

# LIMIT THEOREMS FOR AGGREGATED LINEAR PROCESSES

M. JIRAK,\* *Graz University of Technology*

## Abstract

In this paper we develop an asymptotic theory of aggregated linear processes, and determine in particular the limit distribution of a large class of linear and nonlinear functionals of such processes. Given a sample  $\{Y_1^{(N)}, \dots, Y_n^{(N)}\}$  of the normalized  $N$ -fold aggregated process, we describe the limiting behavior of statistics  $T_{N,n} = T_{N,n}(Y_1^{(N)}, \dots, Y_n^{(N)})$  in both of the cases  $n/N(n) \rightarrow 0$  and  $N(n)/n \rightarrow 0$ , assuming either a ‘limiting long- or short-memory’ condition on the underlying linear process.

**Keywords:** Aggregation; short memory; long memory; invariance principle; random coefficient MA( $\infty$ ); martingale decomposition

2010 Mathematics Subject Classification: Primary 60F17  
Secondary 60G99

## 1. Introduction

Let  $\{X_k, k \in \mathbb{Z}\}$  be a stationary process with mean 0 and finite variances, let  $\{X_k^{(j)}, k \in \mathbb{Z}\}$ ,  $j = 1, 2, \dots$ , be independent copies of  $\{X_k, k \in \mathbb{Z}\}$ , and consider the process

$$\bar{Y}_k^{(N)} = \frac{X_k^{(1)} + \dots + X_k^{(N)}}{N}, \quad k \in \mathbb{Z}. \quad (1.1)$$

Such processes, called *aggregated* processes, appear in many problems of statistics, natural sciences, and economics. For example, macroeconomical data obtained by averaging production data from individual companies are normalized aggregated data. The fundamental statistical problem of aggregation theory is to draw, given a sample  $\{\bar{Y}_1^{(N)}, \dots, \bar{Y}_n^{(N)}\}$  of the aggregated process for some large  $n$  and  $N$ , conclusions for the structure of the underlying processes  $\{X_k^{(j)}, k \in \mathbb{Z}\}$  (‘disaggregation’) and use these to describe the asymptotic behavior of  $\{\bar{Y}_k^{(N)}, k \in \mathbb{Z}\}$ . Besides its obvious importance for econometrics, aggregation has a crucial importance from a purely probabilistic point of view as well. By a seminal paper of Granger [22], aggregating  $N$  random coefficient (autoregressive) AR(1) processes can lead, as  $N \rightarrow \infty$ , to fractionally integrated long-memory processes. Similar examples show that aggregated processes can have a much richer structure and more complicated dynamics than their elementary (‘micro-level’) components. This opens a new way to analyze complex processes by constructing such processes from simple ‘building blocks’ via aggregation. In particular, aggregated random parameter AR processes have been studied extensively in the literature; see the important contributions by Celov *et al.* [10], Dacunha-Castelle and Fermín [13], Dacunha-Castelle and Oppenheim [14], Horváth and Leipus [24], Oppenheim and Viano [36], Puplinskaite and Surgailis [41], [42], Zaffaroni [51], and the references therein. Aggregation or

Received 6 October 2011; revision received 18 July 2012.

\* Current address: Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany.

superpositions in continuous time were discussed in [1], [2], [3], [6], [7], and [32]. For aggregating ARCH-type processes, see [17], [20], [28], [30], [52], and [53]. A somewhat related field are so-called superlinear processes (cf. [47] and the references therein), where processes are instantaneously aggregated by nonrandom weights. Such processes have served as a rich source of examples and counterexamples for the central limit theorem and weak invariance principle for stationary processes. However, the substantial difference between random and nonrandom weights means that different treatments are required.

Let

$$\{Y_k^{(N)}, k \in \mathbb{Z}\} := \{\sqrt{N}\bar{Y}_k^{(N)}, k \in \mathbb{Z}\}.$$

From the multivariate central limit theorem, it follows that the finite-dimensional distributions of  $\{Y_k^{(N)}, k \in \mathbb{Z}\}$  converge as  $N \rightarrow \infty$  to those of the associated Gaussian process  $\{\xi_k, k \in \mathbb{Z}\}$ , whenever some model assumption on  $\{Y_k^{(N)}, k \in \mathbb{Z}\}$  is made. In its simplest analytic form, the disaggregation problem calls for determining the distribution of the process  $\{X_k, k \in \mathbb{Z}\}$  if we know the distribution (e.g. the covariances  $\phi_k$  or spectral density) of the associated process. In case of aggregated AR(1) processes with random parameter  $a$ , a simple explicit solution can be given, but, for aggregated AR( $p$ ) processes with  $p$  random parameters  $a_1, \dots, a_p$ ,  $p \geq 2$  (which are usually dependent due to stationarity and causality assumptions), this is a much harder problem, leading to a Fredholm-type integral equation, which is known to be an ill-posed problem. Equivalently, the moments  $\mathbb{E}[a_j^l]$ ,  $1 \leq j \leq p$ ,  $l = 1, 2, \dots$ , satisfy an infinite nonlinear system of equations, whose structure is so complicated that even numerical solutions are difficult to obtain. In the case of  $p = 1$ , replacing  $\phi_k$  with their estimators  $\hat{\phi}_{k,n,N}$  based on the sample  $\{Y_1^{(N)}, \dots, Y_n^{(N)}\}$  leads to consistent estimators of the moments  $\mathbb{E}[a^l]$  in terms of  $\{Y_1^{(N)}, \dots, Y_n^{(N)}\}$ . For example, in the random coefficient AR(1) case Horváth and Leipus [24] gave the estimator

$$Z_{n,N} = \frac{\sum_{k=2}^n Y_k^{(N)} Y_{k-1}^{(N)} - \sum_{k=4}^n Y_k^{(N)} Y_{k-3}^{(N)}}{\sum_{k=1}^n (Y_k^{(N)})^2 - \sum_{k=3}^n Y_k^{(N)} Y_{k-2}^{(N)}} \quad (1.2)$$

for  $\mathbb{E}[a]$ ; consistent estimators for the density of  $a$  were given in [11] and [31], and, for parametric moment estimators, see [9]. Note that the least square estimator is not consistent in the case of aggregated processes, due to the nonergodicity of  $\{Y_k^{(N)}, k \in \mathbb{Z}\}$ . Determining the asymptotic distribution of such estimators requires studying nonlinear functionals of the sample  $\{Y_1^{(N)}, \dots, Y_n^{(N)}\}$ , a delicate problem whose nature sensitively depends on the relative order of the sample size  $n$  and aggregation number  $N$ , and whose solution is known only in special cases. For example, for  $N = N(n)$  viewed as a function of  $n$ , asymptotic normality of the moment estimators in the AR(1) case was proved in [24] in the case  $n/N(n) \rightarrow \infty$ ; analogous results for the density estimator were given in [11] and [31] in the case  $N = \infty$ , i.e. when the estimators are based on the associated process  $\{\xi_k, k \in \mathbb{Z}\}$ . Similar problems arise when considering panel data; see, e.g. Phillips and Moon [40], who discussed related phenomena in the case  $n/N(n) \rightarrow \infty$ .

Let  $d$  be a positive integer. The purpose of the present paper is to provide a weak invariance principle for multivariate functions  $f(Y_k^{(N)}, \dots, Y_{k-d+1}^{(N)})$  of aggregated linear processes with an arbitrary (possibly infinite) number of random parameters. More precisely, we will study the asymptotic behavior of  $\sum_{k=1}^{\lfloor nt \rfloor} f(Y_k^{(N)}, \dots, Y_{k-d+1}^{(N)})$  as  $n, N \rightarrow \infty$ , considering all interesting cases for the orders of magnitude of  $n$  and  $N$ . Our results will cover processes with short as well as long memory, naturally with a completely different class of admitted

functionals  $f(Y_k^{(N)}, \dots, Y_{k-d+1}^{(N)})$ . For the random coefficients of the underlying process  $X_k = \sum_{j=1}^{\infty} \alpha_j \varepsilon_{k-j}$ , we will assume throughout the paper that

(B)  $\sup_{i \in \mathbb{N}_0} |\alpha_i| \leq C$  almost surely (a.s.)

for some constant  $C > 0$ . Our basic assumption in the short memory case will be

(SM)  $\sum_{j=1}^{\infty} \|\alpha_j\| < \infty$ ,

where  $\|\cdot\|_2$  denotes the  $L^2$ -norm. This implies that not only the process  $\{X_k, k \in \mathbb{Z}\}$ , but also the Gaussian process  $\{\xi_k, k \in \mathbb{Z}\}$ , obtained as the limit of the normalized aggregated process  $\{Y_k^{(N)}, k \in \mathbb{Z}\}$  as  $N \rightarrow \infty$ , has short memory. In the case of long memory, we require that (SM) fails, but

(LM)  $\max_{1 \leq h \leq d} \sum_{i=0}^{\infty} \|\alpha_i - \alpha_{i+h}\| < \infty$  and  $\sum_{i=0}^{\infty} \mathbb{E}[|\alpha_i \alpha_{i+k}|] = \mathcal{O}(k^{-\beta})$ ,  $\beta > 0$ ,

holds, where  $d$  is the dimension of the functional  $f$  we are studying. Condition (LM) allows a slower decrease of the coefficients  $\alpha_j$ , as the degenerate example  $\alpha_j = 1/j$  shows; on the other hand, the admissible class of functionals is much smaller in this case. We do not make any assumption on the dependence of the random coefficients  $\alpha_j$ . Our main results will show that, under either condition, the partial sums of finite-dimensional functionals of  $\{Y_k^{(N)}, k \in \mathbb{Z}\}$  satisfy a weak invariance principle, a fact from which the limit distribution of many important statistics of aggregated processes can be deduced immediately.

The proofs of our results use martingale techniques, specifically, we will apply the decomposition method of Gordin [21], a powerful technique leading to sharp results for large classes of weakly dependent processes; see, e.g. [37] and [50] for recent contributions. The main technical difficulties in our paper are connected with extending the martingale decomposition method for the two-parameter situation provided by aggregation, and involve in particular the handling of the nonergodic setting. As will be clear from the proofs, our method applies to a large class of nonlinear processes as well, e.g. processes allowing a Volterra expansion with sufficiently rapidly decreasing coefficients. This class includes, among others, many ARCH- and GARCH-type processes. However, to keep our exposition at a reasonable length, in this paper we deal only with linear processes. We note that some of the results of the present paper were announced, without proof, in the conference proceedings [27].

Our paper is structured as follows. In Section 2 we introduce notation and prove some preliminary results. The main results are presented in Section 3, along with some additional remarks and comments. In Section 4 we apply the results to aggregated AR(1) sequences, where we revisit and extend results from the literature. The proofs are split into two parts. First, some necessary tools are presented in Section 5.1. Based on these tools, the proofs of the main results are given in Section 5.2.

## 2. Preliminaries

Let

$$X_k = \sum_{i=0}^{\infty} \alpha_i \varepsilon_{k-i}, \quad k \in \mathbb{Z},$$

be a linear process with random coefficients  $\alpha_i$  subject to conditions (SM) or (LM). Assume that  $\{\varepsilon_k, k \in \mathbb{Z}\}$  are independent, identically distributed random variables with  $\mathbb{E}[\varepsilon_k] = 0$ ,  $0 < \sigma_{\varepsilon}^2 = \mathbb{E}[\varepsilon_k^2] < \infty$ , and that the random vectors  $(\alpha_0, \alpha_1, \dots)$  and  $(\varepsilon_k, k \in \mathbb{Z})$  are independent.

Let

$$X_k^{(j)} = \sum_{i=0}^{\infty} \alpha_i^{(j)} \varepsilon_{k-i}^{(j)}, \quad k \in \mathbb{Z}, j = 1, 2, \dots,$$

be independent copies of  $\{X_k, k \in \mathbb{Z}\}$ , i.e. let  $\varepsilon_k^{(j)}, k \in \mathbb{Z}, j = 1, 2, \dots$ , be independent random variables distributed as the  $\varepsilon_k$  and let the arrays  $(\alpha_0^{(j)}, \alpha_1^{(j)}, \dots), j \in \mathbb{N}$ , be independent, distributed as  $(\alpha_0, \alpha_1, \dots)$  and also independent of the  $\varepsilon_k^{(j)}$ . Let  $\{Y_k^{(N)}, k \in \mathbb{Z}\}$  be the normalized aggregated process, i.e.

$$Y_k^{(N)} = \frac{X_k^{(1)} + \dots + X_k^{(N)}}{\sqrt{N}}, \quad k \in \mathbb{Z}, \quad (2.1)$$

and denote by  $\mathbf{Y}^{(N)} = (Y_k^{(N)}, \dots, Y_{k-d+1}^{(N)})^\top$  the  $d$ -dimensional aggregated vector. Define the  $\sigma$ -algebras  $\mathcal{G}_t^{(N)}, \mathcal{A}^{(N)}$ , and  $\mathcal{F}_t^{(N)}$  by

$$\begin{aligned} \mathcal{G}_t^{(N)} &= \sigma(\varepsilon_k^{(j)}, k \leq t, k \in \mathbb{Z}, 1 \leq j \leq N), \\ \mathcal{A}^{(N)} &= \sigma(\alpha_k^{(j)}, k \in \mathbb{N}_0, 1 \leq j \leq N), \\ \mathcal{F}_t^{(N)} &= \sigma(\mathcal{G}_t^{(N)} \cup \mathcal{A}^{(N)}). \end{aligned}$$

When there is no danger of confusion, we will omit the superscript  $(N)$ . We will frequently use the abbreviation  $\|\cdot\|_p = (\mathbb{E}[|\cdot|^p])^{1/p}$ ,  $p < \infty$ ; in particular, we denote the  $L_2$ -norm with  $\|\cdot\| := \|\cdot\|_2$ .

Note that conditions  $\sigma_\varepsilon^2 < \infty$  and (SM) or (LM) imply that the aggregated process  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$  is stationary in the weak sense; in other words,

$$\text{cov}(Y_i^{(N)}, Y_j^{(N)}) = \text{cov}(Y_1^{(N)}, Y_{j-i-1}^{(N)}) =: \phi_{j-i} \quad \text{for all } N \in \mathbb{N},$$

and, in particular, that the covariances  $\phi_h, h = 0, 1, 2, \dots$ , exist. For  $x \in \mathbb{R}^d$ , let  $|x|$  denote the usual Euclidean norm, and denote by  $\xrightarrow{w}$  weak convergence and by  $\xrightarrow{\text{FDD}}$  convergence of finite-dimensional distributions. In addition, by  $\xrightarrow{\mathbb{D}[0,1]}$  we denote weak convergence in the space of the càdlàg functions on the interval  $[0, 1]$  equipped with the Skorokhod topology.

Let  $\eta_k, k \in \mathbb{Z}$ , be a stationary  $d$ -dimensional Gaussian sequence. For a Borel-measurable function  $f: \mathbb{R}^d \mapsto \mathbb{R}$ , we formally define

$$\begin{aligned} (\sigma_\eta^f)^2 &:= \mathbb{E}[(f(\eta_1) - \mathbb{E}[f(\eta_1)])^2] \\ &+ 2 \sum_{k=1}^{\infty} \mathbb{E}[(f(\eta_1) - \mathbb{E}[f(\eta_1)])(f(\eta_{k+1}) - \mathbb{E}[f(\eta_{k+1})])]. \end{aligned} \quad (2.2)$$

Moreover, for fixed  $m \in \mathbb{N}$ , by  $\xi_k^{(m)} = \{\xi_k^{(m)}, k \in \mathbb{Z}\}$  we denote the weak limit of the  $m$ -dependent approximation  $\mathbf{Y}_k^{(N)} - \mathbb{E}[\mathbf{Y}_k^{(N)} | \mathcal{G}_{k-m}^{(N)}]$ , i.e.

$$\xi_k^{(m)} := \lim_{N \rightarrow \infty} \mathbf{Y}_k^{(N)} - \mathbb{E}[\mathbf{Y}_k^{(N)} | \mathcal{G}_{k-m}^{(N)}].$$

In particular, we formally define

$$\sigma_{\xi,*}^f := \lim_{m \rightarrow \infty} \sigma_{\xi^{(m)}}^f, \quad (2.3)$$

where we point out that we do not demand  $\sigma_{\xi,*}^f = \sigma_{\xi}^f$ .

### 3. Main results

As already mentioned, our main result is a functional central limit theorem for the normalized aggregated process  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$ . Note that if  $N$  is fixed and  $n \rightarrow \infty$ , we are dealing with a stationary sequence. There is a vast literature on functional central limit theorems dealing with this particular case. For more recent results, see [15], [16], [34], [37], [38], and [50]; excellent surveys can be found in [35] and [39]. On the other hand, we have seen that if  $n$  is fixed and  $N$  tends to  $\infty$ , then  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}} \xrightarrow{\text{FDD}} \{\xi_k\}_{k \in \mathbb{Z}}$ , which is a stationary ergodic Gaussian sequence. Limit theorems for such processes have been established (among others) by Taqqu [45], [46], Dobrushin and Major [18], Giraitis and Surgailis [19], and Arcones [4], [5]. So it is natural to expect that, if  $N = N(n)$  is a function of  $n$  and dominates  $n$  strong enough in the sense that  $\lim_n n/N(n) \rightarrow 0$  fast enough, we will get similar results.

Let  $g(s)$ ,  $s \in [0, 1]$ , be a Lipschitz continuous function, and set

$$\Lambda_n = \Lambda_n^{(N)} := \frac{1}{\sqrt{n}} \sum_{i=0}^n g\left(\frac{i}{n}\right) (f(Y_i^{(N)}) - \mathbb{E}[f(Y_i^{(N)}) \mid \mathcal{A}^{(N)}]),$$

where  $f$  is a map from  $\mathbb{R}^d$  to  $\mathbb{R}$ . Moreover, for  $L \geq 1$ , we define the  $d$ -dimensional polynomials  $Q_i(\mathbf{x})$ ,  $1 \leq i \leq L$ ,  $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ , as

$$Q_i(\mathbf{x}) = (x_j - x_{j+h_i})^{q_i} P_i(\mathbf{x}), \quad 1 \leq j, j + h_i \leq d, q_i \geq 0, \quad (3.1)$$

where  $P_i(\mathbf{x})$  are  $d$ -dimensional polynomials. For convenience, we restate conditions (B), (SM), and (LM).

(B)  $\sup_{i \in \mathbb{N}_0} |\alpha_i| \leq C$  a.s.

(SM)  $\sum_{i=0}^{\infty} \|\alpha_i\| < \infty$ .

(LM)  $\max_{1 \leq h \leq d} \sum_{i=0}^{\infty} \|\alpha_i - \alpha_{i+h}\| < \infty$  and  $\sum_{i=0}^{\infty} \mathbb{E}[|\alpha_i \alpha_{i+k}|] = \mathcal{O}(k^{-\beta})$ ,  $\beta > 0$ .

**Remark 3.1.** Note that (SM) implies that  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$  is a short-memory process, since we have  $\sum_{h=0}^{\infty} |\phi_h| < \infty$ . Contrarily, (LM) allows for long memory; see Section 4 for examples.

In the case where the aggregated process  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$  exhibits short memory, the following result is derived.

**Theorem 3.1.** Let  $\varepsilon$  be a copy of the random variables  $\{\varepsilon_k^{(j)}\}_{k \in \mathbb{Z}}$ ,  $1 \leq j \leq N$ , and let  $f$  be a map from  $\mathbb{R}^d$  to  $\mathbb{R}$ , satisfying the condition

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq C \sum_{i=1}^L |Q_i(\mathbf{x}) - Q_i(\mathbf{y})|,$$

where the  $Q_i(\mathbf{x})$  are as in (3.1) for some  $C > 0$  and  $L \geq 1$ . Let  $D = 4 \max_{1 \leq i \leq L} \deg(Q_i)$ . If both  $n$  and  $N$  tend to  $\infty$ , and condition (SM) and  $\mathbb{E}[\varepsilon^D] < \infty$  are valid, then

$$\Lambda_{[nt]}^{(N)} \xrightarrow{\mathbb{D}[0,1]} \sigma_{\xi}^f \int_0^t g(s) dW_s, \quad 0 \leq t \leq 1, \quad (3.2)$$

where  $W_t$  is a Brownian motion and  $\sigma_{\xi}^f$  is as in (2.2), where  $\xi = \{\xi_k, k \in \mathbb{Z}\}$  denotes the associated Gaussian sequence of  $\mathbf{Y}^{(N)} = \{\mathbf{Y}_k^{(N)}, k \in \mathbb{Z}\}$ . If  $N$  is fixed and  $n \rightarrow \infty$ , then it suffices to require  $\mathbb{E}[\varepsilon^{D/2}] < \infty$ , and instead of  $\sigma_{\xi}^f$ , we have an  $\mathcal{A}^{(N)}$ -measurable random variable (RV)  $\sigma_{\mathcal{A}^{(N)}}$  in (3.2), which is independent of  $\{W_t\}_{0 \leq t \leq 1}$ .

The above result is quite general, and it seems to be very difficult to prove a corresponding result where condition (LM) is valid but (SM) is not. However, by slightly narrowing the class of functions satisfying (3.1), the following result is obtained.

**Theorem 3.2.** *Suppose that the conditions of Theorem 3.1 are valid, with the exception that we only require condition (LM) instead of (SM). If, in addition,  $\min_{1 \leq i \leq L} q_i \geq 1$  in (3.1) then*

$$\Lambda_{[nt]}^{(N)} \xrightarrow{\mathbb{D}[0,1]} \sigma_{\xi,*}^f \int_0^t g(s) dW_s, \quad 0 \leq t \leq 1, \quad (3.3)$$

where  $W_t$  is a Brownian motion, and  $\sigma_{\xi,*}^f$  is as in (2.3) and exists. If  $N$  is fixed and  $n \rightarrow \infty$ , then it suffices to require  $\mathbb{E}[\varepsilon^{D/2}] < \infty$ , and instead of  $\sigma_{\xi,*}^f$ , we have an  $\mathcal{A}^{(N)}$ -measurable RV  $\sigma_{\mathcal{A}^{(N)}}^f$  in (3.3), which is independent of  $\{W_t\}_{0 \leq t \leq 1}$ .

**Remark 3.2.** Note that even though  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$  may exhibit long memory, we still obtain a Brownian motion as the limit distribution. This phenomena is related to the concept of the *Hermite rank of a function*; see, e.g. [4], [5], and [18].

**Remark 3.3.** One may suspect that even if  $\min_{1 \leq i \leq L} q_i \geq 1$  then we also have the relation  $\sigma_{\xi,*}^f = \sigma_{\xi}^f$ , despite the possible inclusion of long memory. This conjecture is fueled by [26, Lemma 3.86] (and also Proposition 4.1 below). General computations for  $\sigma_{\xi,*}^f$  can be found in, e.g. [5]; however, these do not seem to lead to simpler expressions in the present context.

**Remark 3.4.** Note that if  $N$  is fixed and  $n \rightarrow \infty$ , then  $\mathcal{A}^{(N)}$  is an invariant  $\sigma$ -algebra with respect to the usual shift operators  $\{\theta_k\}_{k \in \mathbb{Z}}$  that can be associated with the stationary process  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$ . In particular,  $\sigma_{\mathcal{A}^{(N)}}$  can then be represented as

$$\lim_{n \rightarrow \infty} \mathbb{E}[(\Lambda_n^{(N)})^2 \mid \mathcal{A}^{(N)}] = \sigma_{\mathcal{A}^{(N)}}^2 \int_0^1 g^2(s) ds.$$

Note the presence of weights  $g(i/n)$  in (3.3). Weighted statistics have proven to be very helpful in many fields, particularly in the case of weighted empirical processes (see, e.g. [29] and [43]), which is addressed in Corollary 3.3 below. In addition, such weights can also be used to account for additional seasonal or nonseasonal effects of the underlying process; see, e.g. [25] and the references therein.

In practical applications, the conditional expectation  $\mathbb{E}[f(Y_k^{(N)}) \mid \mathcal{A}^{(N)}]$  may be difficult to handle, and one would want to know when it could be replaced by  $\mathbb{E}[f(Y_k^{(N)})]$ . This question essentially boils down to the relation between  $n$  and  $N$ , more precisely, if we write  $N = N(n)$  as a function of  $n$ , whether  $\lim_{n \rightarrow \infty} n/N(n) = 0$  or  $\lim_{n \rightarrow \infty} N(n)/n = 0$ . Both cases have economic backgrounds: Granger [22] argued that in many economical problems  $N$  is much larger than  $n$ , whereas Beran *et al.* [8] gave examples of the opposite case. To clarify the situation in the case that (SM) holds, we need a preliminary result. To this end, let

$$f(x_1, \dots, x_d) = \sum_{s=0}^M \lambda_s T_s(x_1, \dots, x_d) = \sum_{\substack{s=0 \\ 2 \mid \deg T_s}}^M \lambda_s T_s(x_1, \dots, x_d) + \sum_{\substack{s=0 \\ 2 \nmid \deg T_s}}^M \lambda_s T_s(x_1, \dots, x_d)$$

be a  $d$ -dimensional polynomial, where  $\lambda_s \in \mathbb{R}$ ,  $M < \infty$ , and  $T_s(x_1, \dots, x_d) = \prod_{i=0}^{d-1} x_i^{q_i^{(s)}}$  are the corresponding monomials. For each  $0 \leq s \leq M$ , we have the relation

$$\mathbb{E}[T_s(\xi_k)] = g_{T_s}(\phi_0, \dots, \phi_{d-1}),$$

where  $g_{T_s}(x_1, \dots, x_d)$  is a  $d$ -dimensional polynomial; see, e.g. [4] and [5]. In addition, set

$$a_h^{(r)} = \mathbb{E}[\varepsilon_1^2] \sum_{i=0}^{\infty} \alpha_i^{(r)} \alpha_{i+h}^{(r)}, \quad r = 1, 2, \dots, N, \quad h = 0, 1, \dots, d-1.$$

**Proposition 3.1.** Assume that (SM) is valid and that the additional conditions of Theorem 3.2 hold. If  $f$  is a  $d$ -dimensional polynomial then we have

$$\sqrt{N}(\mathbb{E}[f(\mathbf{Y}_k^{(N)}) | \mathcal{A}^{(N)}] - \mathbb{E}[f(\mathbf{Y}_k^{(N)})]) \xrightarrow{w} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = \text{var} \left( \sum_{i=0}^{d-1} \phi_i^{-1} (a_i^{(1)} - \phi_i) \sum_{\substack{s=0 \\ 2 \mid \deg T_s}}^M \lambda_s q_i^{(s)} g_{T_s, \xi}(\phi_0, \dots, \phi_{d-1}) \right),$$

with the notation introduced above.

Note that  $\sigma^2 = 0$  if the polynomial  $f$  consists of only monomials with odd degree. Proposition 3.1 yields the following corollary.

**Corollary 3.1.** Let  $f$  be a  $d$ -dimensional polynomial. Assume that (SM) is valid and that the additional conditions of Theorem 3.2 hold. Then

- (i) if  $\lim_{n \rightarrow \infty} n/N(n) = 0$ , we can replace  $\mathbb{E}[f(\mathbf{Y}_k^{(N)}) | \mathcal{A}^{(N)}]$  in Theorem 3.2 with  $\mathbb{E}[f(\mathbf{Y}_k^{(N)})]$ ,
- (ii) if  $\lim_{n \rightarrow \infty} N(n)/n = 0$ , we have

$$\frac{\sqrt{N}}{n} \sum_{i=0}^n g\left(\frac{i}{n}\right) (f(\mathbf{Y}_{i,d}^{(N)}) - \mathbb{E}[f(\mathbf{Y}_{i,d}^{(N)})]) \xrightarrow{w} \int_0^1 g(s) ds \mathcal{N}(0, \sigma^2),$$

where  $\sigma^2$  is as in Proposition 3.1.

**Remark 3.5.** With some additional (notational) effort, one may also cover the case where  $f$  is a polynomial, as in Theorem 3.2 in the case that (LM) holds. A special case is treated in Corollaries 4.1 and 4.2 below.

If the function  $f$  is not a polynomial, and/or only (LM) holds, we can still provide the following result.

**Corollary 3.2.** Suppose that the assumptions of Theorem 3.2 hold and that

$$\lim_{n \rightarrow \infty} \frac{n}{\sqrt{N(n)}} = 0.$$

Then we can replace  $\mathbb{E}[f(\mathbf{Y}_k^{(N)}) | \mathcal{A}^{(N)}]$  in Theorem 3.2 with  $\mathbb{E}[f(\mathbf{Y}_k^{(N)})]$  or  $\mathbb{E}[f(\xi_k)]$ .

By increasing the aggregation level  $N$  with respect to  $n$ , we can partially extend Theorem 3.2.

**Theorem 3.3.** Let  $f$  be a map from  $\mathbb{R}^d$  to  $\mathbb{R}$  such that

$$\sup_x \frac{|f(x)|}{1 + |x|^{p+1}} < \infty \quad \text{for some } p > 0, \quad |f(x) - f(y)| \leq C(Q(x) - Q(y)),$$

for all  $x, y \in \mathbb{R}^d$  up to a finite number of points, where  $C$  is a positive constant and  $Q(x)$  is a  $d$ -dimensional polynomial. Suppose that  $\lim_{n \rightarrow \infty} n/\sqrt{N(n)} = 0$ , and that the assumptions of

Theorem 3.1 hold, together with  $\mathbb{E}[|\varepsilon|^m] < \infty$  for all  $m$ . Then

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{[nt]} g\left(\frac{i}{n}\right) (f(Y_{i,d}^{(N)}) - \mathbb{E}[f(Y_{i,d}^{(N)})]) \xrightarrow{\text{FDD}} \sigma_{\xi}^f \int_0^t g(s) dW_s, \quad 0 \leq t \leq 1,$$

where  $\sigma_{\xi}^f$  is as in (2.2).

If  $f(x_1, \dots, x_d) = \mathbf{1}_{\{x_1 \leq t_1, \dots, x_d \leq t_d\}}$  for fixed  $t_1, \dots, t_d \in \mathbb{R}$ , we obtain the following result for the weighted empirical distribution function  $F_n^{(g)}(t_1, \dots, t_d)$ .

**Corollary 3.3.** *Under the conditions of Theorem 3.3, the weighted empirical distribution function*

$$F_n^{(g)}(t_1, \dots, t_d) = \frac{1}{n} \sum_{k=1}^n g\left(\frac{k}{n}\right) \mathbf{1}_{\{Y_k^{(N)} \leq t_1, \dots, Y_{k+d-1}^{(N)} \leq t_d\}}$$

satisfies

$$\begin{aligned} & \sqrt{n} \left( F_n^{(g)}(t_1, \dots, t_d) - n^{-1} \sum_{k=1}^n g\left(\frac{k}{n}\right) \mathbb{P}(\xi_1 \leq t_1, \dots, \xi_d \leq t_d) \right) \\ & \xrightarrow{w} \sigma_{\xi}^1 \sqrt{\int_0^1 g^2(s) ds} \mathcal{N}(0, 1), \end{aligned}$$

where  $\{\xi_k\}_{k \in \mathbb{Z}}$  denotes the associated process of  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$ , and  $\sigma_{\xi}^1$  is as in (2.2) with  $f = \mathbf{1}_{\{x_1 \leq t_1, \dots, x_d \leq t_d\}}$ .

Theorem 3.2 also allows us to give a weighted version of Theorem 2 of [4].

**Corollary 3.4.** *Let  $g(s)$  be a continuous function on  $[0, 1]$ , and let  $f$  be a Borel-measurable map from  $\mathbb{R}^d$  to  $\mathbb{R}$ . Let  $\{\xi_k\}_{k \in \mathbb{Z}}$  be a zero-mean stationary Gaussian sequence with covariances  $\phi_k = \mathbb{E}[\xi_0 \xi_k]$ ,  $k \in \mathbb{Z}$ , satisfying  $\limsup_n n^{-1} \sum_{k,l=1}^n |\phi_{|k-l||} < \infty$ . Then, under  $\mathbb{E}[f(\xi_{0,d})^2] < \infty$ , we have*

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right) (f(\xi_k) - \mathbb{E}[f(\xi_k)]) \xrightarrow{\text{FDD}} \sigma_{\xi}^f \int_0^t g(s) dW_s.$$

#### 4. Aggregated AR(1) processes and long memory

In this short section we will briefly touch on the disaggregation problem if the micro-level processes  $\{X_k^{(j)}, k \in \mathbb{Z}\}_{j=1, \dots, N}$  are AR(1) processes. In practice, it is often reasonable to approximate individual data by simple time series models, such as, e.g. AR(1) and GARCH(1, 1) (see [12], [33], and [51]). Very often, more complex models neither provide an advantage in accuracy nor in efficiency of the estimates, and usually are difficult to study from the theoretical point of view. Consider now the aggregated process defined by (1.1), where  $\{X_k^{(j)}, k \in \mathbb{Z}\}_{j=1, \dots, N}$  are independent, identically distributed random coefficient AR(1) processes. That is, for any  $j = 1, \dots, N$ ,

$$X_k^{(j)} = a^{(j)} X_{k-1}^{(j)} + \varepsilon_k^{(j)}, \quad k \in \mathbb{Z},$$

where  $\varepsilon_k^{(j)}$ ,  $k \in \mathbb{Z}$ ,  $j = 1, 2, \dots$ , and the copies  $\varepsilon_k$  are independent, identically distributed RVs with zero mean and variance  $\sigma^2 < \infty$ , and the  $a^{(j)}$ ,  $j = 1, 2, \dots$ , are independent, identically



distributed RVs. We assume that  $\{\varepsilon_k^{(j)}, k \in \mathbb{Z}, j = 1, 2, \dots\}$  and  $\{a^{(j)}, j = 1, 2, \dots\}$  are independent. Finally, let  $a$  be an independent RV having the same distribution as  $(a^{(j)})$ ,  $j = 1, 2, \dots$ .

Most of the recent results dealing with the disaggregation problem considered aggregated AR(1) processes; see, e.g. [8], [10], [11], [24], and [31]. For some insights on AR(2) processes, we refer the reader to [27]. A profound discussion of these results is beyond the scope of the present paper, and will be given together with some extensions elsewhere. Here and now, we focus on a special function  $f$ , which is a key ingredient in the disaggregation problem:

$$\begin{aligned} f(Y_k^{(N)}, Y_{k-2}^{(N)}, Y_{k-h}^{(N)}, Y_{k-h-2}^{(N)}) \\ = Y_k^{(N)}(Y_{k-h}^{(N)} - Y_{k-h-2}^{(N)} - \mathbb{E}[a^h]Y_k^{(N)} + \mathbb{E}[a^h]Y_{k-2}^{(N)}), \quad h = 1, 2, \dots \end{aligned} \quad (4.1)$$

The importance of this function stems from the fact that asymptotic results concerning moment and density estimators associated to the RV  $a$  may be derived from results regarding this particular function. In fact, we can show that (cf. [11] and [27])

$$\begin{aligned} \mathbb{E}[a^{2n}] &= \frac{\phi_{2n} - \phi_{2n+2}}{\phi_0 - \phi_2}, \quad n \in \mathbb{N}_0, \\ \mathbb{E}[a^{2n-1}] &= \frac{\phi_{2n-1} - \phi_{2n+1}}{\phi_0 - \phi_2}, \quad n \geq 1, \end{aligned}$$

where  $\sigma^2 = \mathbb{E}[\varepsilon_1^2]$ . Hence, we can estimate the moments  $\mathbb{E}[a^h]$  via

$$Z_{n,N,h} = \frac{\sum_{k=h+1}^n Y_k^{(N)} Y_{k-h}^{(N)} - \sum_{k=h+3}^n Y_k^{(N)} Y_{k-h-2}^{(N)}}{\sum_{k=1}^n (Y_k^{(N)})^2 - \sum_{k=3}^n Y_k^{(N)} Y_{k-2}^{(N)}}.$$

Note in particular that  $Z_{n,N,1} = Z_{n,N}$ , where  $Z_{n,N}$  is given in (1.2). Using Theorem 3.2, we will derive the asymptotic distribution of  $Z_{n,N,h}$  under quite general conditions, given below.

**Assumption 4.1.** Suppose that the following assertions hold.

- (i)  $\mathbb{E}[\varepsilon_1^8] < \infty$ .
- (ii)  $|a| \leq 1$  a.s.
- (iii)  $\sum_{j=0}^{\infty} \|(1 - a^2)a^j\| < \infty$ .
- (iv)  $|\mathbb{E}[a^k/(1 - a^2)]| < \infty$ .

We obtain the following result.

**Proposition 4.1.** If both  $n$  and  $N$  tend to  $\infty$ , we have, under Assumption 4.1,

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor nt \rfloor} g\left(\frac{i}{n}\right) (f(Y_i^{(N)}) - \mathbb{E}[f(Y_i^{(N)}) | \mathcal{A}^{(N)}]) \xrightarrow{w} \sigma_{\xi}^f \int_0^t g(s) dW_s, \quad 0 \leq t \leq 1,$$

in  $\mathbb{D}[0, 1]$ , where  $W_t$  is a Brownian motion and  $\sigma_{\xi}^f$  is as in (2.2).

Note that Assumption 4.1(i), (ii), and (iv) allow for long memory, since  $\text{cov}(Y_k^{(N)}, Y_l^{(N)}) = \mathbb{E}[\varepsilon^2] \mathbb{E}[a^{k-l}(1 - a^2)^{-1}]$  for  $k \geq l$ . Assumption 4.1(iii) can be investigated more easily by introducing a density function, which is done in, e.g. [24] and [31] when discussing the

disaggregation problem. In this context, the following type of density functions  $\varphi(x)$  are commonly used:

$$\varphi(x) = (1+x)^{1-d_1}(1-x)^{1-d_2}\psi(x), \quad x \in [-1, 1], \quad d_1, d_2 < 1.$$

Here  $\psi(x)$  is continuous on  $[-1, 1]$  and does not vanish at  $-1$  and  $1$ . For simplicity, we will consider the case where

$$\varphi(x) = (1-x^2)^{1-2d}\psi(x), \quad x \in [-1, 1], \quad d < \frac{1}{2}.$$

In this case, the condition  $\mathbb{E}[1/(1-|a|)^3] < \infty$  of Horváth and Leipus [24] implies that  $d < -\frac{1}{2}$ , which is a short-memory condition. Evaluating Assumption 4.1(iii), we obtain

$$\|(1-a^2)a^j\|^2 = \mathcal{O}(\Gamma[3-2d]\Gamma[\frac{1}{2}+j]\Gamma^{-1}[3-2d+\frac{1}{2}+j]) = \mathcal{O}(j^{-3+2d}).$$

Hence, we require  $\sum_{j=1}^{\infty} j^{-3/2+d} < \infty$ , which yields  $d < \frac{1}{2}$ . This is also the condition of Celov *et al.* [11]. Note that this also implies that the aggregated process  $\{Y_k^{(N)}\}_{k \geq 1}$  exhibits a long-memory behavior if  $0 < d < \frac{1}{2}$ . Proposition 4.1 allows us to derive the following corollaries, which extend the results presented in [24] and [27].

**Corollary 4.1.** *Let  $N = N(n)$  be an increasing function in  $n$ , and assume that  $g(s)$  is a Lipschitz continuous function on  $[0, 1]$  such that  $\int_0^1 g(s) ds \neq 0$ . Then, under Assumption 4.1, we have, for  $\lim_{n \rightarrow \infty} N(n)/n = 0$ ,*

$$\sqrt{N} \frac{Z_{n,N,h} - \mathbb{E}[a^h]}{\sqrt{Z_{n,N,2} - Z_{n,N,1}^2}} \xrightarrow{w} \mathcal{N}(0, 1).$$

Note that the variance of the limit distribution does not depend on the choice of the function  $g(s)$ , which is rather unexpected. However, this feature is lost in the second case.

**Corollary 4.2.** *Assume that  $g(s)$  is a Lipschitz continuous function on  $[0, 1]$  such that  $\int_0^1 g(s) ds \neq 0$ . Then, under Assumption 4.1, we have, for  $\lim_{n \rightarrow \infty} n/N(n) = 0$ ,*

$$\sqrt{n}(Z_{n,N,h} - \mathbb{E}[a^h]) \xrightarrow{w} \frac{\sqrt{\int_0^1 g(s)^2 ds}}{\int_0^1 g(s) ds} \mathcal{N}(0, \sigma_f^2), \quad (4.2)$$

where  $\sigma_f = \sigma_{\xi}^f$ ,  $\xi = \{\xi_k, k \in \mathbb{Z}\}$  denotes the associated Gaussian process, and  $\sigma_{\xi}^f$  is as in (2.2).

A distribution-free limit can be established by estimating  $\sigma_f^2$  with a Bartlett-type estimator; we omit the details. Also, note that the variance of the limit distribution in (4.2) is greater than or equal to  $\sigma_f^2$  by the Cauchy–Schwarz inequality, and we have equality if and only if  $g(s)$  is a.s. constant. Thus,  $g(s) \equiv 1$  yields the estimator with minimal variance.

## 5. Proofs

The main tool in our proofs will be martingale decomposition, and we will make essential use of the ideas in [26] and [39]. To verify the conditions for the martingale approximation, we will use a second main class of tools, namely approximation results related to the central limit theorem for a stationary Gaussian process. Unless otherwise stated, in all proofs,  $C$  always denotes a generic constant that may vary from one formula to another.

### 5.1. Some tools

We postpone the proof of Theorem 3.2 to Section 5.2, and first establish the other results of Section 3. Depending on assumptions (SM) and (LM), we will require one of the following results.

**Lemma 5.1.** *Let  $g$  be a bounded function on  $[0, 1]$ , and let  $f$  be a map from  $\mathbb{R}^d$  to  $\mathbb{R}$ . Let  $\xi_k = (\xi_k^{(1)}, \dots, \xi_k^{(d)})^\top$  be a zero-mean stationary Gaussian sequence in  $\mathbb{R}^d$  with covariances  $r^{p,q}(k) = \mathbb{E}[\xi_0^{(p)} \xi_k^{(q)}]$ ,  $1 \leq p, q \leq d$ , satisfying  $\limsup_n n^{-1} \sum_{k,l=1}^n |r^{p,q}(k-l)| < \infty$ . Then, under  $\mathbb{E}[f(\xi_k)^2] < \infty$ , we have*

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right) (f(\xi_k) - \mathbb{E}[f(\xi_k)]) \right\| \leq Cd\sqrt{t} \|f(\xi_0)\| \quad \text{for all } n,$$

where the constant  $C$  depends only on the covariances.

**Lemma 5.2.** *Let  $\xi_k = (\xi_k^{(1)}, \dots, \xi_k^{(d)})^\top$  be a stationary Gaussian sequence in  $\mathbb{R}^d$  with covariances  $r^{p,q}(k) = \mathbb{E}[\xi_0^{(p)} \xi_k^{(q)}]$ ,  $1 \leq p, q \leq d$ , satisfying  $\limsup_k k^\beta |r^{p,q}(k)|/L(k) = b_{p,q} < \infty$ ,  $0 < \beta < 1$ , for each  $1 \leq p, q \leq d$ , where  $L(k)$  is a slowly varying function. Then, under  $\mathbb{E}[f(\xi_k)^2] < \infty$ , we have*

$$\left\| a_n^{-1} \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right) (f(\xi_k) - \mathbb{E}[f(\xi_k)]) \right\| \leq dC\sqrt{t} \|f(\xi_0)\| \quad \text{for all } n,$$

where  $a_n = n^{1-\tau\beta/2}$  and  $\tau$  is the rank of the function  $f$ .

For a detailed treatment of the rank of a function, we refer the reader to [4] or [18]. In our case, we will always choose  $\tau = 1$ , since we are only interested in an upper bound (see Lemma 5.6 below).

Both lemmas are slightly changed versions of results in [4, Theorem 2, Theorem 6].

*Proofs of Lemmas 5.1 and 5.2.* Lemma 5.1 is obtained from [4, Lemma 1], following the proof of [4, Theorem 2]. Lemma 5.2 is obtained similarly.

We set  $\mathbf{Y}_k^{(N)} = (Y_k^{(N)}, \dots, Y_{k+d-1}^{(N)})^\top$  and  $\xi_k = (\xi_k, \dots, \xi_{k+d-1})^\top$ , where  $\{\xi_k\}_{k \in \mathbb{Z}}$  is the associated Gaussian sequence. In addition, we set

$$\sigma_{n,\xi}^f = \left\| \frac{1}{\sqrt{n}} \sum_{k=0}^n \{f(\xi_k) - \mathbb{E}[f(\xi_k)]\} \right\|, \quad \sigma_{n,Y}^f = \left\| \frac{1}{\sqrt{n}} \sum_{k=0}^n \{f(\mathbf{Y}_k^{(N)}) - \mathbb{E}[f(\mathbf{Y}_k^{(N)})]\} \right\|.$$

For the purposes of the next lemma, let  $\{\mathbf{Z}_k\}_{k \in \mathbb{N}}$  be a sequence of independent random vectors in  $\mathbb{R}^d$  with zero mean and identity covariance matrix. Define

$$S_N = \frac{1}{\sqrt{N}} \sum_{k=1}^N \mathbf{Z}_k, \quad \delta_N = N^{-1/2} \mathbb{E}[|\mathbf{Z}_1|^3], \quad \eta_{N,p} = N^{-(p-1)/2} \mathbb{E}[|\mathbf{Z}_1|^{p+1}].$$

Let  $\mathcal{P}_N$  denote the distribution of  $S_N$ , and  $\mathcal{N}$  the  $d$ -dimensional standard normal distribution. For a Borel-measurable function  $f$  from  $\mathbb{R}^d$  to  $\mathbb{R}$ , we set

$$\|f\|_p^* = \sup_x \frac{|f(x)|}{1 + |x|^{p+1}}.$$

Note that if  $\|f\|_p^* < \infty$  holds for some function  $f(x_1, \dots, x_d)$  then this remains true for  $g$  with  $g(x_1, \dots, x_d) = f(A(x_1, \dots, x_d)^\top)$ , where  $A$  is a linear transformation.

The following result, stated here as a lemma, is due to Sweeting [44].

**Lemma 5.3.** *Let  $p \geq 3$  be an integer, and let  $f$  be a Borel-measurable function on  $\mathbb{R}^d$  such that  $\|f\|_p^* < \infty$  holds. Assume that  $\mathbb{E}[|Z_1|^{p+1}] < \infty$ . Then there exists a constant  $C(d, p) < \infty$  such that, for all  $N \geq 1$ ,  $\mathbb{E}[|S_N|^{p+1}] \leq C(d, p)(1 + \eta_{N,p})$ . Moreover, we have*

$$\left| \int_{\mathbb{R}^d} f d(\mathcal{P}_N - \mathcal{N}) \right| \leq C(d, p) \left( \|f\|_p^* (\delta_N + \eta_{N,p}) + \int_{\mathbb{R}^d} \omega_f^{C\delta_N} d\mathcal{N} \right),$$

where  $C$  is an absolute constant and  $\omega_f^\delta(x) = \sup\{|f(y_1) - f(y_2)| : |y_i - x| < \delta, i = 1, 2\}$ .

Given  $f : \mathbb{R}^d \mapsto \mathbb{R}$ , define  $\hat{f} : \mathbb{R}^{2d} \mapsto \mathbb{R}$  by

$$\hat{f}(x_1, \dots, x_{2d}) = f(x_1, \dots, x_d) f(x_{d+1}, \dots, x_{2d}).$$

**Proposition 5.1.** *Let  $Y_k^{(N)} = N^{-1/2} \sum_{j=1}^N X_k^{(j)}$  be the aggregated process as in Definition 1.1, and let  $f$  be a Borel-measurable function from  $\mathbb{R}^d$  to  $\mathbb{R}$  such that*

$$|f(x) - f(y)| \leq C|Q(x) - Q(y)| \quad (5.1)$$

for all  $x, y \in \mathbb{R}^d$  up to a finite number of points, where  $Q$  is a  $d$ -dimensional polynomial with degree  $q$ . Assume that conditions (B) and either (SM) or (LM) hold, and that, for some  $p \geq \max\{3, q - \frac{1}{2}\}$ , we have  $\mathbb{E}[|X_1|^{2p+1}] < \infty$ . Then we have

$$\left| \mathbb{E} \left[ \frac{1}{n} \sum_{k,l=1}^{[nt]} f(\xi_k) f(\xi_{l,d}) \right] - \mathbb{E} \left[ \frac{1}{n} \sum_{k,l=1}^{[nt]} f(Y_k^{(N)}) f(Y_{l,d}^{(N)}) \right] \right| \leq C_1 \frac{C(2d, 2p)n}{\sqrt{N}} \quad (5.2)$$

and

$$\left| \mathbb{E} \left[ \frac{1}{n} \sum_{k=1}^{[nt]} f(\xi_k) \right] - \mathbb{E} \left[ \frac{1}{n} \sum_{k=1}^{[nt]} f(Y_k^{(N)}) \right] \right| \leq C_2 \frac{C(2d, 2p)}{\sqrt{N}}, \quad (5.3)$$

where  $C_1$  and  $C_2$  depend only on the covariance structure, and  $C(2d, 2p)$  is a constant depending only on  $d$  and  $p$ .

*Proof.* It suffices to show (5.2) since the same argument applies to (5.3). Owing to (5.1), we readily verify that  $\|f\|_p^* < \infty$  for  $p \geq \max\{3, q - \frac{1}{2}\}$ . Moreover, for a  $d$ -dimensional continuous function  $h$  satisfying (5.1), we have

$$\sup\{|h(y_1) - h(y_2)| : |y_i - x| \leq \delta, i = 1, 2\} \leq \delta \sup_{\{|y_i - x| \leq \delta, i=1,2\}} |R(y_1, y_2)|, \quad (5.4)$$

where  $R$  is a  $2d$ -dimensional polynomial with finite degree. Note that either condition (SM) or (LM) implies the existence of the covariance matrix  $\Lambda(k, 2d)$  of the vector  $(\xi_k, \dots, \xi_{k+d-1}, \xi_0, \dots, \xi_{d-1})^\top$ . To make Lemma 5.3 applicable, we define the functions  $\hat{f}_k(x)$  as  $\hat{f}_k(x) := \hat{f}(L(k, 2d)^{1/2}x)$ . Since  $\mathbb{E}[\xi_0^2] < \infty$ , the elements of  $L(k, 2d)^{1/2}$  are all bounded, one readily verifies that there exists a constant  $C$  such that

$$\|\hat{f}_k\|_{2p}^* \leq C \quad \text{for all } k, \quad (5.5)$$

and a similar bound as in (5.4) is also obtained. Let  $(x_1, \dots, x_m)$  be the finite set of discontinuity points of  $\hat{f}_k$ , and denote by  $\mathcal{B}_\delta(x_i)$  the closed,  $2d$ -dimensional  $\delta$ -ball around  $x_i$ ,  $1 \leq i \leq m$ . Then, owing to (5.4) and (5.5),

$$\begin{aligned} \left| \int_{\mathbb{R}^{2d}} \omega_{\hat{f}_k}^{C\delta_N} d\mathcal{N} \right| &\leq \left| \int_{\mathbb{R}^{2d} \setminus \bigcup_{i=1}^m \mathcal{B}_\delta(x_i)} \omega_{\hat{f}_k}^{C\delta_N} d\mathcal{N} \right| + \left| \int_{\bigcup_{i=1}^m \mathcal{B}_\delta(x_i)} \omega_{\hat{f}_k}^{C\delta_N} d\mathcal{N} \right| \\ &\leq C\delta_N \quad \text{for all } k. \end{aligned}$$

An application of Lemma 5.3 yields

$$\begin{aligned} &|\mathbb{E}[f(\xi_k)f(\xi_{l,d})] - \mathbb{E}[f(Y_k^{(N)})f(Y_{l,d}^{(N)})]| \\ &= \left| \int_{\mathbb{R}^{2d}} \hat{f}_k d(\mathcal{P}_n - \mathcal{N}) \right| \\ &\leq C(2d, 2p) \left( \|\hat{f}\|_{2p}^*(\delta_N + \eta_{N,2p}) + \int_{\mathbb{R}^{2d}} \omega_{\hat{f}_k}^{C\delta_N} d\mathcal{N} \right) \\ &\leq C(2d, 2p) (\|\hat{f}\|_{2p}^*(\delta_N + \eta_{N,2p}) + C\delta_N). \end{aligned}$$

Thus, we obtain

$$\begin{aligned} &\left| \mathbb{E} \left[ \frac{1}{n} \sum_{k,l=1}^{[nt]} f(Y_k^{(N)})f(Y_{l,d}^{(N)}) \right] - \mathbb{E} \left[ \frac{1}{n} \sum_{k,l=1}^{[nt]} f(\xi_k)f(\xi_{l,d}) \right] \right| \\ &\leq \frac{1}{n} \sum_{k,l=1}^{[nt]} \left| \int_{\mathbb{R}^{2d}} \hat{f}_k d(\mathcal{P}_n - \mathcal{N}) \right| \\ &\leq nt^2 C(2d, 2p) \sup_k (\|\hat{f}_k\|_p^*(\delta_N + \eta_{2p,N}) + C\delta_N) \\ &\leq C_1 \frac{C(2d, 2p)n}{\sqrt{N}}, \end{aligned}$$

and, similarly,

$$\left| \frac{1}{n} \sum_{k,l=1}^{[nt]} \mathbb{E}[f(Y_k^{(N)})] \mathbb{E}[f(Y_{l,d}^{(N)})] - \frac{1}{n} \sum_{k,l=1}^{[nt]} \mathbb{E}[f(\xi_k)] \mathbb{E}[f(\xi_{l,d})] \right| \leq C_2 \frac{C(d, p)n}{\sqrt{N}};$$

hence, the claim follows.

Let us assume again that the conditions of Proposition 5.1 hold for a  $d$ -dimensional function  $f$ ,  $\|f\|_p^* < \infty$  for  $0 \leq p < \infty$ . Let  $f_q$  be a polynomial approximation of degree  $q$  of  $f$ , with  $q \geq p$ . Then we can write  $f = f_q + f_{>q}$ , and, by  $\|f\|_p^* < \infty$ , we have  $\|f_{>q}\|_q^* < \infty$ . Hence, we obtain the following result.

**Corollary 5.1.** *Let  $f$  be a  $d$ -dimensional function such that  $\|f\|_p^* < \infty$ ,  $0 \leq p < \infty$ , and assume that the conditions of Proposition 5.1 are satisfied together with (SM). Then*

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} f_{>q}(Y_k^{(N)}) \right\|^2 \leq C \frac{C(2d, 2q)n}{\sqrt{N}} + C \|f_{>q}(\xi_{0,d})\|^2,$$

where  $\{\xi_k\}_{k \in \mathbb{Z}}$  is the associated sequence of  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$ .

*Proof.* It follows from Proposition 5.1 that

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} f_{>q}(Y_k^{(N)}) \right\|^2 = C \frac{C(d, q)n}{\sqrt{N}} + \left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} f_{>q}(\xi_k^{(N)}) \right\|^2.$$

One readily verifies that (SM) implies that the associated sequence  $\{\xi_k\}_{k \in \mathbb{Z}}$  satisfies the covariance condition of Lemma 5.1; hence, we obtain  $\|n^{-1/2} \sum_{k=1}^{\lfloor nt \rfloor} f_{>q}(\xi_k^{(N)})\|^2 \leq C \|f_{>q}(\xi_{0,d})\|^2$ , which completes the proof.

## 5.2. Proofs of the main theorems

This subsection is devoted to the proofs of Theorems 3.2 and 3.3, Proposition 3.1, and Corollaries 3.1, 3.2, and 3.4. Unless otherwise stated, in all the proofs,  $C$  always denotes a dynamic constant that may vary from one formulae to another. The proof of Theorem 3.2 consists of two steps. The first step provides a martingale representation, where we will rely heavily on results presented in [39]. The second step is to verify conditions enabling us to use martingale limit theorems, which will involve estimates for the square bracket of martingales. Throughout this subsection, we will assume that the assumptions of Theorem 3.2 hold, and, for all the proofs, we assume without loss of generality that  $L = 1$ , i.e.  $|f(x) - f(y)| \leq C|Q(x) - Q(y)|$ , where  $Q(x)$  is a  $d$ -dimensional polynomial and  $C > 0$ . For notational convenience, we will omit the superscript  $(N)$  in the notation  $Y_k^{(N)} = Y_k$  for the aggregated process in this subsection, except when we really want to emphasize the dependence on  $N = N(n)$ . Let  $k \geq t$ , write  $X_k^{(j)}$  as

$$X_k^{(j)} = \sum_{i=0}^{k-t-1} \alpha_i^{(j)} \varepsilon_{k-i}^{(j)} + \alpha_{k-t}^{(j)} \varepsilon_t^{(j)} + \sum_{i=k-t+1}^{\infty} \alpha_i^{(j)} \varepsilon_{k-i}^{(j)} = X_{k,>t}^{(j)} + \alpha_{k-t}^{(j)} \varepsilon_t^{(j)} + X_{k,<t}^{(j)},$$

and observe that  $X_{k,>t}^{(j)}$  is independent of  $\mathcal{G}_t^{(N)}$ . We introduce the notation

$$\begin{aligned} Y_{k,>t}^{(N)} &= \frac{1}{\sqrt{N}} \sum_{j=1}^N X_{k,>t}^{(j)}, & Y_{k,<t}^{(N)} &= \frac{1}{\sqrt{N}} \sum_{j=1}^N X_{k,<t}^{(j)}, \\ Y_{k,\geq t}^{(N)} &= \frac{1}{\sqrt{N}} \sum_{j=1}^N X_{k,>t}^{(j)} + \alpha_{k-t}^{(j)} \varepsilon_t^{(j)}, & Y_{k,=t}^{(N)} &= \frac{1}{\sqrt{N}} \sum_{j=1}^N \alpha_{k-t}^{(j)} \varepsilon_t^{(j)}, \\ Y_{k,\leq t}^{(N)} &= \frac{1}{\sqrt{N}} \sum_{j=1}^N X_{k,<t}^{(j)} + \alpha_{k-t}^{(j)} \varepsilon_t^{(j)}. \end{aligned}$$

For a process  $U_i$ , we define the projection operator  $\mathcal{P}_k(U_i) = \mathbb{E}[U_i \mid \mathcal{F}_k] - \mathbb{E}[U_i \mid \mathcal{F}_{k-1}]$ . In connection with the projection operator and the previous notation, we will use the coupling method. This method has been used by Wu [48], [49] and Dedecker and Doukhan [15], among others. The method consists of coupling a sequence of RVs with a special copy of that sequence, and is particularly helpful when dealing with projections of the type of  $\mathcal{P}_k(U_i)$ . To this end, given an innovation  $\varepsilon_k^{(j)}$ , define  $(\varepsilon_k^{(j)})'$  as another innovation with the same law, independent of all  $\varepsilon_k^{(j)}$ ,  $1 \leq j \leq N$ ,  $k \in \mathbb{Z}$ . In this spirit, we also introduce  $Y_{k,>t}^{(N)'}$ ,  $Y_{k,<t}^{(N)'}$ ,  $\dots$ , which are defined as before with the innovations  $(\varepsilon_k^{(j)})'$ .

**Lemma 5.4.** Let  $W_k^{(d)} = K(Y_k^{(N)}, \dots, Y_{k+d-1}^{(N)}) - \mathbb{E}[K(Y_k^{(N)}, \dots, Y_{k+d-1}^{(N)}) \mid \mathcal{A}^{(N)}]$ , where  $K$  satisfies  $|K(\mathbf{y}) - K(\mathbf{x})| \leq C|\mathcal{Q}(\mathbf{y}) - \mathcal{Q}(\mathbf{x})|$  with  $\mathbf{x} = (x_1, \dots, x_d)$  and  $\mathbf{y} = (y_1, \dots, y_d) \in \mathbb{R}^d$ , such that

$$\mathcal{Q}(\mathbf{x}) = (x_j - x_{j+h})^q P(\mathbf{y}), \quad 1 \leq j, h \leq d, q \geq 0,$$

where  $P(\mathbf{y})$  is a polynomial with degree  $D$ . Moreover, we assume that  $\mathbb{E}[\varepsilon_1^{2D}] < \infty$ . If, in addition, (SM) holds if  $q \geq 0$  or (LM) holds if  $q \geq 1$ , then

- (i)  $\mathbb{E}[W_k^{(d)} \mid \mathcal{F}_{-\infty}^{(N)}] = 0$  for all  $N$ , where  $\mathcal{F}_{-\infty}^{(N)} := \bigcap_{i \in \mathbb{N}} \mathcal{F}_{-i}^{(N)}$ ,
- (ii)  $\sum_{j=0}^{\infty} \|\mathcal{P}_{k-j}(W_k^{(d)})\| < \infty$  for all  $k$ .

*Proof.* We first verify (i). Owing to Kolmogorov's zero-one law, it follows that  $\mathcal{G}_{-\infty}^{(N)} := \bigcap_{i \in \mathbb{N}} \mathcal{G}_{-i}^{(N)}$  is the trivial  $\sigma$ -algebra  $\{\Omega, \emptyset\}$  for each  $N$ . Hence, we deduce that  $\mathcal{F}_{-\infty}^{(N)} = \mathcal{A}^{(N)}$ , and, thus, by the definition of  $W_k^{(d)}$ , the claim follows.

To prove (ii), we note that

$$\|\mathcal{P}_{k-j}(W_k^{(d)})\| = \|\mathbb{E}[W_k^{(d)} \mid \mathcal{F}_{k-j}] - \mathbb{E}[W_k^{(d)} \mid \mathcal{F}_{k-j-1}]\|.$$

Hence, we will take a closer look at  $\mathbb{E}[W_k^{(d)} \mid \mathcal{F}_t] - \mathbb{E}[W_k^{(d)} \mid \mathcal{F}_{t-1}]$ , with  $t = k - j$ . We set  $\Lambda_k = (Y_k^{(N)}, \dots, Y_{k+d-1}^{(N)}) = (Y_{k,>t}^{(N)} + Y_{k,=t}^{(N)} + Y_{k,<t}^{(N)}, \dots, Y_{k+d-1,>t}^{(N)} + Y_{k+d-1,=t}^{(N)} + Y_{k+d-1,<t}^{(N)})$  and  $\Lambda_k^{t'} = (Y_{k,>t}^{(N)} + Y_{k,=t}^{(N)}, \dots, Y_{k+d-1,>t}^{(N)} + Y_{k+d-1,=t}^{(N)} + Y_{k+d-1,<t}^{(N)})$ . Using the coupling method, it follows that

$$\begin{aligned} |\mathbb{E}[W_k^{(d)} \mid \mathcal{F}_t] - \mathbb{E}[W_k^{(d)} \mid \mathcal{F}_{t-1}]| &= |\mathbb{E}[K(\Lambda_k) \mid \mathcal{F}_t] - \mathbb{E}[K(\Lambda_k^{t'}) \mid \mathcal{F}_{t-1}]| \\ &= |\mathbb{E}[K(\Lambda_k) - K(\Lambda_k^{t'}) \mid \mathcal{F}_t]| \\ &\leq \mathbb{E}[C|\mathcal{Q}(\Lambda_k) - \mathcal{Q}(\Lambda_k^{t'})| \mid \mathcal{F}_t]. \end{aligned}$$

Jensen's inequality yields

$$\begin{aligned} \|\mathbb{E}[C|\mathcal{Q}(\Lambda_k) - \mathcal{Q}(\Lambda_k^{t'})| \mid \mathcal{F}_t]\|_2 &\leq (\|\mathbb{E}[C^2(\mathcal{Q}(\Lambda_k) - \mathcal{Q}(\Lambda_k^{t'}))^2 \mid \mathcal{F}_t]\|_1)^{1/2} \\ &= C(\mathbb{E}[(\mathcal{Q}(\Lambda_k) - \mathcal{Q}(\Lambda_k^{t'}))^2])^{1/2}. \end{aligned} \quad (5.6)$$

Now suppose that  $q \geq 0$  and (SM) holds. Then, owing to the structures of  $\Lambda_k$  and  $\Lambda_k^{t'}$ , all the terms in (5.6) cancel, except for those which contain powers of  $Y_{t,k}^{(N)}$  and  $Y_{t,k}^{(N),t'}$ . Hence, we obtain

$$\|\mathbb{E}[W_k^{(d)} \mid \mathcal{F}_t] - \mathbb{E}[W_k^{(d)} \mid \mathcal{F}_{t-1}]\| \leq C_1 \max_{1 \leq h \leq 2D} \|\alpha_j\|^h \leq C \|\alpha_j\|$$

for all  $k, j, N$ , where  $C_1 > 0$  is a fixed constant, independent of  $k, j$ , and  $N$ . The last inequality follows from the fact that we have  $|\alpha_j|^h \leq C^{h-1}|\alpha_j|$  for  $j = 1, 2, \dots$ .

Now suppose that  $q \geq 1$  and (LM) holds. Then, arguing in the same manner, we obtain

$$\begin{aligned} \|\mathbb{E}[W_k^{(d)} \mid \mathcal{F}_t] - \mathbb{E}[W_k^{(d)} \mid \mathcal{F}_{t-1}]\| &\leq C_1 \max_{1 \leq q \leq 2D} \max_{1 \leq h \leq d} \|\alpha_j - \alpha_{j+h}\|^q \\ &\leq C \max_{1 \leq h \leq d} \|\alpha_j - \alpha_{j+h}\| \end{aligned}$$

for all  $k, j, N$ , where  $C_1 > 0$  is a fixed constant, independent of  $k, j$ , and  $N$ . Hence, the claim follows.

Owing to Lemma 5.4, we obtain the following martingale difference decomposition for  $W_k^{(d)} : W_k^{(d)} = \sum_{i=0}^{\infty} \mathcal{P}_{k-i}(W_k^{(d)})$ . This allows us to use the machinery developed in [39]. Let  $g(s)$  be a Lipschitz continuous function on  $[0, 1]$ , and set

$$\Theta_k = \sum_{i=k}^{k+m-1} g\left(\frac{i}{n}\right) \{\mathbb{E}[W_i^{(d)} | \mathcal{F}_k] - \mathbb{E}[W_i^{(d)} | \mathcal{F}_{i-m}]\},$$

where  $m$  is an arbitrary positive integer (which tends to  $\infty$  later; see [39, Lemma 2]). Define the martingale

$$M_k = \sum_{i=1}^k (\Theta_i - \mathbb{E}[\Theta_i | \mathcal{F}_{i-1}]) \quad \text{and} \quad M_t^{(n)} = \frac{1}{\sqrt{n}} M_{[nt]}, \quad 0 \leq t \leq 1, \quad (5.7)$$

which is a martingale with respect to the filtration  $\mathcal{H}_t := \mathcal{F}_{[nt]}$ . One readily deduces from Lemma 5.3 that the sequence  $\{Y_k^{(N)}\}_{k \in \mathbb{Z}}$  is uniformly integrable; hence, owing to [39, Proposition 8] (see also [39, Lemma 2]), the limiting behavior of

$$\frac{1}{\sqrt{n}} S_{[nt]} = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} g\left(\frac{k}{n}\right) W_k^{(d)}$$

is governed by the martingale  $M_t^{(n)}$ . Hence, it remains to verify and apply various martingale limit theorems. We will do so in a series of lemmas. We first mention an identity, which is a key result for the proofs. For simplicity, we only state it in the one-dimensional case.

**Lemma 5.5.** *Let  $i \geq k \geq l > -\infty$  and  $j \geq l$ . Then it holds that*

$$\mathbb{E}[f(Y_i^{(N)}) | \mathcal{F}_k] \mathbb{E}[f(Y_j^{(N)}) | \mathcal{F}_l] = \mathbb{E}[f(Y_i^{(N)}) f(Y_{j \leq l}^{(N)}) + f(Y_{j > l}^{(N),'}) | \mathcal{F}_k];$$

in particular, it holds that

$$\mathbb{E}[f(Y_i^{(N)}) | \mathcal{A}^{(N)}] \mathbb{E}[f(Y_j^{(N)}) | \mathcal{A}^{(N)}] = \mathbb{E}[f(Y_i^{(N)}) f(Y_{j \leq l}^{(N),'}) | \mathcal{A}^{(N)}].$$

*Proof.* Using the coupling argument, we have

$$\begin{aligned} \mathbb{E}[f(Y_i^{(N)}) | \mathcal{F}_k] \mathbb{E}[f(Y_j^{(N)}) | \mathcal{F}_l] &= \mathbb{E}[\mathbb{E}[f(Y_j^{(N)}) | \mathcal{F}_l] f(Y_i^{(N)}) | \mathcal{F}_k] \\ &= \mathbb{E}[\mathbb{E}[f(Y_{j > l}^{(N),'}) + f(Y_{j \leq l}^{(N)}) | \mathcal{F}_l] f(Y_i^{(N)}) | \mathcal{F}_k] \\ &= \mathbb{E}[\mathbb{E}[f(Y_{j > l}^{(N),'}) + f(Y_{j \leq l}^{(N)}) f(Y_i^{(N)}) | \mathcal{F}_l] | \mathcal{F}_k] \\ &= \mathbb{E}[f(Y_i^{(N)}) f(Y_{j > l}^{(N),'}) + f(Y_{j \leq l}^{(N)}) | \mathcal{F}_k]. \end{aligned}$$

The second identity follows in the same manner.

**Lemma 5.6.** *Let  $M_t^{(n)}$  be the martingale defined by (5.7), and assume that  $g$  is a Lipschitz continuous function on  $[0, 1]$ . Then the following statements hold.*

- (i) *If  $N$  is fixed and  $n \rightarrow \infty$ , then  $[M^{(n)}, M^{(n)}]_t \xrightarrow{\mathbb{P}} \sigma_{N,m} \int_0^t g(s)^2 ds$ , where  $\sigma_{N,m}$  is an  $\mathcal{A}^{(N)}$ -measurable random variable.*
- (ii) *If both  $n$  and  $N$  tend to  $\infty$ , then  $[M^{(n)}, M^{(n)}]_t \xrightarrow{\mathbb{P}} (\sigma_{\xi^m}^f)^2 \int_0^t g(s)^2 ds$ , where  $(\sigma_{\xi^m}^f)^2 \in \mathbb{R}^+$ , depending only on  $m$ .*



*Proof.* We have  $[M^{(n)}, M^{(n)}]_t = [M, M]_{[nt]}n^{-1} = n^{-1}(M_0^2 + \sum_{k=0}^{[nt]}(M_{k+1} - M_k)^2)$ ; hence, we need to take a closer look at  $\Delta M_k := M_k - M_{k-1} = \Theta_k - \mathbb{E}[\Theta_k \mid \mathcal{F}_{k-1}]$ . For fixed  $n$  and  $N$ , we have

$$\begin{aligned}\Delta M_k &= g\left(\frac{k}{n}\right)\left(\sum_{i=k}^{k+m-1} \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_k] - \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_{k-1}]\right) \\ &\quad + \sum_{i=k}^{k+m-1} \left(g\left(\frac{i}{n}\right) - g\left(\frac{k}{n}\right)\right)(\mathbb{E}[W_i^{(d)} \mid \mathcal{F}_k] - \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_{k-1}]) \\ &=: A_{n,k} + B_{n,k}.\end{aligned}$$

Owing to the Lipschitz continuity of  $g(s)$  and Lemma 5.4, we readily verify that  $\|B_{n,k}\| = \mathcal{O}(mn^{-1})$ . Hence, we can replace  $(\Delta M_k)^2$  with  $A_{n,k}$ ; thus,

$$n[M^{(n)}, M^{(n)}]_t = M_0^2 + \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right)^2 \left(\sum_{i=k}^{k+m-1} \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_k] - \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_{k-1}]\right)^2 + \mathcal{O}(m).$$

(i) Put  $\hat{\eta}_{k,m}^2 = \{\sum_{i=k}^{k+m-1} \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_k] - \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_{k-1}]\}^2$ . It follows from the ergodic theorem that  $n^{-1}(M_0^2 + \sum_{k=0}^{[nt]}(\hat{\eta}_{k,m}^2)) \rightarrow t\sigma_{N,m}$  a.s. Hence, arguing as in [39, Theorem 9], the claim follows.

(ii) The proof of (ii) is more involved. Using the results from Section 5.1 together with Lemmas 5.1 and 5.5 we will show that  $\|[M^{(n)}, M^{(n)}]_t - (\sigma_{\xi,m}^f)^2 \int_0^t g(s)^2 ds\|_2 \rightarrow 0$ , as both  $n$  and  $N$  tend to  $\infty$ , which implies (ii). To this end, put

$$\mathcal{K}^{\mathcal{A}^{(N)}}(Y_{k,\leq t}, \dots, Y_{k+d-1,\leq t}, Y_{k,<t}, \dots, Y_{k+d-1,<t}) := \mathbb{E}[W_k^{(d)} \mid \mathcal{F}_t] - \mathbb{E}[W_k^{(d)} \mid \mathcal{F}_{t-1}],$$

and note that

$$\begin{aligned}\mathcal{K}^{\mathcal{A}^{(N)}}(x_1, \dots, x_{2d}) &= \mathbb{E}[K(Y_{k,>t}^{(N)} + x_1, \dots, Y_{k+d-1,>t}^{(N)} + x_d) \\ &\quad - K(Y_{k,>t-1}^{(N)} + x_{d+1}, \dots, Y_{k+d-1,>t-1}^{(N)} + x_{2d}) \mid \mathcal{A}^{(N)}].\end{aligned}$$

Now let

$$\begin{aligned}&\left(\sum_{i=k}^{k+m-1} \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_k] - \mathbb{E}[W_i^{(d)} \mid \mathcal{F}_{k-1}]\right)^2 \\ &= \left(\sum_{i=k}^{k+m-1} \mathcal{K}^{\mathcal{A}^{(N)}}(\cdot)\right)^2 \\ &=: \mathcal{H}^{\mathcal{A}^{(N)}}(Y_{k,\leq k}, \dots, Y_{k+d-1,\leq k}, Y_{k,<k}, \dots, Y_{k+d-1,<k}, Y_{k+1,\leq k}, \dots, Y_{k+d,\leq k}, \dots) \\ &=: \mathcal{H}_k^{\mathcal{A}^{(N)}},\end{aligned}$$

and note that  $\mathcal{H}^{\mathcal{A}^{(N)}}(\cdot)$  is a map from  $\mathbb{R}^{2dm} \rightarrow \mathbb{R}$  for fixed  $(\alpha_k^{(j)}, k \in \mathbb{N}, 1 \leq j \leq N)$ . Moreover, it holds that

$$\begin{aligned}I(n, t) &:= \mathbb{E}\left[\left(\sum_{k=1}^{[nt]} g\left(\frac{k}{n}\right)^2 \{\hat{\eta}_{k,m}^2 - \mathbb{E}[\hat{\eta}_{k,m}^2]\}\right)^2\right] \\ &= \sum_{k,l=1}^{[nt]} \left\{g\left(\frac{k}{n}\right)g\left(\frac{l}{n}\right)\mathbb{E}[\mathcal{H}_k^{\mathcal{A}^{(N)}}\mathcal{H}_l^{\mathcal{A}^{(N)}}] - \mathbb{E}[\mathcal{H}_k^{\mathcal{A}^{(N)}}]\mathbb{E}[\mathcal{H}_l^{\mathcal{A}^{(N)}}]\right\}.\end{aligned}$$

By squaring out  $\mathcal{H}_k^{\mathcal{A}^{(N)}}$  and  $\mathcal{H}_l^{\mathcal{A}^{(N)}}$ , and applying Lemma 5.5 multiple times, we have

$$\mathbb{E}[\mathcal{H}_k^{\mathcal{A}^{(N)}} \mathcal{H}_l^{\mathcal{A}^{(N)}}] = \mathbb{E}[H_k H_l], \quad \mathbb{E}[\mathcal{H}_k^{\mathcal{A}^{(N)}}] \mathbb{E}[\mathcal{H}_l^{\mathcal{A}^{(N)}}] = \mathbb{E}[H_k] \mathbb{E}[H_l],$$

where  $H_k = H(Y_{k, \leq k}, \dots)$  and  $H$  is a map from  $\mathbb{R}^{6dm} \rightarrow \mathbb{R}$ , independent of  $k$  and  $N$ . Comparing this to the function  $\mathcal{H}^{\mathcal{A}^{(N)}}$ , note that the dimension has tripled; however, the unpleasant conditioning on the  $\sigma$ -algebra  $\mathcal{A}^{(N)}$  has disappeared. In addition, using the identity  $a^2 - b^2 = (a - b)(a + b)$  and arguing similarly as in the proof of Lemma 5.4, we obtain

$$|H(x) - H(y)| \leq |x - y| R(x, y), \quad \sup_x \frac{|H(x)|}{1 + |x|^l} < C m^2, \quad x, y \in \mathbb{R}^{6dm},$$

for  $l = 2 \deg Q$ , where  $R(x, y)$  is a polynomial. Note, in particular, that we also have  $|\int_{\mathbb{R}^{2dm}} \omega_H^{\mathcal{C}_N} d\mathcal{N}| \leq C m^2 \delta_N$ . Thus, proceeding as in the proof of Proposition 5.1 we can rewrite  $I(n, t)$  as

$$\begin{aligned} & \frac{1}{n^2} \sum_{k,l=1}^n g\left(\frac{k}{n}\right) g\left(\frac{l}{n}\right) \{\mathbb{E}[H(\xi_k) H(\xi_l)] - \mathbb{E}[H(\xi_k)] \mathbb{E}[H(\xi_l)]\} + R_{N,m} \\ &= \mathbb{E} \left[ \left( \frac{1}{n} \sum_{k=1}^{[nt]} g\left(\frac{k}{n}\right) \{H(\xi_k) - (\sigma_{\xi_m}^f)^2\} \right)^2 \right] + R_{N,m}, \end{aligned}$$

where  $\mathbb{E}[H(\xi_0)^2] \leq C m^4$ ,  $\xi_k = (\xi_k^{(1)}, \dots, \xi_k^{(6dm)})$  is a  $6dm$ -dimensional stationary Gaussian sequence with covariance matrix  $r_k$  satisfying the conditions of Lemma 5.1 (or those of Lemma 5.2, which are more general),  $\sigma_{\xi_m}^f$  is defined as  $(\sigma_{\xi_m}^f)^2 = \mathbb{E}(H(\xi_0))$ , and  $R_{N,m} = O(N^{-1/2} m^{D+3})$ . It now follows from Lemma 5.1 (or Lemma 5.2) that

$$\mathbb{E} \left[ \left( \frac{1}{n} \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right) \{H(\xi_k) - (\sigma_{\xi_m}^f)^2\} \right)^2 \right] \leq \frac{C(2dm)^2}{n^{\beta/2}} \mathbb{E}[H(\xi_0)^2] \leq \frac{C}{n^{\beta/2}} m^6$$

for some  $\beta > 0$ , and, since  $n^{-1} \sum_{k=0}^{[nt]} g(k/n)^2 \rightarrow \int_0^t g(s)^2 ds$ , the claim follows.

**Lemma 5.7.** *If  $g$  is a Lipschitz continuous function on  $[0, 1]$  then, under the same assumptions as in Theorem 3.2, we have*

$$\mathbb{E} \left[ \sum_{k \leq Tn} (n^{-1/2} \Delta M_k)^2 \mathbf{1}_{\{|n^{-1/2} \Delta M_k| > \varepsilon\}} \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty \text{ for all fixed } \varepsilon > 0, T \geq 0.$$

*Proof.* For fixed  $n$  and  $N$ , we have, from the Lipschitz continuity of  $g$ ,

$$\begin{aligned} |\Delta M_k| &\leq \sum_{i=k}^{k+m-1} \left( \left| g\left(\frac{k}{n}\right) \right| + \frac{C|i-k|}{n} \right) |\mathbb{E}[W_i^{(d)} | \mathcal{F}_k] - \mathbb{E}[W_i^{(d)} | \mathcal{F}_{k-1}]| \\ &\leq C \sum_{i=k}^{k+m-1} |\mathbb{E}[W_i^{(d)} | \mathcal{F}_k] - \mathbb{E}[W_i^{(d)} | \mathcal{F}_{k-1}]| \\ &=: \Omega_{k,m}. \end{aligned}$$

Note that  $\Omega_{k,m}$  is stationary for fixed  $m$  and  $N$ , and that  $\mathbb{E}[\Omega_{k,m}^4] \leq C_m$  for all  $N$ , since we have  $\mathbb{E}[\varepsilon^D] < \infty$ . We define  $V(\delta)_T = \sum_{i=0}^T \Omega_{i,m}^2 \mathbf{1}_{\{|\Omega_{i,m}| > \delta\}}$ ,  $T \in \mathbb{N}$ , which is an

additive process in  $T$  for fixed  $m$  and  $N$ , so  $\mathbb{E}[V(\delta)_T] = T\mathbb{E}[V(\delta)_1]$  for  $T \in \mathbb{N}$ . We have, for  $\delta = \varepsilon\sqrt{n}$ ,  $n^{-1}\mathbb{E}[\sum_{i \leq Tn} (\Omega_{i,m})^2 \mathbf{1}_{\{|\Omega_{i,m}| > \delta\}}] = n^{-1}\mathbb{E}[V(\varepsilon\sqrt{n})_{Tn}] = T\mathbb{E}[V(\varepsilon\sqrt{n})_1]$ . The Cauchy–Schwarz inequality together with the Markov inequality gives

$$\mathbb{E}[V(\varepsilon\sqrt{n})_1] \leq (\varepsilon^2 n)^{-1} \sum_{i=0}^T \mathbb{E}[\Omega_{i,m}^4] = \frac{T}{\varepsilon^2 n} \mathbb{E}[\Omega_{i,m}^4].$$

Hence, it follows that  $V(\varepsilon\sqrt{n})_{Tn} \rightarrow 0$  as  $n \rightarrow \infty$ , and, since  $\Omega_{k,m} \geq \Delta M_k$ , we conclude that  $\mathbb{E}[\sum_{k \leq Tn} (n^{-1/2} \Delta M_k)^2 \mathbf{1}_{\{n^{-1/2} |\Delta M_k| > \varepsilon\}}] \rightarrow 0$  as  $n \rightarrow \infty$  for all  $\varepsilon > 0$ .

**Lemma 5.8.** *Let  $f: \mathbb{R}^d \mapsto \mathbb{R}$  be a continuous function. Then there exists a sequence  $N = N(n)$  with  $\lim_n n/N(n) = 0$  such that*

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n \{f(Y_k^{(N)}), \dots, Y_{k+d-1}^{(N)}\} - \mathbb{E}[f(Y_k^{(N)}), \dots, Y_{k+d-1}^{(N)})] \xrightarrow{D} \mathcal{N}(0, \sigma^2),$$

where  $\sigma = \sigma_\xi^f$  and  $\sigma_\xi^f$  is defined in (2.2).

*Proof.* Let  $\mathcal{L}(\cdot, \cdot)$  denote the Lévy distance, i.e.

$$\mathcal{L}(X, Y) = \inf\{\varepsilon > 0: F(x) \leq G(x + \varepsilon) + \varepsilon \text{ and } G(x) \leq F(x + \varepsilon) + \varepsilon \text{ for all } x\},$$

where  $F$  and  $G$  are distribution functions of the RVs  $X$  and  $Y$ . Define  $f_{k,N} = f(Y_k^{(N)}, \dots, Y_{k+d-1}^{(N)}) - \mathbb{E}[f(Y_k^{(N)}, \dots, Y_{k+d-1}^{(N)})]$ , and let  $f_{k,\xi}$  be the corresponding associated Gaussian version. Then we have, owing to [4, Theorem 2], for some  $\sigma^2 \geq 0$ ,  $\mathcal{L}[n^{-1/2} \sum_{k=1}^n f_{k,\xi}, \mathcal{N}(0, \sigma^2)] \leq \varepsilon_n$ , with  $\lim_n \varepsilon_n \rightarrow 0$ . For each  $n$ , according to the multivariate central limit theorem, we can choose an  $N(n)$  such that  $\mathcal{L}[n^{-1/2} \sum_{k=1}^n f_{k,N}, n^{-1/2} \sum_{k=1}^n f_{k,\xi}] \leq \varepsilon_{Nn}$ , with  $\lim \varepsilon_{N(n)} \rightarrow 0$ , and  $\lim_n n/N(n) = 0$ . Since

$$\begin{aligned} & \mathcal{L}\left[\frac{1}{\sqrt{n}} \sum_{k=1}^n f_{k,N}, \mathcal{N}(0, \sigma^2)\right] \\ & \leq \mathcal{L}\left[\frac{1}{\sqrt{n}} \sum_{k=1}^n f_{k,\xi}, \mathcal{N}(0, \sigma^2)\right] + \mathcal{L}\left[\frac{1}{\sqrt{n}} \sum_{k=1}^n f_{k,N}, \frac{1}{\sqrt{n}} \sum_{k=1}^n f_{k,\xi}\right], \end{aligned}$$

the claim follows.

*Proof of Theorem 3.2.* Owing to Lemmas 5.6 and 5.7, we can apply [26, Theorem VIII.5.9] (see also [26, Lemma 3.86]) to the martingale  $M_t^n$ , which gives  $M_t^n \xrightarrow{w} \sigma_{m,N} \int_0^t g(s) dW_s$  in  $\mathbb{D}[0, 1]$  (for fixed  $N$  and  $n \rightarrow \infty$ ), and  $M_t^n \xrightarrow{w} \sigma_{\xi m}^f \int_0^t g(s) dW_s$  in  $\mathbb{D}[0, 1]$  (if both  $n$  and  $N$  tend to  $\infty$ ) for all  $m = 1, 2, 3, \dots$ , where  $\sigma_{\xi m}^f$  is defined as in Lemma 5.6. Now [39, Lemma 2] yields  $n^{-1/2} S_{[nt]} \xrightarrow{w} \sigma_m \int_0^t g(s) dW_s$  for fixed  $N$  and  $n \rightarrow \infty$ , and  $n^{-1/2} S_{[nt]} \xrightarrow{w} \sigma_{\xi,*}^f \int_0^t g(s) dW_s$  if both  $n$  and  $N$  tend to  $\infty$ .

*Proof of Theorem 3.1.* We can proceed exactly as in the proof of Theorem 3.2, with the exception that we need to compute  $\sigma_{\xi f}^f$  in this particular case. However, owing to Lemma 5.6, it suffices to compute the value of  $\sigma_{\xi}^f$  for a special sequence  $N_0(n)$ . Let  $N(n)$  be such that  $\lim_n n/N(n)^2 = 0$ . Then according to Corollary 3.2, we can replace the conditional expectation  $\mathbb{E}[f(Y_k^{(N)}) | \mathcal{A}^{(N)}]$  with the expectation  $\mathbb{E}[f(Y_k^{(N)})]$ , and now, from Lemma 5.8 in connection with [4, Theorem 2], it holds that, for some sequence  $N_0(n)$ ,  $\sigma_{\xi}^f$  is as in Definition 2.2.

*Proof of Proposition 3.1.* We only prove the claim in a simple case, the general case can be handled in the same manner, but involves spacious computations. Let  $\deg(f)$  be the degree of the one-dimensional polynomial  $f(x) = x^{\deg(f)}$ , and put  $m = \lfloor \deg(f)/2 \rfloor$ . Let us first assume that  $\deg(f)$  is even. Then it follows that

$$\begin{aligned} & \mathbb{E}[f(Y_k^{(N)}) \mid \mathcal{A}^{(N)}] \\ &= N^{-m}(2m-1)!! \sum_{1 \leq j_1 \neq j_2 \neq \dots \neq j_m \leq N} \mathbb{E}[(X_k^{(j_1)})^2 (X_k^{(j_2)})^2 \dots (X_k^{(j_m)})^2 \mid \mathcal{A}^{(N)}] \\ & \quad + O_p(N^{-1}) \\ &= N^{-m}(2m-1)!! \sum_{1 \leq j_1 \neq j_2 \neq \dots \neq j_m \leq N} a_0^{(j_1)} a_0^{(j_2)} \dots a_0^{(j_m)} + O_p(N^{-1}), \end{aligned}$$

where  $a_h^{(r)} = \mathbb{E}[\varepsilon_1^2] \sum_{i=0}^{\infty} \alpha_i^{(r)} \alpha_{i+h}^{(r)}$ ,  $r = 1, 2, \dots, h = 0, 1, \dots, d-1$ . Hence, we obtain

$$\sqrt{N}(\mathbb{E}[f(Y_k^{(N)}) \mid \mathcal{A}^{(N)}] - \mathbb{E}[f(Y_k^{(N)})]) = \sqrt{N}(m-1)!! U_N + O_p(N^{-1/2}),$$

where

$$U_N = \frac{1}{\binom{N}{m}} \sum_{1 \leq j_1 < j_2 < \dots < j_m \leq N} a_0^{(j_1)} \dots a_0^{(j_m)} - \mathbb{E}[a_0^{(j_1)} \dots a_0^{(j_m)}],$$

which is an example of a  $U$ -statistic. The theory of  $U$ -statistics (see, e.g. [23]) yields  $\sqrt{N}U_N \xrightarrow{w} \mathcal{N}(0, \sigma^2)$  with  $\sigma^2 = m^2 \mathbb{E}^{2m-2}[a_0^1] \text{var}[a_0^1]$ . On the other hand, by appropriately adding and subtracting  $\mathbb{E}[a_0^{(j_i)}] = \mathbb{E}[a_0^{(1)}]$ ,  $1 \leq i \leq m$ , we obtain

$$\begin{aligned} & N^{-m} \sum_{1 \leq j_1 \neq j_2 \neq \dots \neq j_m \leq N} \prod_{i=1}^m (a_0^{(j_i)} - \mathbb{E}[a_0^{(j_i)}] + \mathbb{E}[a_0^{(j_i)}]) \\ &= N^{-1} m \sum_{j=1}^N (a_0^{(j)} - \mathbb{E}[a_0^{(j)}]) \mathbb{E}^{m-1}[a_0^{(1)}] + \mathbb{E}^m[a_0^{(1)}] + O_p(N^{-1}). \end{aligned}$$

The central limit theorem now also yields the claim. Now consider the case where  $\deg(f)$  is odd. Then it follows that

$$\begin{aligned} & \mathbb{E}[f(Y_k^{(N)}) \mid \mathcal{A}^{(N)}] - O_p(N^{-1}) \\ &= N^{-m-1/2} 2m(2m-2) \dots \sum_{\substack{1 \leq j_1 \neq j_2 \neq \dots \neq j_m \leq N, \\ j_1 = j_{m+1}, \dots, j_m = j_{m+1}}} a_0^{(j_1, j_{m+1})} a_0^{(j_2, j_{m+1})} \dots a_0^{(j_m, j_{m+1})}, \end{aligned}$$

where  $a_0^{(r,s)} = \mathbb{E}(\varepsilon_1^2) \sum_{i=0}^{\infty} (\alpha_i^{(r)})^2$  if  $r \neq s$  and  $a_0^{(r,r)} = a_0^{(r,r)} = \mathbb{E}(\varepsilon_1^3) \sum_{i=0}^{\infty} (\alpha_i^{(r)})^3$  if  $r = s$ . Arguing as in the even case, we obtain

$$\sqrt{N}(\mathbb{E}[f(Y_k^{(N)}) \mid \mathcal{A}^{(N)}] - \mathbb{E}[f(Y_k^{(N)})]) = O_p(N^{-1/2}),$$

which completes the proof.

*Proof of Corollary 3.1.* Statement (i) follows immediately from  $\lim_{n \rightarrow \infty} n/N(n) = 0$  and  $\sqrt{N}(\mathbb{E}[f(\mathbf{Y}_k^{(N)}) \mid \mathcal{A}^{(N)}] - \mathbb{E}[f(\mathbf{Y}_k^{(N)})]) \xrightarrow{w} \mathcal{N}(0, \sigma^2)$ . To prove (ii), observe that

$$\begin{aligned} & \frac{\sqrt{N}}{n} \sum_{i=0}^{[nt]} g\left(\frac{i}{n}\right) (f(\mathbf{Y}_k^{(N)}) - \mathbb{E}[f(\mathbf{Y}_k^{(N)})]) \\ &= \sqrt{\frac{N}{n}} A_{nt,N} + \sqrt{N} (\mathbb{E}[f(\mathbf{Y}_{0,d}^{(N)}) \mid \mathcal{A}^{(N)}] - \mathbb{E}[f(\mathbf{Y}_{0,d}^{(N)})]) \frac{1}{n} \sum_{i=0}^{[nt]} g\left(\frac{i}{n}\right), \end{aligned}$$

where

$$A_{nt,N} = n^{-1/2} \sum_{i=0}^{[nt]} g\left(\frac{i}{n}\right) (f(\mathbf{Y}_k^{(N)}) - \mathbb{E}[f(\mathbf{Y}_k^{(N)}) \mid \mathcal{A}^{(N)}]) \xrightarrow{w} \sigma_\xi^f \int_0^t g(s) dW_s.$$

Hence, it follows that  $\sqrt{N/n} A_{nt,N} \xrightarrow{\mathbb{P}} 0$ , and, thus, the limiting behavior is determined by  $\sqrt{N}(\mathbb{E}[f(\mathbf{Y}_{0,d}^{(N)}) \mid \mathcal{A}^{(N)}] - \mathbb{E}[f(\mathbf{Y}_{0,d}^{(N)})]) n^{-1} \sum_{i=0}^{[nt]} g(i/n)$ . In view of  $n^{-1} \sum_{i=0}^{[nt]} g(i/n) \rightarrow \int_0^t g(s) ds$ , the claim follows.

*Proof of Corollary 3.2.* Owing to Lemma 5.3, we have  $\sqrt{n}|\mathbb{E}[f(\mathbf{Y}_k^{(N)})] - \mathbb{E}[f(\xi_k)]| \leq C\sqrt{n/N}$ ; hence, it suffices to establish the claim for  $\mathbb{E}[f(\mathbf{Y}_k^{(N)})]$ . Since  $g(s)$  is bounded, the Minkowski and Jensen inequalities yield

$$\begin{aligned} & \mathbb{E} \left| \sum_{k=0}^{[nt]} g\left(\frac{i}{n}\right) [\mathbb{E}[f(\mathbf{Y}_k^{(N)})] - \mathbb{E}[f(\mathbf{Y}_k^{(N)}) \mid \mathcal{A}^{(N)}]] \right| \\ & \leq C \sum_{i=0}^{[nt]} \|\mathbb{E}[f(\mathbf{Y}_k^{(N)})] - \mathbb{E}[f(\mathbf{Y}_k^{(N)}) \mid \mathcal{A}^{(N)}]\|. \end{aligned}$$

We will now again use a coupling argument. Let  $\{\zeta_k^{(j)}\}_{k \geq 1}$  be an independent copy of  $\{\varepsilon_k^{(j)}\}_{k \geq 1}$ , and define  $\mathbf{Y}_k^{(N),*} = (Y_k^{(N),*}, \dots, Y_{k+d-1}^{(N),*})^\top$  by replacing  $\varepsilon_k^{(j)}$  with  $\zeta_k^{(j)}$  in  $\mathbf{Y}_k^{(N)} = (Y_k^{(N)}, \dots, Y_{k+d-1}^{(N)})^\top$  (the  $\alpha_i^{(j)}$ s remain the same). Then we have

$$\begin{aligned} \mathbb{E}[\mathbb{E}^2[f(\mathbf{Y}_k^{(N)}) \mid \mathcal{A}]] &= \mathbb{E}[\mathbb{E}[f(\mathbf{Y}_k^{(N)})] \mathbb{E}[f(\mathbf{Y}_k^{(N)}) \mid \mathcal{A}^{(N)}] \mid \mathcal{A}^{(N)}]] \\ &= \mathbb{E}[\mathbb{E}[f(\mathbf{Y}_k^{(N)}) f(\mathbf{Y}_k^{(N),*}) \mid \mathcal{A}]] \\ &= \mathbb{E}[f(\mathbf{Y}_k^{(N)}) f(\mathbf{Y}_k^{(N),*})]. \end{aligned}$$

Let  $\{\xi_k, \xi_k^*\}_{k \geq 1}$  be the (two-dimensional) associated processes of  $\{\mathbf{Y}_k^{(N)}, \mathbf{Y}_k^{(N),*}\}_{k \geq 1}$ , and put  $\xi_k = (\xi_k, \dots, \xi_{k+d-1})^\top$  and  $\xi_k^* = (\xi_k^*, \dots, \xi_{k+d-1}^*)^\top$ . The conditions of Theorem 3.2 imply that the assumptions of Lemma 5.3 are met; hence, arguing similarly as in Proposition 5.1, we obtain  $|\mathbb{E}[f(\mathbf{Y}_k^{(N)}) f(\mathbf{Y}_k^{(N),*})] - \mathbb{E}[f(\xi_k) f(\xi_k^*)]| \leq C/\sqrt{N}$ , where  $C_2$  does not depend on  $n$ . In addition, since  $\mathbb{E}[Y_k^{(N)} Y_k^{(N),*}] = 0$  for all  $k \in \mathbb{Z}$ , we can assume that  $\xi_k$  and  $\xi_k^*$  are independent, and, thus, it holds that

$$\mathbb{E}[f(\xi_k) f(\xi_k^*)] = \mathbb{E}[f(\xi_k)] \mathbb{E}[f(\xi_k^*)] = \mathbb{E}^2[f(\xi_k)].$$

In the same manner, we obtain  $|\mathbb{E}^2[f(\mathbf{Y}_k^{(N)})] - \mathbb{E}^2[f(\xi_k)]| \leq C/\sqrt{N}$ . Putting everything together, we obtain  $\|\mathbb{E}[f(\mathbf{Y}_k^{(N)})] - \mathbb{E}[f(\mathbf{Y}_k^{(N)}) \mid \mathcal{A}^{(N)}]\|_2 \leq CN^{-1/4}$ , and, hence, it follows

that  $n^{-1/2} \sum_{k=1}^{[nt]} \|\mathbb{E}[f(Y_k^{(N)})] - \mathbb{E}[f(Y_k^{(N)}) | \mathcal{A}]\|_2 = O(\sqrt{n}/N^{1/4})$ , which completes the proof.

*Proof of Theorem 3.3.* Owing to  $\|f\|_p^* < \infty$ , we have  $\mathbb{E}[f(\xi_{0,d})^2] < \infty$  and, thus,  $f$  can be represented as an infinite sum of Hermite polynomials. Let  $f_p$  be the sum of the first  $p$  polynomials. Then we have, after dropping the variables of  $f$ ,  $f_p$ , and  $f_{>p}$  for the simplicity of the formulae,

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right) (f - \mathbb{E}[f]) \\ &= \frac{1}{\sqrt{n}} \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right) (f_p - \mathbb{E}[f_p]) + \frac{1}{\sqrt{n}} \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right) (f_{>p} - \mathbb{E}[f_{>p}]) \\ &= A_{p,n}(t) + A_{>p,n}(t). \end{aligned}$$

Since  $f_p$  is a polynomial, it follows from Theorem 3.2 together with Corollary 3.1 that

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{[nt]} g\left(\frac{k}{n}\right) (f_p - \mathbb{E}[f_p]) \xrightarrow{w} \sigma_{\xi}^{f_p} \int_0^t g(s) dW_s.$$

Moreover, Corollary 5.1 (which remains valid since  $g$  is bounded) yields

$$\|A_{p,n}(t) - A_{\infty,n}(t)\|_2 = \|A_{>p,n}(t)\|_2 \leq \varepsilon_p^{(1)} + \varepsilon_{N,n,p}^{(2)},$$

where  $\varepsilon_p^{(1)} \rightarrow 0$  and  $\varepsilon_{N,n,p}^{(2)} \rightarrow 0$  if  $C(2d, 2p)n/\sqrt{N} \rightarrow 0$ , and  $C(2d, 2p)$  is as in Proposition 5.1. Thus, by Slutsky's lemma, it suffices to treat  $A_{p,n}(t)$ . It follows from [4, Theorem 2] that  $\sigma_{\xi}^{f_p} \rightarrow \sigma_{\xi}^f$ , where  $\sigma_{\xi}^{f_p}$  and  $\sigma_{\xi}^f$  are defined in (2.2). Now let  $h$  be a continuous, bounded function. Then the dominated convergence theorem implies that

$$\left| \mathbb{E} \left[ h \left( \sigma_{\xi}^{f_p} \int_0^t g(s) dW_s \right) \right] - \mathbb{E} \left[ h \left( \sigma_{\xi}^f \int_0^t g(s) dW_s \right) \right] \right| \leq \varepsilon_p^{(2)},$$

with  $\varepsilon_p^{(2)} \rightarrow 0$ . Thus, it follows from the Minkowski inequality and Theorem 3.2 that

$$\begin{aligned} & \left| \mathbb{E}[h(A_{p,n})] - \mathbb{E} \left[ h \left( \sigma_{\xi}^f \int_0^t g(s) dW_s \right) \right] \right| \\ & \leq \left| \mathbb{E} \left[ h \left( \sigma_{\xi}^{f_p} \int_0^t g(s) dW_s \right) \right] - \mathbb{E} \left[ h \left( \sigma_{\xi}^f \int_0^t g(s) dW_s \right) \right] \right| \\ & \quad + \left| \mathbb{E}[h(A_{p,n})] - \mathbb{E} \left[ h \left( \sigma_{\xi}^{f_p} \int_0^t g(s) dW_s \right) \right] \right| \\ & \leq \varepsilon_p^{(2)} + \varepsilon_{N,n,p}^{(3)} + \varepsilon_{n,p}^{(4)}, \end{aligned}$$

where  $\lim_n \varepsilon_{n,p}^{(4)} \rightarrow 0$  for each  $p$ . In view of  $n/\sqrt{N} \rightarrow 0$  we can choose a sequence  $p \rightarrow \infty$  such that  $C(2d, 2p)n/\sqrt{N} \rightarrow 0$ , which establishes the claim in dimension one. The general case follows in the same manner in connection with the Cramér–Wold device.

*Proof of Corollary 3.4.* In order to establish the convergence of the finite-dimensional distributions, we can repeat the proof of Theorem 3.3 verbatim. However, since we no longer require any approximation results for aggregated processes, we only need  $\mathbb{E}[f(\xi_0)^2] < \infty$  instead of  $\mathbb{E}[f(\xi_0)^4] < \infty$ ; this is especially the case in Theorem 3.2, in particular in Lemmas 5.6 and 5.7.

### 5.3. Proofs of Section 4

*Proof of Proposition 4.1.* For the proof, it suffices to verify the conditions of Theorem 3.2. We have  $\deg(f) = 2$ ; hence, we obtain  $D = 8$ . Since  $|a| \leq 1$  a.s. per Assumption 4.1, we obtain

$$\sum_{i=0}^{\infty} \|\alpha_i - \alpha_{i+2}\| \leq C_1 \sup_{h \in \mathbb{N}_0} \sum_{i=0}^{\infty} \|a^{i+h}(1-a^2)\| \leq C_2 \sum_{i=0}^{\infty} \|a^i(1-a^2)\|,$$

which is finite due to Assumption 4.1. Finally, it is obvious that  $f$  may be written as in (3.1) with  $L = 2$  and  $\min_{1 \leq i \leq L} q_i = 1$ .

**Remark 5.1.** If the function  $f$  is as in (4.1), we may extend Proposition 3.1 (by mimicking the proof) to the case where (LM) is valid instead of (ii).

*Proof of Corollary 4.1.* Proceeding as in the proof of Corollary 3.1, we obtain, due to Remark 5.1,

$$\sqrt{N}(Z_{n,N,h} - \mathbb{E}[a^h]) \xrightarrow{D} \mathcal{N}(0, \text{var}(a)).$$

This implies in particular that  $\sqrt{Z_{n,N,2} - Z_{n,N,1}^2}$  is a consistent estimator for  $\text{var}(a)$ ; hence, the claim follows from Slutsky's lemma.

*Proof of Corollary 4.2.* As in the proof of Corollary 4.1, we may proceed as in the proof of Corollary 3.1, using Remark 5.1.

### Acknowledgement

The author thanks the anonymous referee for valuable remarks and comments that helped to improve the presentation and results of this paper.

### References

- [1] ANH, V. V., HEYDE, C. C. AND LEONENKO, N. N. (2002). Dynamic models of long-memory processes driven by Lévy noise. *J. Appl. Prob.* **39**, 730–747.
- [2] ANH, V. V., KNOPOVA, V. P. AND LEONENKO, N. N. (2004). Continuous-time stochastic processes with cyclical long-range dependence. *Austral. N. Z. J. Statist.* **46**, 275–296.
- [3] ANH, V. V., LEONENKO, N. N. AND McVINISH, R. (2001). Models for fractional Riesz-Bessel motion and related processes. *Fractals* **9**, 329–346.
- [4] ARCONES, M. A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Prob.* **22**, 2242–2274.
- [5] ARCONES, M. A. (2000). Distributional limit theorems over a stationary Gaussian sequence of random vectors. *Stoch. Process. Appl.* **88**, 135–159.
- [6] BARNDORFF-NIELSEN, O. E. (2000). Superposition of Ornstein-Uhlenbeck type processes. *Teor. Veroyat. Primen.* **45**, 289–311. English translation: *Theory Prob. Appl.* **45** (2001), 175–194.
- [7] BARNDORFF-NIELSEN, O. E. AND LEONENKO, N. N. (2005). Spectral properties of superpositions of Ornstein-Uhlenbeck type processes. *Methodology Comput. Appl. Prob.* **7**, 335–352.
- [8] BERAN, J. AND OCKER, D. (2000). Temporal aggregation of stationary and nonstationary farima (p, d, 0) models. CoFE discussion paper, University of Konstanz.
- [9] BERAN, J., GHOSH, S. AND SCHÜTZNER, M. (2010). From short to long memory: aggregation and estimation. *Comput. Statist. Data Anal.* **54**, 2432–2442.
- [10] CELOV, D., LEIPUS, R. AND PHILIPPE, A. (2007). Time series aggregation, disaggregation, and long memory. *Liet. Mat. Rink.* **47**, 466–481. English translation: *Lithuanian Math. J.* **47**, 379–393.
- [11] CELOV, D., LEIPUS, R. AND PHILIPPE, A. (2010). Asymptotic normality of the mixture density estimator in a disaggregation scheme. *J. Nonparametric Statist.* **22**, 425–442.
- [12] CHONG, T. T.-L. (2006). The polynomial aggregated AR(1) model. *Econometrics J.* **9**, 98–122.
- [13] DACUNHA-CASTELLE, D. AND FERMÍN, L. (2006). Disaggregation of long memory processes on  $\mathcal{C}^\infty$  class. *Electron. Commun. Probab.* **11**, 35–44.

- [14] DACUNHA-CASTELLE, D. AND OPPENHEIM, G. (2001). Mixtures, aggregations and long-memory. Tech. Rep. Université de Paris-Sud.
- [15] DEDECKER, J. AND DOUKHAN, P. (2003). A new covariance inequality and applications. *Stoch. Process. Appl.* **106**, 63–80.
- [16] DEDECKER, J. AND RIO, E. (2000). On the functional central limit theorem for stationary processes. *Ann. Inst. H. Poincaré Prob. Statist.* **36**, 1–34.
- [17] DING, Z. AND GRANGER, C. W. J. (1996). Modeling volatility persistence of speculative returns: a new approach. *J. Econometrics* **73**, 185–215.
- [18] DOBRUSHIN, R. L. AND MAJOR, P. (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrscheinlichkeitsth.* **50**, 27–52.
- [19] GIRAITIS, L. AND SURGAILIS, D. (1985). CLT and other limit theorems for functionals of Gaussian processes. *Z. Wahrscheinlichkeitsth.* **70**, 191–212.
- [20] GIRAITIS, L., LEIPUS, R. AND SURGAILIS, D. (2010). Aggregation of the random coefficient GLARCH(1,1) process. *Econometric Theory* **26**, 406–425.
- [21] GORDIN, M. I. (1969). The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR* **188**, 739–741.
- [22] GRANGER, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *J. Econometrics* **14**, 227–238.
- [23] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19**, 293–325.
- [24] HORVÁTH, L. AND LEIPUS, R. (2009). Effect of aggregation on estimators in AR(1) sequence. *Test* **18**, 546–567.
- [25] HURD, H. L. AND MIAMEE, A. (2007). *Periodically Correlated Random Sequences*. Wiley-Interscience, Hoboken, NJ.
- [26] JACOD, J. AND SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes* (Fundamental Principles Math. Sci. **288**), 2nd edn. Springer, Berlin.
- [27] JIRAK, M. (2011). Asymptotic behavior of weakly dependent aggregated processes. *Period. Math. Hungar.* **62**, 39–60.
- [28] KAZAKEVIČIUS, V., LEIPUS, R. AND VIANO, M.-C. (2004). Stability of random coefficient ARCH models and aggregation schemes. *J. Econometrics* **120**, 139–158.
- [29] KOUL, H. L. (1992). *Weighted Empirical and Linear Models* (IMS Lecture Notes Monogr. Ser. **21**). Institute of Mathematical Statistics, Hayward, CA.
- [30] LEIPUS, R. AND VIANO, M.-C. (2002). Aggregation in ARCH models. *Liet. Mat. Rink.* **42**, 68–89.
- [31] LEIPUS, R., OPPENHEIM, G., PHILIPPE, A. AND VIANO, M.-C. (2006). Orthogonal series density estimation in a disaggregation scheme. *J. Statist. Planning Inf.* **136**, 2547–2571.
- [32] LEONENKO, N. N. AND TAUFER, E. (2005). Convergence of integrated superpositions of Ornstein-Uhlenbeck processes to fractional Brownian motion. *Stochastics* **77**, 477–499.
- [33] LEWBEL, A. (1994). Aggregation and simple dynamics. *Amer. Econom. Rev.* **84**, 905–918.
- [34] MAXWELL, M. AND WOODROOFE, M. (2000). Central limit theorems for additive functionals of Markov chains. *Ann. Prob.* **28**, 713–724.
- [35] MERLEVÈDE, F., PELIGRAD, M. AND UTEV, S. (2006). Recent advances in invariance principles for stationary sequences. *Prob. Surveys* **3**, 1–36.
- [36] OPPENHEIM, G. AND VIANO, M.-C. (2004). Aggregation of random parameters Ornstein-Uhlenbeck or AR processes: some convergence results. *J. Time Ser. Anal.* **25**, 335–350.
- [37] PELIGRAD, M. AND UTEV, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *Ann. Prob.* **33**, 798–815.
- [38] PELIGRAD, M. AND UTEV, S. (2006). Central limit theorem for stationary linear processes. *Ann. Prob.* **34**, 1608–1622.
- [39] PELIGRAD, M. AND UTEV, S. (2006). Invariance principle for stochastic processes with short memory. In *High Dimensional Probability* (IMS Lecture Notes Monogr. Ser. **51**), Institute of Mathematical Statistics, Beachwood, OH, pp. 18–32.
- [40] PHILLIPS, P. C. B. AND MOON, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica* **67**, 1057–1111.
- [41] PUPLINSKAITĖ, D. AND SURGAILIS, D. (2009). Aggregation of random-coefficient AR(1) process with infinite variance and common innovations. *Lithuanian Math. J.* **49**, 446–463.
- [42] PUPLINSKAITĖ, D. AND SURGAILIS, D. (2010). Aggregation of a random-coefficient AR(1) process with infinite variance and idiosyncratic innovations. *Adv. Appl. Prob.* **42**, 509–527.
- [43] SHORACK, G. R. AND WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. John Wiley, New York.
- [44] SWEETING, T. J. (1977). Speeds of convergence for the multidimensional central limit theorem. *Ann. Prob.* **5**, 28–41.



- [45] TAQQU, M. S. (1974/75). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie* **31**, 287–302.
- [46] TAQQU, M. S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrscheinlichkeitstheorie* **50**, 53–83.
- [47] VOLNÝ, D., WOODROOFE, M. AND ZHAO, O. (2011). Central limit theorems for superlinear processes. *Stoch. Dyn.* **11**, 71–80.
- [48] WU, W. B. (2005). Nonlinear system theory: another look at dependence. *Proc. Natl. Acad. Sci. USA* **102**, 14150–14154.
- [49] WU, W. B. (2007). Strong invariance principles for dependent random variables. *Ann. Prob.* **35**, 2294–2320.
- [50] WU, W. B. AND WOODROOFE, M. (2004). Martingale approximations for sums of stationary processes. *Ann. Prob.* **32**, 1674–1690.
- [51] ZAFFARONI, P. (2004). Contemporaneous aggregation of linear dynamic models in large economies. *J. Econometrics* **120**, 75–102.
- [52] ZAFFARONI, P. (2007). Aggregation and memory of models of changing volatility. *J. Econometrics* **136**, 237–249.
- [53] ZAFFARONI, P. (2007). Contemporaneous aggregation of GARCH processes. *J. Time Ser. Anal.* **28**, 521–544.