(s) = \sum_{k\geq 1}(A_k - A_{k+1})s^k = \sum_{k\geq 1} A_k(s^k - s^{k-1}) + A_1 = Q_A(s)(1-s^{-1}) + 1.$$

□

In the following example we shall apply Tauberian's Theorem to verify (24). For sequenceses (a_n) and (b_n), denote $a_n \sim b_n$, if $a_n/b_n \to 1$ as $n \to \infty$.

EXAMPLE 8. Assume $A_k \sim c_0 k^{-\alpha}$ with $c_0 > 0$, $0 < \alpha < 1/2$ and $A_0 = 1$, then $A_0 + \ldots + A_n \sim c_0(1-\alpha)^{-1} n^{1-\alpha}$. If $Q_A(s)$ has no zero root for $|s| \leq 1$, by Tauberian's Theorem, $Q_A(s) \sim c_1(1-s)^{\alpha-1}$, $s \to 1-$ where $c_1 = c_0\Gamma(2-\alpha)(1-\alpha)^{-1}$. Thus

$$Q_A(s)(s^{-1}-1) \sim c_1 s^{-1}(1-s)^{\alpha}, \quad s \to 1-. \tag{25}$$

Since $2\alpha < 1$, $|1-e^{i\theta}|^{-2\alpha}$ is integrable for θ around a neighbor of 0 and 2π. By (25) we have (24). If $\alpha \geq 1/2$, then the integral in (24) is ∞ and τ_n are no longer square summable. For example, if $a_k \sim c_1 k^{-1-\alpha}$ with $\alpha \geq 1/2$ and $\sum_{k=1}^{\infty} a_k = 1$, the stationary solution does not exist. □

An analagous result can be derived for process (16). For $t \in \mathbb{Z}$, $n \in \mathbb{N}$ define recursively

$$X_t^{(t)} = 0, \quad X_t^{(t-n)} = \mathbb{E}(f_0(\epsilon_{t-1}, \epsilon_{t-2}, \ldots)|\mathcal{G}_{t-n}) + \sum_{k=1}^{n} f_k(\epsilon_{t-1}, ..., \epsilon_{t-k}) X_{t-k}^{(t-n)}. \tag{26}$$

Let $\eta_j = f_0(\epsilon_j, \epsilon_{j-1}, \ldots)$. As in (32) let functional dependence measure $\delta_{j,p} = \|\eta_j - \eta_{j,0}\|_p$, $j \geq 0$.

COROLLARY 4. *For process (16), assume that $f_0(\epsilon_0, \epsilon_{-1}, \ldots)$, $f_k(\epsilon_0)$, $k \geq 1$ are in $\mathcal{L}^p$ with $p \geq 2$ and $\mathbb{E}(f_0(\epsilon_1, \epsilon_2, ...)) = 0$. Denote $a_k = \|f_k(\epsilon_1, ..., \epsilon_k)\|_p$ and $b_k = \delta_{k,p}$. Assume $\sum_{k\geq 1} a_k = 1$, $\sum_{k\geq 1} b_k^2 < \infty$ and (53). Then (16) exists a stationary $\mathcal{L}^p$ solution.*

Remark 4. When $\mathbb{E}f_0(\epsilon_0, \epsilon_1, ...) = 0$, we have

$$\sum_{k=0}^{\infty} \|\mathbb{E}\big(X_0^{(-k-1)}|\mathcal{G}_{-k}\big) - X_0^{(-k)}\|_p < \infty, \tag{27}$$

which is an analogue of Condition 2 for process (16) and it also implies the cumulative influence of initial state is finite. From the recursive equation

$$X_t^{(t-n)} - \mathbb{E}(X_t^{(t-n-1)}|\mathcal{G}_{t-n}) = \sum_{k=1}^{n} f_k(\epsilon_{t-1}, ..., \epsilon_{t-k})\big(X_{t-k}^{(t-n)} - \mathbb{E}(X_{t-k}^{(t-n-1)}|\mathcal{G}_{t-n})\big),$$

and that the initial value $X_t^{(t)} - \mathbb{E}(X_t^{(t-1)}|\mathcal{G}_t) = -\mathbb{E}f_0(\epsilon_0, \epsilon_1, ...) = 0$, we have $X_t^{(t-n)} = \mathbb{E}(X_t^{(t-n-1)}|\mathcal{G}_{t-n})$. Thus (27) directly follows. □

2.4. Simultaneous autoregressive schemes

In this section we shall consider stationary distribution for spatial models. Linear spatial processes were studied in Whittle [33] and Besag [3] among others. For the form (10), which can be viewed as a bilateral version of Model (1), we adopt an idea which is similar to the loopy propagation commonly used in machine learning. First set the initial values to be zero, and then update them based on previous results:

$$X_t^{\{0\}} = 0, \quad X_t^{\{k+1\}} = G_{\epsilon_t}(\ldots, X_{t-1}^{\{k\}}, X_{t+1}^{\{k\}}, \ldots) \quad t \in \mathbb{Z},\, k = 0, 1, \ldots. \tag{28}$$

Similarly, we can set the initial value and update them for the general form (8). Under suitable conditions on G (cf Condition 3), $X_t^{\{k\}}$ has a limit as $k \to \infty$; cf Theorem 3.

Condition 3. *There exist constants $a_v \geq 0$, $v \in \mathbb{Z}^d$, and $p \geq 1$, such that*

$$|G_{\epsilon_0}(\{w_v,\ v \neq \mathbf{0}\}) - G_{\epsilon_0}(\{s_v,\ v \neq \mathbf{0}\})| \leq \sum_{v \neq \mathbf{0}} a_v |w_v - s_v| \tag{29}$$

holds for all $w_v, s_v \in \mathbb{R}$, and $\|G_{\epsilon_0}(\mathbf{o})\|_p < \infty$, where $\mathbf{o} = \{o_v,\ v \neq \mathbf{0}\}$ with all $o_v = 0$.

Theorem 3. *Assume Condition 3 with some $p \geq 1$ and the contraction condition $\sum_{k \neq \mathbf{0}} a_k < 1$. Then there exists a unique $\mathcal{L}^p$ stationary solution for (8).*

Example 9. Let $h_v, k_v, v \in \mathbb{Z}^d$, be real coefficients. Consider the spatial threshold AR model

$$X_k = \sum_{v \neq \mathbf{0}} (h_v \max(X_{k-v}, 0) + k_v \min(X_{k-v}, 0)) + \epsilon_k, \quad k \in \mathbb{Z}^d,$$

with $\epsilon_k \in \mathcal{L}^p$, $p \geq 1$. Then Condition 3 holds with $a_v = \max(|h_v|, |k_v|)$. This example is a spatial generalization of threshold AR processes of Tong [31]. □

2.5. Extension to non-stationary processes

Though stationary models work well in many cases, they may be unsuitable for more complicated situations: when the location domain has boundary or is non-lattice, due to different configuration of neighborhood, there is hardly a geometric or physical base to assume stationarity. For those irregular cases, the function G (8) can be location dependent. Paulik, Das and Loh [24], Brunsdon, Fotheringham and Charlton [4] studied linear cases and Jenish and Prucha [19, 20] derived LLN and CLT for nonlinear situations under mixing or near-epoch dependence. Here we consider the model:

$$X_t = G_{(t)}(\{X_v,\ v \in \Theta_t\}; \epsilon_t), \qquad t \in \Theta, \tag{30}$$

where Θ is a set with countably many points, $\Theta_t \subseteq \Theta$ which may change with t and the data generating mechanism $G_{(t)}$ is a real-valued measurable function.

The above setting may appear in practice, for instance if the lattice exists boundary, or if we are not dealing with regular lattice but just certain undirected graphs etc, then the relative configuration for each point will be different and it is no longer appropriate to assume same function G for every point.

For this more general situation, under certain uniform bounded conditions on $G_{(v)}$, there exists measurable function H_t such that $X_t = H_t(\xi)$ satisfies the system (30) where $\xi = (\epsilon_t)_{t\in\Theta}$.

Condition 4. *Assume there exists coefficients $a_{t,s} \geq 0$, $\rho < 1$, $M < \infty$ and $p \geq 1$ such that the data generating mechanism $G_{(t)}$ satisfies*

$$|G_{(t)}(\{w_s, s \in \Theta_t\}; \epsilon_t) - G_{(t)}(\{v_s, s \in \Theta_t\}; \epsilon_t)| \leq \sum_{s\in\Theta_t} a_{t,s}|w_s - v_s|,$$

$$\textit{with} \quad \sup_{t\in\Theta} \sum_{s\in\Theta_t} a_{t,s} \leq \rho \quad \textit{and} \quad \sup_{t\in\Theta} \|G_{(t)}(\ldots, 0, \ldots; \epsilon_0)\|_p \leq M.$$

Corollary 5. *If $(G_{(t)})_{t\in\Theta}$ satisfy Condition 4, then there exists a measurable function H_t such that $X_t = H_t(\xi) \in \mathcal{L}^p$ and (30) holds.*

3. Functional dependence measures and asymptotic results

3.1. Functional dependence measures

In this section, we shall compute functional dependence measure introduced by Wu [34] for the processes (1) and (8). In view of (4) and (9), we consider the form

$$X_i = g(\epsilon_{i-j}, j \in \Xi), \quad i \in \Theta, \tag{31}$$

where data generating mechanism g is a real-valued measurable function such that X_i is properly defined, and $\epsilon_i, i \in \Theta$, are i.i.d. random variables. For model (1) with representation (4), $\Xi = \{0, 1, 2, \ldots\}$ and $\Theta = \mathbb{Z}$. For spatial process on lattice in $\mathbb{Z}^d$, both Ξ and Θ are $\mathbb{Z}^d$.

Assume that $X_i \in \mathcal{L}^p,\ p > 0$. Let $\epsilon'_j, \epsilon_i,\ i, j \in \Theta$, be i.i.d. random variables. In view of (31), X_t is a random variable constructed on the underlying random sequence $(\epsilon_i,\ i \in \Theta)$. Therefore instead of directly describing the relationship between X_s and X_t, we use functional dependence measure to capture the extent to which X_t depends on the underlying random variables (ϵ_i). Change ϵ_{t-i} to ϵ'_{t-i} and keep other ϵ_i unchanged, we get a copy of X_t which is denoted by $X_{t,t-i}$. By stationarity, the functional dependence measure

$$\delta_{i,p} = \|X_t - X_{t,t-i}\|_p = \|X_i - X_{i,0}\|_p. \tag{32}$$

To deal with functional dependence measures, we need the following Theorem 4 which concerns magnitudes of the convolved sequences. The result is of independent interest. Case (v) provides an explicit decay speed of the functional dependence measure and it implies that the bound in case (ii) is sharp.

THEOREM 4. *Let $(u_k)_{k\geq 1}$ and $(v_k)_{k\geq 1}$ be nonnegative sequences with $A := \sum_{k=1}^{\infty} u_k < 1$ and $B := \sum_{k=1}^{\infty} v_k < \infty$. Define $(\tau_i)_{i\geq 0}$ recursively by $\tau_0 = c \geq 0$ and for $n \geq 1$,*

$$\tau_n = \sum_{k=1}^{n} u_k \tau_{n-k} + v_n. \tag{33}$$

Let $\delta = 1 - A$. Then (i) $\sum_{k=1}^{\infty} \tau_k = (B + c)/\delta$. For $\theta > 1$ we have the following: (ii) if $u_n + v_n = O(n^{-\theta})$, then $\tau_n = O(n^{-\theta})$; (iii) if $c > 0$, $u_n + v_n \asymp n^{-\theta}$, then $\tau_n \asymp n^{-\theta}$; (iv) if $\sum_{k=n}^{\infty}(u_k + v_k) = O(n^{-(\theta-1)})$, then $\sum_{k=n}^{\infty} \tau_k = O(n^{-(\theta-1)})$; and (v) if $v_n \equiv 0$ and $u_n \sim Cn^{-\theta}$ for some $C > 0$, then $\tau_n \sim (Cc/\delta^2)n^{-\theta}$.

In the following two sections, we will apply this dependence measure into our models and derive relationship of decay rate between the functional dependence measure and Lipschitz coefficients. The functional dependence measure of the underlying procedures can be quite useful for further deriving asymptotic properties; cf Section 3.2–3.4.

3.1.1. *AR(∞) and random coefficient processes*

In this subsection we shall apply Theorem 4 to relate Lipchitz coefficients in (11) or (29) with functional dependence measures; cf. Corollary 6, Corollary 7. In view of Example 10 below, the bounds in these corollaries are sharp.

COROLLARY 6. *Let $(X_i)_{i\in\mathbb{Z}}$ be a nonlinear AR(∞) process (1). Assume conditions in Theorem 1 are satisfied with $p \geq 1$ and coefficients $(a_i)_{i\in\mathbb{N}}$. Recall the functional dependence measure $\delta_{i,p} = \|X_i - X_{i,0}\|_p$. Then (i) $\sum_{i\geq 0} \delta_{i,p} < \infty$. For $\theta > 1$ we have the following: (ii) if $a_n = O(n^{-\theta})$, then $\delta_{n,p} = O(n^{-\theta})$; (iii) if $\sum_{m\geq n} a_m = O(n^{1-\theta})$, then $\sum_{m\geq n} \delta_{m,p} = O(n^{1-\theta})$.*

EXAMPLE 10. For the special linear AR(∞) process:

$$X_t = \sum_{k\geq 1} a_k X_{t-k} + \epsilon_t,$$

where $a_k \geq 0$ and (ϵ_t) are i.i.d random variables. The corresponding MA(∞) process is

$$X_t = \sum_{k\geq 0} \tau_k \epsilon_{t-k}, \tag{34}$$

where τ_k satisfies the recursion

$$\tau_n = \sum_{i=1}^{n} a_k \tau_{n-k}, \quad \tau_0 = 1. \tag{35}$$

This is a special case of (33) with $u_n = a_n$, $v_n = 0$. By Theorem 4(v), if $a_n \sim cn^{-\theta}$ holds for some constant $c > 0$, then $\tau_n \sim cn^{-\theta}(1 - \sum_{k\geq 1} a_k)^{-2}$. $\square$

Before stating the proof of Corollary 6, we need some notation. Let $\epsilon_i, \epsilon'_j, i, j \in \mathbb{Z}$, be i.i.d. random variables, $\xi_t = (\epsilon_t, \epsilon_{t-1}, ...)$, $\xi_{t,0} = (\varepsilon_t, \ldots, \varepsilon_1, \varepsilon'_0, \varepsilon_{-1}, \ldots)$ and the coupled process $X_{t,0} = g(\xi_{t,0})$. Denote $\xi_i^j = (\epsilon_i, \epsilon_{i+1}, ..., \epsilon_j),\ i \leq j$.

Proof of Corollary 6. By Theorem 1, (1) has a stationary solution of form (4). Observe that the process $(X_{t,0})$ satisfy the recursion

$$X_{t,0} = F_{\epsilon_t}(X_{t-1,0}, X_{t-2,0}, \ldots) \text{ when } t \neq 0, \text{ and } X_{0,0} = F_{\epsilon'_0}(X_{-1}, X_{-2}, \ldots).$$

Since ϵ_t is independent of ξ_{t-1}, Condition 1 implies that, for $t \geq 1$,

$$\left(\mathbb{E}\left(|X_{t,0} - X_t|^p \middle| \xi_{t-1}\right)\right)^{1/p} \leq \sum_{k=1}^{\infty} a_k |X_{t-k,0} - X_{t-k}| = \sum_{k=1}^{t} a_k |X_{t-k,0} - X_{t-k}|.$$

By Minkowski's inequality,

$$\delta_{t,p} \leq \sum_{k=1}^{t} a_k \delta_{t-k,p}. \tag{36}$$

Let $u_n = a_n$, $v_n = 0$, $n \geq 1$, $\tau_0 = \delta_{0,p}$. Clearly by (36), $\delta_{i,p} \leq \tau_i \delta_{0,p}, i \geq 0$. Result follows from Theorem 4 (i) (ii) and (iv). $\square$

We can similarly have a corresponding result for the process (16).

COROLLARY 7. *Let $(X_i)_{i\in\mathbb{Z}}$ be the random coefficient AR(∞) process defined in (16). Assume that conditions in Corollary 2 are satisfied with $p \geq 1$ and coefficients $f_0(\xi_0)$, $f_k(\xi_1^k)$, $k \geq 1$. Denote $\theta_{k,p} = \|f_0(\xi_{k-1}) - f_0(\xi_{k-1,0})\|_p$ and*

$$a_k = \|f_k(\xi_1^k)\|_p, \quad b_k = \theta_{k,p} + \sum_{i>k} \|f_k(\xi_{k-1}^{k-i}) - f_k(\xi_{k-1,0}^{k-i})\|_p \|X_0\|_p. \tag{37}$$

Assume $b_k < \infty$. Then (i) $\sum_{i\geq 0} \delta_{i,p} < \infty$. For $\beta > 1$ we have the following: (ii) if $a_n + b_n = O(n^{-\beta})$, then $\delta_{n,p} = O(n^{-\beta})$; (iii) if $\sum_{m\geq n}(a_m + b_m) = O(n^{1-\beta})$, then $\sum_{m\geq n} \delta_{m,p} = O(n^{1-\beta})$.

Remark 5. For the PGARCH and asymmetric PGARCH, $\theta_{t,p} = 0$ for any t. □

Proof of Corollary 7. Change ϵ_0 to ϵ_0' and we can similarly obtain a new sequence $X_{t,0}$ which satisfies

$$\|X_t - X_{t,0}\|_p \leq \theta_{t,p} + \sum_{k=1}^{t-1} \|f_k(\xi_1^k)\|_p \|X_{t-k} - X_{t-k,0}\|_p + \sum_{k \geq t} \big\|f_k(\xi_{t-1}^{t-k}) - f_k(\xi_{t-1,0}^{t-k})\big\|_p \|X_0\|_p.$$

Let $u_n = a_n$ and $v_n = b_n$ as in (37). Then the result follows from Theorem 4. □

3.1.2. Functional dependence measure for simultaneous autoregressive schemes

Consider the process (8). Recall the generating mechanism of $(X_t^{\{k\}})_{t\in\mathbb{Z}^d}$ in (28) and let $(\delta_t^{\{k\}})_{t\in\mathbb{Z}^d}$ be the functional dependence measure, that is $\delta_t^{\{k\}} = \|X_t^{\{k\}} - X_{t,0}^{\{k\}}\|_p$. Write $X \in \mathcal{L}^p$, if $\|X\|_p < \infty$.

Condition 5. *Let $p \geq 1$. For the data generating mechanism G_{ϵ_0} in (8), assume that*

$$g(\epsilon_0, \epsilon_0') := \sup_{w_v, v\in\mathbb{Z}^d} |G_{\epsilon_0}(\{w_v,\ v\in\mathbb{Z}^d\}) - G_{\epsilon_0'}(\{w_v,\ v\in\mathbb{Z}^d\})| \in \mathcal{L}^p.$$

Condition 5 holds, for instance, when $G_{\epsilon_0}(\{w_v,\ v\in\mathbb{Z}^d\}) = F(\{w_v,\ v\in\mathbb{Z}^d\}) + f(\epsilon_0)$ with $f(\epsilon_0) \in \mathcal{L}^p$. Without loss of generality, set $\|g(\epsilon_0, \epsilon_0')\|_p = 1$. Define $(\tau_t^{\{k\}})_{t\in\mathbb{Z}^d}$, $k = 0, 1, \ldots$, recursively through

$$\tau_t^{\{0\}} = 0, \qquad \tau_t^{\{k+1\}} = \sum_{v\neq 0} a_v \tau_{t+v}^{\{k\}} + \mathbf{1}_{\{t=0\}}. \tag{38}$$

Then the functional dependence measure $\delta_t^{\{k\}} \leq \tau_t^{\{k\}}$. Under certain conditions we can develop results similar as Theorem 4 for $\tau_t^{\{k\}}$ and therefore also bound $\delta_t^{\{k\}}$. Since $X_t^{\{k\}}$ converges to X_t as $k \to \infty$, $\delta_t^{\{k\}}$ goes to δ_t, results are thus established for $(\delta_t)_{t\in\mathbb{Z}^d}$, where δ_t is the functional dependence measure of the stationary $\mathcal{L}^p$ process (8).

Proposition 1. *Assume that function G satisfies Conditions 3 and 5 with $p \geq 1$ and sequence $(a_v)_{v\in\mathbb{Z}^d}$ satisfy $\sum_{v\neq 0} a_v < 1$, then $\sum_{v\in\mathbb{Z}^d} \delta_v < \infty$. If moreover assume $a_v = O(|v|^{-d+1-\beta})$ for some $\beta > 1$, then $\delta_v = O(|v|^{-\beta})$.*

For the line transect model (10) with $d = 1$, we have $\delta_{v,p} = O(|v|^{-\beta})$, $\beta > 1$, provided that $a_v = O(|v|^{-\beta})$, which is the same as Theorem 4. For higher dimension $\mathbb{Z}^d$, we require an extra factor $|v|^{1-d}$.

EXAMPLE 11. If the process only contains finite order, then we can have geometric moment contraction ([35]) under milder conditions. To be specific, for $X_t = G_{\epsilon_t}(X_s,\ s \in \Theta),\ t \in \mathbb{Z}^d$, assuming that the index set Θ has only finitely many vectors in $\mathbb{Z}^d$-$\{\mathbf{0}\}$, then $r = \max\{|v|,\ v \in \Theta\}$ is finite and functional dependence measure of X_t decays geometrically, provided that Condition 3 holds with $\rho = \sum_{v\in\Theta} a_v < 1$. Let $K = [|t|/r]$. Then

$$\begin{aligned}
\|X_t - X_{t,0}\|_p &\le \sum_{i_1\in\Theta} a_{i_1}\|X_{t+i_1} - X_{t+i_1,0}\|_p \\
&\le \sum_{i_1,i_2\in\Theta}\sum a_{i_1}a_{i_2}\|X_{t+i_1+i_2} - X_{t+i_1+i_2,0}\|_p \\
&\le \ldots \le \sum_{i_1,\ldots,i_K\in\Theta}\ldots\sum a_{i_1}\ldots a_{i_K}(2\|X_0\|_p) = 2\rho^K\|X_0\|_p.
\end{aligned}$$

Hence we can have geometric moment contraction $\delta_t = O(\rho^{|t|/r})$. In the nonlinear time series setting; (cf (5) and (6)), Shao and Wu [29] obtained a similar result. □

3.1.3. Functional dependence measure for non-stationary simultaneous autoregressive schemes

Let (Θ, d) be a metric space containing countably many indices. By Corollary 5, we can construct $(X_t)_{t\in\Theta}$ satisfying (30). Interestingly, we can obtain similar results as Proposition 1 for such a system. To account for non-stationarity, we define the functional dependence measure

$$\delta^*_{s,v} = \|X_s - X_{s,v}\|_p, \quad \text{and} \quad \delta^*_t = \sup_{\{(s,v)\in\Theta^2 | d(s,v)=t\}} \delta^*_{s,v}. \tag{39}$$

CONDITION 6. *Let $p \ge 1$. For $G_{(t)}$ in (30), assume that for all $t \in \Theta$,*

$$g_t(\epsilon_t, \epsilon'_t) := \sup_{w_v, v\in\Theta_t} |G_{(t)}(\{w_v,\ v\in\Theta_t\};\epsilon_t) - G_{(t)}(\{w_v,\ v\in\Theta_t\};\epsilon'_t)| \in \mathcal{L}^p,$$

and $M_0 := \sup_{t\in\Theta} \|g_t(\epsilon_0, \epsilon'_0)\|_p < \infty$.

COROLLARY 8. *Assume that G satisfies Conditions 4 and 6 with coefficients $a_{s,t}$. Let $\alpha_l = \sup\{a_{s,t} : d(s,t) = l\}$ and assume $\sup_{t\in\Theta}\sum_{s\in\Theta_t}\alpha_{d(t,s)} = \rho_0 < 1$. Then we have*

$$\sup_{t\in\Theta}\sum_{s\in\Theta_t}\delta^*_{t,s} \le M_0/(1-\rho_0). \tag{40}$$

If moreover assume the cardinal number $\sup_{w\in\Theta} \sharp\{s\in\Theta|\ d(s,w) = t\} \le c_1 t^d$ and $\alpha_l \le c_2 l^{-(d+\beta)}$ for some constants c_1, c_2, $d > 0$ and $\beta > 1$, then

$$\delta^*_t = O(t^{-\beta}). \tag{41}$$

If additionally $r := \sup_{t\in\Theta}\sup_{s\in\Theta_t} d(s,t) < \infty$, then $\delta^*_t = O(\phi_0^{t/r})$, where $\phi_0 = \sup_{t\in\Theta}\sum_{s\in\Theta_t} a_{t,s} < 1$.

3.2. Central limit theorem

Theorem 1 in El Machkouri, Volnỳ and Wu [11] asserts that, if the functional dependence measure $\delta_{i,2}$ for the process (9) is summable, namely $\sum_{i\in\mathbb{Z}^d}\delta_{i,2}<\infty$, then for $S_{\Gamma_n}=\sum_{v\in\Gamma_n}(X_v-\mathbb{E}(X_v))$, where $\Gamma_n\subset\mathbb{Z}^d$ satisfies $|\Gamma_n|\to\infty$ and $\sigma_n^2=\mathbb{E}(S_{\Gamma_n}^2)\to\infty$, the Levy distance

$$L\big(S_{\Gamma_n}/\sqrt{|\Gamma_n|},\ \mathcal{N}(0,\sigma_n^2/|\Gamma_n|)\big)\to\infty \qquad \text{as } n\to\infty.$$

Notice that the above CLT holds without specifying any other requirement on Γ_n. The summability of $(\delta_{i,2})$ follows from Theorem 4. If additionally $|\partial\Gamma_n|/|\Gamma_n|$ goes to zero, then

$$S_{\Gamma_n}/\sqrt{|\Gamma_n|}\xrightarrow{\mathcal{L}}\mathcal{N}(0,\sigma^2),\ \text{ where } \sigma^2=\sum_{v\in\mathbb{Z}^d}\operatorname{Cov}(X_0,X_v). \tag{42}$$

For the process (8), by Proposition 1, if Condition 5 holds with $p=2$ and $\sum_{v\neq 0}a_v<1$, then $\sum_{i\in\mathbb{Z}^d}\delta_{i,2}<\infty$ and the above CLT holds.

3.3. Quenched central limit theorem

In certain applications such as MCMC, the process starts at values that do not follow the stationary distribution. This leads to the idea of quenched CLT; see Volnỳ and Woodroofe [32]. For process (1) due to its infinite order, we cannot generate it directly. However we can generate sequence $(X_k^\circ)_{k\geq 1}$ through

$$X_0^\circ=G_{\epsilon_0}(0,0,...),\quad X_k^\circ=G_{\epsilon_k}(X_{k-1}^\circ,\ldots,X_0^\circ,0,\ldots),\ k\geq 1.$$

Theorem 5 provides a CLT for such sequences.

THEOREM 5. *For process (1), assume Condition 1 with $p\geq 2$ and $\sum_{k\geq 1}a_k<1$. Then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n(X_i^\circ-\mathbb{E}(X_i^\circ))\to N(0,\sigma^2),\quad \textit{where}\quad \sigma^2=\sum_{v\in\mathbb{Z}}\operatorname{Cov}(X_0,X_v). \tag{43}$$

Similar assumptions can be applied to the process (16).

COROLLARY 9. *If there exists some $p\geq 2$ such that $f_0(\xi_0), f_i(\xi_1^i), i\geq 1$ are in $\mathcal{L}^p$ and (19) holds, then for process (16) we have (43).*

Proof of Theorem 5. By Theorem 1, there exists a stationary solution in $\mathcal{L}^p$. Let $c_0=\max(\|X_0\|_p,\ \|X_0^\circ\|_p)$. Condition 1 leads to recursive inequality of functional dependence $\|X_n-X_{n,n-k}\|_p\leq\sum_{i=1}^k a_i\|X_{n-i}-X_{n-i,n-k}\|_p$. Similar recursion holds for X_n°. Therefore functional dependence $\|X_0-X_{0,-k}\|_p$ and $(\|X_n^\circ-X_{n,n-k}^\circ\|_p)_{n\geq 0}$ are all bounded by τ_k with

$$\tau_n=\sum_{i=1}^n a_i\tau_{n-i},\quad \tau_0=2c_0.$$

By Theorem 4, $\sum_{k\geq 1}\tau_k < \infty$. Let $D_i = X_i - X_i^\circ - \mathbb{E}(X_i) + \mathbb{E}(X_i^\circ)$ and $d_k = \sup_{i\geq k}\|D_i\|$, then for any $k\in\mathbb{N}$ we have

$$\begin{aligned}\frac{1}{\sqrt{n}}\Big\|\sum_{i=1}^{n} D_i\Big\| &\leq \frac{1}{\sqrt{n}}\sum_{j=0}^{\infty}\Big\|\sum_{i=k}^{n} P_{i-j}D_i\Big\| + \frac{1}{\sqrt{n}}\Big\|\sum_{i=1}^{k-1} D_i\Big\| \\ &\leq \sum_{j=0}^{\infty}\min\{2\tau_j, d_k\} + \frac{k}{\sqrt{n}}d_1. \end{aligned} \tag{44}$$

Since $d_k \to 0$, (44) goes to 0 as $n\to\infty$. Thus it remains to show the CLT for summation of (X_i). From Theorem 4 (i), Condition 1 leads to the summability of functional dependence measure $(\|X_0 - X_{0,-k}\|_p)_{k\geq 0}$. Then by (42) in Subsection 3.2, the CLT holds. The proof for Corollary 9 is similar. □

3.4. An invariance principle

Section 2 provides sufficient conditions for the existence of stationary distributions for process (1) and (16). Based on the functional dependence measures in Section 3.1, we can further derive Gaussian approximation results.

THEOREM 6. *Let (a_k) be either (i) the constants in Condition 1 for process (1) or (ii) $a_k = u_k + v_k$ defined in (37) for process (16), $p>2$. Assume $\sum_{k\geq 1} a_k < 1$ and*

$$\sum_{m=k}^{\infty} a_m = O(k^{1-\beta}),\ \text{where } \beta > \begin{cases} 2 & \text{if } 2<p\leq 4\\ 1+\frac{p^2-4+(p-2)\sqrt{p^2+20p+4}}{8p} & \text{if } p>4. \end{cases} \tag{45}$$

Then there exists a probability space $(\Omega_c, \mathcal{A}_c, \mathcal{P}_c)$ on which we can define a process X_i^c with $S_n^c = \sum_{i=1}^n (X_i^c - \mathbb{E}X_i^c)$, and a standard Wiener process $W_c(.)$, such that $(X_i^c)_{i\in\mathbb{Z}} \overset{\mathcal{D}}{=} (X_i)_{i\in\mathbb{Z}}$ for (1) (resp. $(X_i^c)_{i\in\mathbb{Z}} \overset{\mathcal{D}}{=} (Y_i)_{i\in\mathbb{Z}}$ for process (16)) and

$$S_n^c - \sigma W_c(n) = o_{a.s.}(n^{1/p}),\quad \text{where } \sigma^2 = \sum_{i\in\mathbb{Z}} \operatorname{cov}(X_0, X_i). \tag{46}$$

Proof of Theorem 6. We shall only prove case (i) since (ii) is similar. By Theorem 1, (1) has an $\mathcal{L}^p$ stationary solution (4). Following steps in Section 3, we can get (36). Since $\sum_{m=k}^{\infty} a_m = O(k^{1-\beta})$, by Theorem 4, we have $\sum_{m=k}^{\infty}\delta_{m,p} = O(k^{1-\beta})$. Thus (46) follows from Corollary 2.1 in Berkes, Liu and Wu [2]. □

Aue, Berkes and Horváth [1] obtained an invariance principle for the process

$$Y_t = g(\epsilon_{t-1})Y_{t-1} + c(\epsilon_{t-1}),$$

which is a special case of our (16) and (17). Assuming that $|1/\Lambda'(\Lambda^{-1}(x))| \leq Cx^\gamma$ and $\Lambda(\sigma_0^2) \geq \omega$ hold for some constants $C, \gamma, \omega > 0$ and Y_0 has a finite $\upsilon > 4(1+\gamma)$ moment, they obtained a strong invariance principle for $S_n = \sum_{i=1}^n Y_i$

with rate $o_{a.s.}(n^\theta)$, where $\theta > 5/12$. Our Theorem 6 provides a much sharper rate. Let $p = v/(1+\gamma)$. Then

$$\begin{aligned}\delta_{k,p} &= \|\Lambda^{-1}(Y_t) - \Lambda^{-1}(Y'_{t,t-k})\|_p \|\epsilon_0\|_p \\ &\le C \left\| (|Y_t|^\gamma + |Y_{t,t-k}|^\gamma)|Y_t - Y'_{t,t-k}| \right\|_p \|\epsilon_0\|_p \\ &\le C \| |Y_t|^\gamma + |Y_{t,t-k}|^\gamma \|_{v/\gamma} \|Y_t - Y'_{t,t-k}\|_v \|\epsilon_0\|_p.\end{aligned}$$

Based on the recursion $\|Y_t - Y'_{t,t-k}\|_v = \rho \|Y_{t-1} - Y'_{t-1,t-k}\|_v$, where $\rho = \|c(\epsilon_0)\|_v < 1$, we get $\|Y_t - Y'_{t,t-k}\|_v = O(\rho^k)$. Therefore Theorem 6 leads to a Gaussian approximation with error rate $o_{a.s.}(n^{1/p})$, which is much sharper than their rate $o_{a.s.}(n^\theta)$ with $\theta > 5/12$ since $p > 4$. Aue, Berkes and Horváth [1] applied their invariance principle to a change point detection problem with weighted CUSUM statistics ([6]). It is expected that our sharper strong invariance principle can lead to an improved convergence rate.

4. Proofs

In this section we shall provide proofs of results stated in the previous sections.

Lemma 1. *Under settings in Theorem 4, $\tau_n \le c + B$ for any $n \in \mathbb{N}$.*

Proof. We shall show by induction that $\tau_k \le c + \sum_{i=1}^k b_i$. It trivially holds when $k = 0$. Suppose it holds for any $k \le n$. Then for $k = n+1$, by (33) and $\sum_{k\ge 1} a_k < 1$, we have

$$\tau_{n+1} \le \sum_{k=1}^{n+1} a_k \Big(c + \sum_{i=1}^{n+1-k} b_i\Big) + b_{n+1} \le c + \sum_{i=1}^{n+1} b_i. \qquad \square$$

Proof of Theorem 1. Recall (15) for $X_t^{(t-n)}$. Thus $X_{t+m-k}^{(t)} \in \mathcal{F}_{t+m-k}$ is independent of ϵ_{t+m} for any $k \ge 1$. Condition 1 implies

$$\begin{aligned}&\Big(\mathbb{E}\big(|X_{t+m}^{(t)} - X_{t+m}^{(t-1)}|^p \big| \mathcal{F}_{t+m-1}\big)\Big)^{1/p} \\ &\quad \le \sum_{k=1}^m a_k |X_{t+m-k}^{(t)} - X_{t+m-k}^{(t-1)}| + a_{m+1}|X_{t-1}^{(t-1)}|. \end{aligned} \tag{47}$$

Since $X_{t+1}^{(t)} - X_t^{(t)} = F_{\epsilon_{t+1}}(0,0,\ldots)$ is in $\mathcal{L}^p$ and $\sum_{k\ge1} a_k < 1$, by induction we know for any $m \ge 1$, $X_{t+m}^{(t)} - X_{t+m}^{(t-1)}$ is in $\mathcal{L}^p$ with

$$\|X_{t+m}^{(t)} - X_{t+m}^{(t-1)}\|_p \le \sum_{k=1}^m a_k \|X_{t+m-k}^{(t)} - X_{t+m-k}^{(t-1)}\|_p \le \|F_{\epsilon_0}(0,0,\ldots)\|_p. \tag{48}$$

Notice for any $k \in \mathbb{Z}$ and $t \ge m$, $X_t^{(m)} - X_t^{(m-1)} \overset{\mathcal{D}}{=} X_{t+k}^{(m+k)} - X_{t+k}^{(m+k-1)}$. Thus (48) leads to

$$\|X_t^{(t-m)} - X_t^{(t-m-1)}\|_p \le \sum_{k=1}^m a_k \|X_t^{(t-m+k)} - X_t^{(t-m+k-1)}\|_p. \tag{49}$$

Let $C = \|F_{\epsilon_0}(0, 0, \ldots)\|_p (1 - \sum_{k \geq 1} a_k)^{-1}$. By (49) and Theorem 4 (i),

$$\sum_{k=0}^{\infty} \|X_t^{(t-k)} - X_t^{(t-k-1)}\|_p \leq C < \infty. \tag{50}$$

Consequently $X_t^{(t-m)}$ converges to some random variable Z_t(say) almost surely and in $\mathcal{L}^p$. It remains to verify that Z_t satisfies (1). If $C = 0$, then 0 is a stationary solution. Assume $C > 0$ from now on. For any $\epsilon > 0$, there exists $M \in \mathbb{N}$ such that $\sum_{k > M} a_k < \epsilon/(6C)$. For this fixed M, there exists $N \in \mathbb{N}$ such that for any $n \geq N$, and $l \leq M$, $\|X_{t-l}^{(t-n)} - Z_{t-l}\|_p \leq \epsilon/3$. By (50), $\|Z_t\|_p$ and $\|X_t^{(m)}\|_p$ are bounded by C. Consequently for any $n \geq N$, observe

$$\begin{aligned}
\|Z_t - F_{\epsilon_t}(Z_{t-1}, Z_{t-2}, \ldots)\|_p &\leq \|Z_t - X_t^{(t-n)}\|_p \\
&\quad + \|X_t^{(t-n)} - F_{\epsilon_t}(Z_{t-1}, Z_{t-2}, \ldots)\|_p \\
&\leq \epsilon/3 + \sum_{k=1}^{M} a_k \|Z_{t-k} - X_{t-k}^{(t-n)}\|_p + \sum_{k > M} a_k 2C \leq \epsilon,
\end{aligned}$$

which proves the existence part. For uniqueness, we shall show by contradiction. If there exists another solution X'_t satisfying (12), then $\|X_t - X'_t\|_p \leq \sum_{k=1}^{\infty} a_k \|X_{t-k} - X'_{t-k}\|_p$. Since $\sum_{k \geq 1} a_k < 1$, the latter leads to a contradiction. ☐

Proof of Corollary 1. Define sequence $X_{t+k}^{(t)}$ as in (15). Apply (13) and since $p < 1$ we have

$$\begin{aligned}
\mathbb{E}|X_t^{(t-n)} - X_t^{(t-n-1)}|^p &\leq \mathbb{E}\Big[\mathbb{E}\Big(|X_t^{(t-n)} - X_t^{(t-n-1)}| \big| \mathcal{F}_{t-1}\Big)^p\Big] \\
&\leq \sum_{k=1}^{n} a_k \mathbb{E}|X_{t-k}^{(t-n)} - X_{t-k}^{(t-n-1)}|^p.
\end{aligned}$$

Similar argument as in Theorem 1 yields that $X_t^{(t-n)}$ converges as $n \to \infty$ to some random variable which is the unique adapted $\mathcal{L}^p$ stationary solution. ☐

Proof of Corollary 2. For any $t \in \mathbb{Z}$, define $(X_k^{(t)})_{k \geq t}$ as the sequence generated through

$$X_t^{(t)} = 0, \qquad X_t^{(t-m)} = f_0(\epsilon_{t-1}, \epsilon_{t-2}, \ldots) + \sum_{k=1}^{m} f_k(\epsilon_{t-1}, ..., \epsilon_{t-k}) X_{t-k}^{(t-m)}. \tag{51}$$

Since $X_t^{(t)} - X_t^{(t-1)} = f_0(\epsilon_{t-1}, \epsilon_{t-2}, \ldots)$ is in $\mathcal{L}^p$ and $X_{t-k}^{(t-m)}$ is independent of ϵ_{t-k+i} for any $i \geq 0$, by (51) we derive

$$\|X_t^{(t-m)} - X_t^{(t-m-1)}\|_p \leq \sum_{k=1}^{m} \|f_k(\epsilon_1, ..., \epsilon_k)\|_p \|X_{t-k}^{(t-m)} - X_{t-k}^{(t-m-1)}\|_p. \tag{52}$$

Notice $X_{t-k}^{(t-n)} - X_{t-k}^{(t-n-1)}$ has the same distribution as $X_t^{(t-n+k)} - X_t^{(t-n+k-1)}$. By the argument in Theorem 1, $X_t^{(t-m)}$ converges as $m \to \infty$ almost surely and in $\mathcal{L}^p$ to some random variable which is the unique predictive stationary distribution. □

Proof of Corollary 3. Define sequence $X_{t+k}^{(t)}$ recursively as (51). Thus $X_t^{(t-n)}$ is independent of ϵ_k, $k \geq t$. Using the elementary inequality $(a+b)^p \leq a^p + b^p$, we have

$$\mathbb{E}|X_t^{(t-n)} - X_t^{(t-n-1)}|^p \leq \sum_{k=1}^{n} \mathbb{E}|f_k(\epsilon_{t-1}, ..., \epsilon_{t-k})|^p \mathbb{E}|X_{t-k}^{(t-n)} - X_{t-k}^{(t-n-1)}|^p.$$

Then the result follows from the similar argument as in the proof of Theorem 1. □

Lemma 2. *For nonnegative sequences $(a_k)_{k\geq 1}, (b_k)_{k\geq 1}$, assume $\sum_{k=1}^{\infty} a_k = 1$ and*

$$\int_0^{2\pi} \Big|\frac{\tau_0 + Q_b(e^{i\theta})}{1 - Q_a(e^{i\theta})}\Big|^2 d\theta < \infty, \text{ where } Q_a(s) = \sum_{n=1}^{\infty} a_n s^n. \tag{53}$$

Then for $(\tau_k)_{k\geq 0}$ generated through

$$\tau_n = \sum_{k=1}^{n} a_k \tau_{n-k} + b_n, \quad \tau_0 > 0, \tag{54}$$

we have the square summability $\sum_{k\geq 0} \tau_k^2 < \infty$.

Proof of Lemma 2. By (54), $Q_\tau(s) = \tau_0 + Q_a(s)Q_\tau(s) + Q_b(s)$. Hence

$$Q_\tau(s) = (\tau_0 + Q_b(s))(1 - Q_a(s))^{-1}. \tag{55}$$

Since $\int_0^{2\pi} |Q_\tau(e^{i\theta})|^2 d\theta = 2\pi \sum_{k\geq 0} \tau_k^2$, we finish the proof. □

Proof of Theorem 2. We shall show that $(X_t^{(t-n)})_{n\geq 0}$ is a Cauchy sequence. Write

$$X_t^{(t-n)} - X_t^{(t-n-m)} = \sum_{k=n}^{n+m-1} \Big(X_t^{(t-k)} - \mathbb{E}(X_t^{(t-k-1)}|\mathcal{G}_{t-k+1})\Big)$$
$$+ \sum_{k=n}^{n+m-1} \Big(\mathbb{E}(X_t^{(t-k-1)}|\mathcal{G}_{t-k+1}) - X_t^{(t-k-1)}\Big) =: \mathrm{I}_1(m,n) + \mathrm{I}_2(m,n).$$

By Condition 2, as $n \to \infty$,

$$\sup_{m\geq 1} \|\mathrm{I}_1(m,n)\|_p \leq \sum_{k\geq n} \|X_t^{(t-k)} - \mathbb{E}(X_t^{(t-k-1)}|\mathcal{G}_{t-k+1})\|_p \to 0.$$

For $\mathrm{I}_2(m,n)$, write $H_n(\epsilon_t,\ldots,\epsilon_{t-n+1}) = X_t^{(t-n)}$, where $H_n(\cdot)$ is a measurable function, and let $Y_t^{(t-n)} = H_n(\epsilon_t,\ldots,\epsilon_{t-n+2},\epsilon'_{t-n+1})$. Then

$$\mathbb{E}(X_t^{(t-k-1)}|\mathcal{G}_{t-k+1}) - X_t^{(t-k-1)} = \mathbb{E}(Y_t^{(t-k-1)} - X_t^{(t-k-1)}|\mathcal{G}_{t-k}).$$

By Condition 1, since $\|X_{t-k}^{(t-n)} - Y_{t-k}^{(t-n)}\|_p = \|X_t^{(t-n+k)} - Y_t^{(t-n+k)}\|_p$, for $n \geq 1$ we have

$$\begin{aligned}\|X_t^{(t-n-1)} - Y_t^{(t-n-1)}\|_p &\leq \sum_{k=1}^{n} a_k \|X_{t-k}^{(t-n-1)} - Y_{t-k}^{(t-n-1)}\|_p \\ &= \sum_{k=1}^{n} a_k \|X_t^{(t+k-n-1)} - Y_t^{(t+k-n-1)}\|_p.\end{aligned}$$

Since $\|X_{t-n}^{(t-n-1)} - Y_{t-n}^{(t-n-1)}\|_p \leq 2\|F_{\epsilon_0}(0,0,...)\|_p$ finite and (23) and Lemma 2 yield

$$\sum_{k\geq 1} \|X_t^{(t-k)} - Y_t^{(t-k)}\|_p^2 < \infty. \tag{56}$$

Since $D_k = \mathbb{E}(X_t^{(t-k-1)}|\mathcal{G}_{t-k+1}) - X_t^{(t-k-1)}$, $k \in \mathbb{Z}$, are martingale differences, by Burkholder inequality, $\|\sum_{k\geq n} D_k\|_p^2 \leq (p-1)\sum_{k\geq n}\|D_k\|_p^2$. Thus by (56) and Jensen's inequality,

$$\lim_{n\to\infty}\sup_m \|\mathrm{I}_2(m,n)\|_p^2 \leq \lim_{n\to\infty}(p-1)\sum_{k\geq n}\|Y_t^{(t-k-1)} - X_t^{(t-k-1)}\|_p^2 = 0.$$

Hence as $k\to\infty$, $X_t^{(t-k)}$ converges to a limit $Z_t \in \mathcal{L}^p$ (say). Similar argument as in proof of Theorem 1 shows that Z_t satisfies (1). □

Proof of Corollary 4. Generate $X_t^{(t-n)}$ by (26). Decompose

$$\begin{aligned}X_t^{(t-n)} - X_t^{(t-n-m)} &= \sum_{k=n}^{n+m-1}\left(X_t^{(t-k)} - \mathbb{E}(X_t^{(t-k-1)}|\mathcal{G}_{t-k})\right) \\ &+ \sum_{k=n}^{n+m-1}\left(\mathbb{E}(X_t^{(t-k-1)}|\mathcal{G}_{t-k}) - X_t^{(t-k-1)}\right) =: \mathrm{I}_1(m,n) + \mathrm{I}_2(m,n).\end{aligned}$$

By (27) in Remark 4, $\sup_{m>0}\|\mathrm{I}_1(m,n)\|_p \to 0$ as $n\to\infty$. For $\mathrm{I}_2(m,n)$ we have recursion

$$\begin{aligned}&\|\mathbb{E}(X_t^{(t-n)}|\mathcal{G}_{t-n+1}) - X_t^{(t-n)}\|_p \\ &\quad\leq \sum_{k=1}^{n}\|f_k(\epsilon_{t-1},...,\epsilon_{t-k})\|_p \left\|\mathbb{E}(X_{t-k}^{(t-n)}|\mathcal{G}_{t-n+1}) - X_{t-k}^{(t-n)}\right\|_p + \delta_{n,p}.\end{aligned}$$

By (53), Lemma 2 leads to the square summability

$$\sum_{k\geq 1} \left\|\mathbb{E}(X_0^{(-k)}|\mathcal{G}_{-k+1}) - X_0^{(-k)}\right\|_p^2 < \infty. \tag{57}$$

Based on (57) and apply Burkholder inequality we get

$$\lim_{n\to\infty}\sup_m \|\mathrm{I}_2(m,n)\|_p^2 \leq \lim_{n\to\infty}(p-1)\sum_{k\geq n}\left\|\mathbb{E}(X_t^{(t-k-1)}|\mathcal{G}_{t-k}) - X_t^{(t-k-1)}\right\|_p^2 = 0.$$

Thus $X_t^{(t-n)}$ converges to some limit Z_t (say). Since $\sum_{k\geq 1}\delta_{k,p}^2 < \infty$, apply Burkholder inequality and similar argument as in Theorem 1, we can verify that Z_t satisfy (16). □

Proof of Theorem 3 and Corollary 5. We only deal with $d = 1$ since the case $d \geq 2$ can be similarly handled. Define the sequence $(X_t^{\{k\}})_{k\geq 0}$ recursively as

$$X_t^{\{0\}} = 0 \text{ and } X_t^{\{k+1\}} = G_{\epsilon_t}(\ldots, X_{t-2}^{\{k\}}, X_{t-1}^{\{k\}}, X_{t+1}^{\{k\}}, X_{t+2}^{\{k\}}, \ldots),\ k = 0, 1, \ldots. \tag{58}$$

By Condition 3, denote $\rho = \sum_{k\neq 0} a_k$ and $C_0 = \|G_{\epsilon_0}(0, 0, \ldots)\|_p$, we have

$$\begin{aligned}\|X_t^{\{k+1\}} - X_t^{\{k\}}\|_p &\leq \Big\| \sum_{i_1\neq 0} a_{i_1} |X_{t+i_1}^{\{k\}} - X_{t+i_1}^{\{k-1\}}| \Big\|_p \\ &\leq \Big\| \sum_{i_1\neq 0} a_{i_1} \sum_{i_2\neq 0} a_{i_2} |X_{t+i_1+i_2}^{\{k-1\}} - X_{t+i_1+i_2}^{\{k-2\}}| \Big\|_p \leq \ldots \\ &\leq \Big(\sum_{i\neq 0} a_i\Big)^k \sup_{t\in\mathbb{Z}} \|X_t^{\{1\}} - X_t^{\{0\}}\|_p = \rho^k C_0. \end{aligned} \tag{59}$$

Since $\rho < 1$, $X_t^{\{k+1\}}$ converges almost surely and in $\mathcal{L}^p$ to some random variable denoted as Z_t and $\|Z_t - X_t^{\{k\}}\|_p \leq \rho^k C_0/(1-\rho)$. Clearly Z_t satisfies (10) by letting $k \to \infty$ in

$$\begin{aligned}&\|Z_t - G_{\epsilon_t}(\ldots, Z_{t-2}, Z_{t-1}, Z_{t+1}, Z_{t+2}\ldots)\|_p \\ &\leq \|Z_t - X_t^{\{k+1\}}\|_p + \|G_{\epsilon_t}(\ldots, X_{t-1}^{\{k\}}, X_{t+1}^{\{k\}}\ldots) - G_{\epsilon_t}(\ldots, Z_{t-1}, Z_{t+1}, \ldots)\|_p \\ &\leq \|Z_t - X_t^{\{k+1\}}\|_p + \sum_{i\neq 0} a_i \sup_{i\in\mathbb{Z}} \|Z_i - X_i^{\{k\}}\|_p \leq 2\rho^k C_0/(1-\rho).\end{aligned}$$

To show the uniqueness, if there exists another process $Y_t \in \mathcal{L}^p$ also satisfies (10), then

$$\|X_0 - Y_0\|_p \leq \sum_{i\neq 0} a_i \|X_i - Y_i\|_p = \rho\|X_0 - Y_0\|_p.$$

Since $\rho < 1$, $X_t = Y_t$ almost surely. The non-stationary case can be similarly dealt with. □

Proof of Theorem 4(i). For any $N \in \mathbb{N}$, by (33), we get

$$\sum_{n=0}^{N} \tau_n = \tau_0 + \sum_{n=1}^{N}\sum_{k=1}^{n} u_k \tau_{n-k} + \sum_{n=0}^{N} v_n \leq (c+B) + (\sum_{k=1}^{N} u_k)(\sum_{n=0}^{N} \tau_n).$$

Thus $\sum_{k=0}^{N} \tau_k \leq (c+B)/\delta$ and therefore $T := \sum_{k=0}^{\infty} \tau_k$ is finite. Applying (33) again and letting $N \to \infty$, we have $T = \tau_0 + B + AT$, implying $T = (c+B)/\delta$.

Proof of (ii). Let $\Delta = 1 - (1+\delta/4)^{-1/\theta}$. Write

$$\tau_n = \sum_{k=1}^{\lfloor \Delta n \rfloor} u_k \tau_{n-k} + \sum_{k=\lfloor \Delta n \rfloor+1}^{n-\lfloor \Delta n \rfloor -1} u_k \tau_{n-k} + \left(\sum_{k=n-\lfloor \Delta n \rfloor}^{n} u_k \tau_{n-k} + v_n \right) =: \mathscr{A}_n + \mathscr{B}_n + \mathscr{C}_n. \tag{60}$$

where $\lfloor x \rfloor$ means biggest integer that is no larger than x. Let $C_0 = \sup_{n\geq 1}(u_n + v_n)n^{\theta}$. Take N_0 big enough such that $C_0 2^{\theta}(\theta-1)^{-1}(N_0\Delta)^{-(\theta-1)} \leq \delta/4$ and let $M = \max\{4C_0(c+B)/\delta^2, (c+B)N_0^{\theta}, 4C_0/\delta\}$. Next we shall apply induction and show that

$$\tau_n \leq M n^{-\theta} \text{ holds for all } n \in \mathbb{N}. \tag{61}$$

By Lemma 1 we have $\tau_n \leq c+B$, and thus $\tau_k \leq Mk^{-\theta}$ for any $k \leq N_0$. Suppose for any $k \leq n-1$, τ_k satisfies (61). Then for $k = n$ we have by the induction hypothesis that

$$\mathscr{A}_n \leq M(n - \lfloor \Delta n \rfloor)^{-\theta} \sum_{k=1}^{\lfloor \Delta n \rfloor} u_k \leq M n^{-\theta}(1-\Delta)^{-\theta}(1-\delta).$$

For the second part, applying Jensen's inequality we derive

$$\begin{aligned}\mathscr{B}_n &\leq C_0 M \sum_{k=\lfloor \Delta n \rfloor+1}^{n-\lfloor \Delta n \rfloor -1} k^{-\theta}(n-k)^{-\theta} = C_0 M n^{-\theta} \sum_{k=\lfloor \Delta n \rfloor+1}^{n-\lfloor \Delta n \rfloor -1} \big(1/k + 1/(n-k)\big)^{\theta} \\ &\leq C_0 M n^{-\theta} 2^{\theta} \sum_{k=\lfloor \Delta n \rfloor+1}^{\infty} k^{-\theta} \leq M n^{-\theta} C_0 2^{\theta}(\theta-1)^{-1} \lfloor \Delta n \rfloor^{-(\theta-1)}.\end{aligned}$$

For the last part, according to part(i), $\sum_{k=0}^{\infty} \tau_k = (c+B)/\delta$ we get

$$\mathscr{C}_n \leq C_0(n - \lfloor \Delta n \rfloor)^{-\theta} \sum_{k=0}^{\lfloor \Delta n \rfloor} \tau_k + Cn^{-\theta} \leq M n^{-\theta} C_0 \Big(\frac{c+B}{(1-\Delta)^{\theta}\delta M} + \frac{1}{M} \Big).$$

Then the induction step is completed in view of

$$\mathscr{A}_n + \mathscr{B}_n + \mathscr{C}_n \leq M n^{-\theta}[(1-\delta)(1+\delta/4) + \delta/4 + \delta/4(1+\delta/4) + \delta/4] \leq M n^{-\theta}.$$

Proof of (iii). By (ii), the upper bound follows. By (33), we get $\tau_n \geq u_n \tau_0 + v_n \geq \min\{1, c\}(u_n + v_n)$, implying the lower bound.

Proof of (iv). Let $c_0 = \max\{(B+c)/(1-A), 1\}$ and $\Gamma_n = \sum_{k=n}^{\infty} \tau_k$. By (33) we have

$$\begin{aligned}\Gamma_n &= \sum_{k=n}^{\infty}\sum_{i=1}^{k} u_i \tau_{k-i} + \sum_{k=n}^{\infty} v_k \\ &= \sum_{i=1}^{n} u_i \Big(\sum_{k=n}^{\infty} \tau_{k-i}\Big) + \sum_{i>n} u_i \Big(\sum_{k=0}^{\infty} \tau_k\Big) + \sum_{k=n}^{\infty} v_k \\ &\le \sum_{i=1}^{n} u_i \Gamma_{n-i} + c_0 \sum_{i=n}^{\infty} (u_i + v_i).\end{aligned}$$

Choose $\Delta > 0$ such that $(1-\Delta)^{-(\theta-1)} = 1 + \delta/4$. Let $C_1 = \sup_{n\ge 1} \sum_{k\ge n}(u_k + v_k)n^{\theta-1}$. Take N_0 big enough such that $C_1 \Delta^{-2(\theta-1)} N_0^{-(\theta-1)} \le \delta/4$ and let $M = \max\{c_0 N_0^{(\theta-1)}, 4c_0(1-\Delta)^{-(\theta-1)}/\delta, 4C_1 c_0/\delta\}$. Similarly as in the proof of (ii) we show by induction that for any $n \in \mathbb{N}$, $\Gamma_n \le M n^{-(\theta-1)}$. By part(i) we have $\Gamma_n \le c_0$, and thus $\Gamma_k \le M k^{-(\theta-1)}$ for any $k \le N_0$. If for any $k \le n-1$, Γ_k satisfies $\Gamma_k \le M k^{-(\theta-1)}$. Then for $k = n$ we have

$$\begin{aligned}&\sum_{i=1}^{\lfloor \Delta n\rfloor} u_i \Gamma_{n-i} + \sum_{i=\lfloor \Delta n\rfloor+1}^{n-\lfloor \Delta n\rfloor-1} u_i \Gamma_{n-i} + \sum_{i=n-\lfloor \Delta n\rfloor}^{n} u_i \Gamma_{n-i} + c_0 \sum_{i\ge n}(u_i + v_i) \\ \le& (1-\delta) M (1-\Delta)^{-(\theta-1)} n^{-(\theta-1)} + MC\Delta^{-2(\theta-1)} n^{-2(\theta-1)} \\ &+ (1-\Delta)^{-(\theta-1)} n^{-(\theta-1)} C c_0 + C c_0 n^{-(\theta-1)} \\ \le& M n^{-(\theta-1)}[(1-\delta)(1+\delta/4) + \delta/4 + \delta/4 + \delta/4] \le M n^{-(\theta-1)},\end{aligned}$$

completing the induction step.

Proof of (v). Let $U = \limsup_{n\to\infty} \tau_n n^{\theta}$ and $L = \liminf_{n\to\infty} \tau_n n^{\theta}$. Without loss of generality set $c = 1$. In I) and II) below, we shall show that, respectively $U \le C/\delta^2$ and $L \ge C/\delta^2$. Hence $U = L = C/\delta^2$ and $\tau_n \sim (C/\delta^2) n^{-\theta}$.

I). We shall show $U \le C/\delta^2$ by contradiction. If not, set $\eta = U - C/\delta^2 > 0$ and $\Delta_0 = \min\{1, \eta\delta/10\}$. Part (ii) leads to $\tau_n = O(n^{-\theta})$, thus U is finite. Consequently for $\Delta_a = \min\{1, \Delta_0 C/(U\delta)\}$, there exists $N_0 \in \mathbb{N}$, s.t. for all $n \ge N_0$, $\tau_n n^{\theta} \le U + \Delta_0$ and $u_n n^{\theta} \le C + \Delta_a$.

Choose $\Delta > 0$ such that $(1-\Delta)^{-\theta} = \min\{1 + \Delta_0/(1+U), 1 + \delta\Delta_0/(1+C)\}$. Let $N_1 \in \mathbb{N}$ be such that $N_1^{-(\theta-1)} \le \Delta_0(\theta-1)2^{-\theta}(U+1)^{-1}(C+1)^{-1}$. Write

$$\tau_n n^{\theta} = \left(\sum_{k=1}^{\lfloor \Delta n\rfloor} + \sum_{k=\lfloor \Delta n\rfloor+1}^{n-\lfloor \Delta n\rfloor-1} + \sum_{k=n-\lfloor \Delta n\rfloor}^{n}\right) u_k \tau_{n-k} n^{\theta} =: \mathscr{A}_n + \mathscr{B}_n + \mathscr{C}_n.$$

Set $N = \max\{\lfloor N_0/\Delta\rfloor, \lfloor N_1/\Delta\rfloor\}$. When $n \ge N$, $\tau_k n^{\theta} = \tau_k k^{\theta}(n/k)^{\theta} \le (U + \Delta_0)(1-\Delta)^{-\theta}$ for any $k \ge (1-\Delta)n$,

$$\mathscr{A}_n \le \Big(\sum_{k=1}^{\lfloor \Delta n\rfloor} u_k\Big)(U+\Delta_0)(1-\Delta)^{-\theta} \le (1-\delta)(U+\Delta_0)(1-\Delta)^{-\theta} \le (1-\delta)U + 2\Delta_0.$$

For the second part, using $u_k \le (C+\Delta_a)k^{-\theta}$ and $\tau_k \le (U+\Delta_0)k^{-\theta}$, we derive

$$\begin{aligned}\mathscr{B}_n &\le (U+\Delta_0)(C+\Delta_a)\sum_{k=\lfloor \Delta n\rfloor+1}^{n-\lfloor \Delta n\rfloor-1} k^{-\theta}(n-k)^{-\theta}n^{\theta}\\ &= (U+\Delta_0)(C+\Delta_a)\sum_{k=\lfloor \Delta n\rfloor+1}^{n-\lfloor \Delta n\rfloor-1}(1/k+1/(n-k))^{\theta} \le \Delta_0.\end{aligned}$$

For the last part, by part(i) that $\sum_{k=1}^{\infty}\tau_k = 1/\delta$ and $u_k n^{\theta} = u_k k^{\theta}(n/k)^{\theta} \le (C+\Delta_a)(1-\Delta)^{-\theta}$ for any $k > n-\lfloor \Delta n\rfloor$, we get

$$\mathscr{C}_n \le (C+\Delta_a)(1-\Delta)^{-\theta}\sum_{k=0}^{\lfloor \Delta n\rfloor}\tau_k \le (C+\Delta_a)(1-\Delta)^{-\theta}/\delta \le (C+\Delta_a)/\delta+\Delta_0.$$

Recall that $\eta = U - C/\delta^2$. Hence

$$\mathscr{C}_n \le (C+\Delta_a)\delta C^{-1}(U-\eta)+\Delta_0 \le U\delta-\eta\delta+2\Delta_0.$$

Since $\eta\delta \ge 10\Delta_0$, combining above results together we conclude

$$\tau_n n^{\theta} = \mathscr{A}_n+\mathscr{B}_n+\mathscr{C}_n \le (1-\delta)U+U\delta-\eta\delta+5\Delta_0 \le U-\eta\delta/2.$$

Since this holds for all $n \ge N$, $U = \limsup_{n\to\infty}\tau_n n^{\theta} \le U-\eta\delta/2 < U$. The latter leads to a contradiction.

II) Similarly, we need to prove $L \ge C/\delta^2$. If not, let $\eta = C/\delta^2 - L > 0$. Take $\Delta_0 = \delta\eta/8$ and $\Delta_a = \Delta_0\delta$. Since $\tau_n n^{\theta} \ge a_n n^{\theta}\tau_0 \sim C\tau_0 > 0$, thus L is strictly larger than 0.

Since $L = \liminf_{n\to\infty}\tau_n n^{\theta}$ and $u_n \sim Cn^{-\theta}$, there exists $N_0 \in \mathbb{N}$, s.t. for all $n \ge N_0$, $\tau_n n^{\theta} \ge L-\Delta_0$ and $a_n n^{\theta} \ge C-\Delta_a$. Since $\sum_{k\ge 0}a_k = 1-\delta$ and $\sum_{k\ge 0}\tau_k = 1/\delta < \infty$, there exists $N_1 \in \mathbb{N}$, such that $\sum_{k=0}^{n}u_k \ge 1-\delta-\Delta_0/L$ and $\sum_{k=0}^{n}\tau_k > 1/\delta-\Delta_0/C$ for all $n \ge N_1$. Take some $0 < \Delta < 0.5$(say $\Delta = 0.1$) and decompose $\tau_n n^{\theta}$ as in case I). Let $N = \max\{\lfloor N_0/\Delta\rfloor, \lfloor N_1/\Delta\rfloor\}$. Since $\tau_k n^{\theta} = \tau_k k^{\theta}(n/k)^{\theta} \ge L-\Delta_0$ when $k \ge (1-\Delta)n$, for any $n \ge N$ we have

$$\mathscr{A}_n \ge \Big(\sum_{k=1}^{\lfloor \Delta n\rfloor}u_k\Big)(L-\Delta_0) \ge (1-\delta)L-2\Delta_0,$$

For the second part, notice that every term is nonnegative, therefore we have $\mathscr{B}_n \ge 0$.

For the last part, since $u_k n^{\theta} = u_k k^{\theta}(n/k)^{\theta} \ge C-\Delta_a$, for $k \ge (1-\Delta)n$, we know

$$\mathscr{C}_n \ge (C-\Delta_a)\sum_{k=0}^{\lfloor \Delta n\rfloor}\tau_k \ge (C-\Delta_a)(1/\delta-\Delta_0/C) \ge \delta(L+\eta)-2\Delta_0.$$

Together we have

$$\mathscr{A}_n + \mathscr{B}_n + \mathscr{C}_n \geq (1-\delta)L + \delta(L+\eta) - 4\Delta_0 \geq L + \delta\eta/2 > L.$$

Since this holds for all $n \geq N$, $L = \liminf_{n\to\infty} \tau_n n^\theta \geq L + \delta\eta/2 > L$. Contradiction. □

LEMMA 3. *Let $\delta = 1 - \sum_{v\in\mathbb{Z}^d} a_v$ and $a_0 = 0$. Recall that $\tau_v^{\{k\}}$ is defined in (38). Under conditions in Proposition 1, we have*

$$S^{\{k\}} = \sum_{v\in\mathbb{Z}^d} \tau_v^{\{k\}} \leq 1/\delta. \tag{62}$$

If we further assume $a_v = O(|v|^{-\beta-d+1})$ for some $\beta > 1$, then

$$\max_{k\in\mathbb{N}} \tau_v^{\{k\}} = O(|v|^{-\beta}). \tag{63}$$

Proof of Lemma 3. Summing (38) over $t \in \mathbb{Z}^d$, we have (62) by induction in view of

$$S^{\{k+1\}} \leq 1 + \sum_{v\in\mathbb{Z}^d}\sum_{s\neq\mathbf{0}} a_s \tau_{v+s}^{\{k\}} \leq 1 + \sum_{s\neq\mathbf{0}} a_s S^{\{k\}} = 1 + (1-\delta)S^{\{k\}} \tag{64}$$

and $S^{\{0\}} = 0$. To show (63), choose ϵ such that $(1-\epsilon)^{-\beta} = 1 + \delta/4$. Let $C = \sup_v a_v(1 \vee |v|)^\beta$ and $M = \max\{2C/\delta^2, C_0^\beta/\delta\}$, where $C_0 = (4C_1/\delta)^{1/(\beta-1)}$ and

$$C_1 = \frac{C2^{2\beta-1}(1-\epsilon)^{d-1}\pi^{d/2}d}{\epsilon^{\beta+d-2}\Gamma(d/2+1)(\beta-1)}.$$

We will show by induction with respect to k that

$$\tau_t^{\{k\}} \leq M|t|^{-\beta}, \text{ for any } t \in \mathbb{Z}^{\mathrm{d}}. \tag{65}$$

When $k = 0$, (65) trivially holds since $\tau_t^{\{0\}} = 0$. Suppose it holds for any $k \leq m$, $m \in \mathbb{N}$. By induction it suffices to verify it for $k = m+1$. Note that for any $|t| \leq C_0$ and any $k \in \mathbb{N}$, $\tau_t^{\{k\}} \leq M|t|^{-\beta}$, since $\tau_t^{\{k\}} \leq S^{\{k\}} \leq 1/\delta$. We now deal with t with $|t| > C_0$. By (38), we have

$$\tau_t^{\{m+1\}} = \left(\sum_{|t+v|\geq(1-\epsilon)|t|} + \sum_{|t+v|\leq\epsilon|t|} + \sum_{\epsilon|t|<|t+v|<(1-\epsilon)|t|}\right) a_v \tau_{t+v}^{\{m\}} =: \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

Since $\sum_v a_v = 1 - \delta$, we have

$$\mathrm{I} \leq M((1-\epsilon)|t|)^{-\beta} \sum_{|t+v|\geq\epsilon|t|} a_v \leq (1-\epsilon)^{-\beta}(1-\delta)M|t|^{-\beta}.$$

In part II we have $|v| \geq (1-\epsilon)|t|$. Hence

$$\mathrm{II} \leq C((1-\epsilon)|t|)^{-\beta} \sum_{|t+v|\leq\epsilon|t|} \tau_{t+v}^{\{m\}} \leq C(1-\epsilon)^{-\beta}|t|^{-\beta}1/\delta.$$

Since $\beta > 1$, using $((x+y)/2)^\beta \le (x^\beta + y^\beta)/2$, we have

$$\begin{aligned}
\mathrm{III} &\le MC \sum_{\epsilon|t| \le |t+v| \le (1-\epsilon)|t|} |v|^{-d+1-\beta} |t+v|^{-\beta} \\
&\le MC \sum_{\epsilon|t| \le |t+v| \le (1-\epsilon)|t|} (\epsilon|t|)^{-(d-1)} |t|^{-\beta} \Big(\frac{1}{|t+v|} + \frac{1}{(|t| - |t+v|)} \Big)^\beta \\
&\le MC 2^\beta \frac{\epsilon^{1-d}}{|t|^{d-1+\beta}} \sum_{\epsilon|t| \le |t+v| \le (1-\epsilon)|t|} |t+v|^{-\beta} \le \frac{MC_1}{|t|^{2\beta-1}}
\end{aligned}$$

Combining all three parts, we derive for $|t| \ge C_0$ that

$$\tau_t^{\{m+1\}} \le ((1-\delta)(1+\delta/4) + (1+\delta/4)\delta/2 + \delta/4)\, M|t|^{-\beta} \le M|t|^{-\beta},$$

which completes the proof. □

Proof of Proposition 1. Note that $\tau_t^{\{k\}} \ge \delta_t^{\{k\}}$. Hence by Lemma 3, $\max_{k \in \mathbb{N}} \delta_t^{\{k\}} = O(|t|^{-\beta})$. Since $\lim_{k\to\infty} \|X_t^{\{k\}} - X_{t,0}^{\{k\}}\|_p = \|X_t - X_{t,0}\|_p$, we have $\delta_t \le \limsup_{k\to\infty} \delta_t^{\{k\}}$, implying that $\delta_t = O(|t|^{-\beta})$. □

Proof of Corollary 8. Let $X_t^{\{0\}} = 0$ for all $t \in \Theta$ and generate the processes $(X_t^{\{k\}})_{t\in\Theta}$ by

$$X_t^{\{k+1\}} = G_{(t)}(X_v^{\{k\}}, v \in \Theta_t; \epsilon_t),\ k = 0, 1, \ldots.$$

Then for any $v \in \Theta$ we have $\delta_{t,v}^{*\{0\}} = 0$ and

$$\delta_{t,v}^{*\{k+1\}} \le \begin{cases} \sum_{s\in\Theta_t} \alpha_{d(t,s)} \delta_{s,v}^{*\{k\}} & t \ne v \\ M_0 + \sum_{s\in\Theta_t} \alpha_{d(t,s)} \delta_{s,v}^{*\{k\}} & t = v. \end{cases}$$

Since $\sum_{s\in\Theta_t} \alpha_{d(t,s)} \le \rho_0 < 1$, similarly as (64) we have by induction that $\sum_{t\in\Theta} \delta_{t,v}^{*\{k\}} \le M_0/(1-\rho_0)$, for any $k \in \mathbb{N}$, $v \in \Theta$. Thus (40) follows. Similar argument as in the proof of Theorem 3 implies that for any $k \in \mathbb{N}$, $\delta_{t,v}^{*\{k\}} \le M d(t,v)^{-\beta}$, where M does not depend on t, v, which proves (41). The last statement can be derived similarly as Remark 11. □

References

[1] Aue, A., Berkes, I. and Horváth, L. (2006). Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli* **12** 583–608. MR2248229

[2] Berkes, I., Liu, W. and Wu, W. B. (2014). Komlós–Major–Tusnády approximation under dependence. *The Annals of Probability* **42** 794–817. MR3178474

[3] Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)* 192–236. MR0373208

[4] Brunsdon, C., Fotheringham, A. S. and Charlton, M. (1998). Spatial nonstationarity and autoregressive models. *Environment and Planning A* **30** 957–973.

[5] Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* **18** 17–39. MR1885348

[6] Csörgö, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis* **18**. John Wiley & Sons Inc. MR2743035

[7] Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM Review* **41** 45–76. MR1669737

[8] Ding, Z., Granger, C. W. and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1** 83–106.

[9] Douc, R., Roueff, F. and Soulier, P. (2008). On the existence of some processes. *Stochastic Processes and Their Applications* **118** 755–761. MR2411519

[10] Duan, J.-C. (1997). Augmented GARCH (p, q) process and its diffusion limit. *Journal of Econometrics* **79** 97–127. MR1457699

[11] El Machkouri, M., Volný, D. and Wu, W. B. (2013). A central limit theorem for stationary random fields. *Stochastic Processes and Their Applications* **123** 1–14. MR2988107

[12] Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer Science & Business Media. MR1964455

[13] Giraitis, L., Kokoszka, P. and Leipus, R. (2000). Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory* **16** 3–22. MR1749017

[14] Giraitis, L., Leipus, R. and Surgailis, D. (2007). Recent advances in ARCH modelling. In *Long Memory in Economics* 3–38. Springer. MR2265054

[15] Giraitis, L., Leipus, R. and Surgailis, D. (2009). ARCH () models and long memory properties. In *Handbook of Financial Time Series* 71–84. Springer.

[16] Granger, C. W. J. and Andersen, A. P. (1978). *An Introduction to Bilinear Time Series Models*. Vandenhoeck & Ruprecht. MR0483231

[17] Hörmann, S. (2008). Augmented GARCH sequences: dependence structure and asymptotics. *Bernoulli* 543–561. MR2544101

[18] Jarner, S. and Tweedie, R. (2001). Locally contracting iterated functions and stability of Markov chains. *Journal of Applied Probability* 494–507. MR1834756

[19] Jenish, N. and Prucha, I. R. (2009). Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics* **150** 86–98. MR2525996

[20] Jenish, N. and Prucha, I. R. (2012). On spatial processes and asymptotic

inference under near-epoch dependence. *Journal of Econometrics* **170** 178–190. MR2955948

[21] Kazakevicius, V. and Leipus, R. (2002). On stationarity in the arch ([infty infinity]) model. *Econometric Theory* **18** 1–16. MR1885347

[22] Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent RV'-s, and the sample DF. I. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **32** 111–131. MR0375412

[23] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica: Journal of the Econometric Society* 347–370. MR1097532

[24] Paulik, M. J., Das, M. and Loh, N. (1992). Nonstationary autoregressive modeling of object contours. *Signal Processing, IEEE Transactions on* **40** 660–675.

[25] Priestley, M. B. (1988). *Nonlinear and Nonstationary Time Series Analysis*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London. MR991969 MR0991969

[26] Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms* **9** 223–252. MR1611693

[27] Rao, T. S. and Gabr, M. (2012). *An Introduction to Bispectral Analysis and Bilinear Time Series Models* **24**. Springer Science & Business Media. MR0757536

[28] Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics* **47** 67–84. MR1087207

[29] Shao, X. and Wu, W. B. (2007). Asymptotic spectral theory for nonlinear time series. *The Annals of Statistics* **35** 1773–1801. MR2351105

[30] Tjøstheim, D. (1994). Non-linear time series: a selective review. *Scandinavian Journal of Statistics* 97–130. MR1294588

[31] Tong, H. (1990). *Non-Linear Time Series: a Dynamical System Approach*. Oxford University Press. MR1079320

[32] Volný, D. and Woodroofe, M. (2014). Quenched central limit theorems for sums of stationary processes. *Statistics & Probability Letters* **85** 161–167. MR3157895

[33] Whittle, P. (1954). On stationary processes in the plane. *Biometrika* 434–449. MR0067450

[34] Wu, W. B. (2005). Nonlinear system theory: another look at dependence. *Proceedings of the National Academy of Sciences of the United States of America* **102** 14150–14154. MR2172215

[35] Wu, W. B. and Shao, X. (2004). Limit theorems for iterated random functions. *Journal of Applied Probability* 425–436. MR2052582