Joint FCLT for Sample Quantile and Measures of Dispersion for Functionals of Mixing Processes

Marcel Bräutigam* and Marie Kratz†

* CREAR, ESSEC Business School
† ESSEC Business School, IDO dept & CREAR, and Labex MME-DII, CY initiative

Abstract

In this paper, we establish a joint (bivariate) functional central limit theorem of the sample quantile and the r-th absolute centred sample moment for functionals of mixing processes. More precisely, we consider $L_2$-near epoch dependent processes that are functionals of either $\phi$-mixing or absolutely regular processes. The general results we obtain can be used for two classes of popular and important processes in applications: The class of augmented GARCH($p,q$) processes with independent and identically distributed innovations (including many GARCH variations used in practice) and the class of ARMA($p,q$) processes with mixing innovations (including, e.g., ARMA-GARCH processes). For selected examples, we provide exact conditions on the moments and parameters of the process for the joint asymptotics to hold.

AMS classification: 60F05; 60F17; 60G10; 62H10; 62H20

Keywords: asymptotic distribution; Bahadur representation; functional central limit theorem; (near epoch) dependence; (augmented) GARCH; ARMA;

1 Introduction and framework

The focus of this paper is on the asymptotic theory of sample estimators of standard statistical quantities as rank, location and dispersion measures, for a very large class of widely used stationary processes. Such a theory is often needed for related statistical inference. The literature for sample quantiles has considerably developed (closely related to empirical processes) from the iid setting toward dependent random variables. In view of applications, it is also important to consider jointly the estimators of such standard statistical quantities. Fields of applications include dynamic systems (networks), (financial) econometrics, reliability, risk theory and management, quantitative finance, etc. For instance, dispersion measures, as e.g. the second centred moment (named volatility in financial fields), are essential in economics for studying market fluctuations. This is why they are often introduced, via conditioning or jointly, in the econometric analysis of macroeconomic and financial data, to take into account the market state. Investigations on the relation between volatility estimators (a measure of dispersion) and Value-at-Risk (a quantile estimator) for time-series models are common in this field; see e.g. [14], [27], [38], [64].

In the literature, the asymptotic behavior of each quantity has been investigated on its own, under some dependence structures. Limit theorems for sample quantiles have been proposed, inter alia, in the case
of m-dependent or \( \phi \)-mixing processes ([52], [53]), infinite-variance or non-linear processes ([30], [61]), and functionals of mixing processes ([60]). They are generally based on the use of the Bahadur representation developed in [4] for independent and identically distributed (iid) random variables (and refined by Kiefer in a sequence of papers), then extended (generally via coupling techniques) to linear and nonlinear processes, with various types of dependence (see e.g. [61] and references therein).

For dispersion measures, there exist results for the mean absolute deviation (MAD) when considering strongly mixing processes ([51]), while the asymptotics of the sample variance can be found for various settings as examples in standard textbooks, see e.g. [1], [15], [57]. When considering more generally the autocovariance function of such estimators, we can mention, exemplarily, results for linear, bilinear or non-linear processes with regularly varying noise ([18], [19]), and for long-memory time series (e.g. [35] - and for a more complete overview [39] and references therein).

In this paper, we establish the joint asymptotics of the sample quantile and the \( r \)-th absolute centred sample moment for functionals of strictly stationary processes \((Z_n)_{n\in\mathbb{Z}}\), namely

\[
X_n = f(Z_{n+k}, k \in \mathbb{Z}), \quad \text{with } f \text{ measurable},
\]

assuming \((Z_n)\) \( \phi \)-mixing or absolutely regular, respectively, and \((X_n)\) near-epoch dependent. We briefly recall these dependence notions further below. Our limit theorem extends existing results on joint asymptotics in the case of an identically and independently distributed (iid) sample (see [14]). In particular, we establish an asymptotic representation for the \( r \)-th absolute centred sample moment for stationary, ergodic, short-memory processes under minimal conditions (extending the results in [51] and [14]). Furthermore, our general results include two important popular classes of linear and nonlinear processes: The class of augmented GARCH\((p,q)\) processes with iid innovations and ARMA\((p,q)\) processes with mixing innovations. The former, introduced by Duan in [23], contains many well-known GARCH processes as special cases, of standard use, e.g. in quantitative finance and financial econometrics. Previous works on univariate CLTs and stationarity conditions for this class of GARCH processes are, among others, [3], [5], [34] and [40]. The class of ARMA processes includes the classical ARMA models with iid or white noise innovations (see e.g. [15], [12] for classical references on the topic), but also with mixing innovations. The earliest explicitly mentioned examples of ARMA processes with mixing innovations date back, to our knowledge, to [58], [59] considering ARMA models with ARCH errors. More widely used in applications is the ARMA-GARCH model, which can be seen as an extension of the GARCH models to also have dynamics in the mean and not only in the variance process. First contributions on such types of processes can be found in [42], [43], [45]; more recent examples of applications are, e.g. [17], [54], [55]. Analogously to the class of augmented GARCH processes, there exist results in the literature on FCLTs for ARMA processes with mixing innovations (or even larger classes of linear processes), as e.g. [17], [49].

Classical and broad notions for (asymptotically vanishing) dependence of a process are known under the name of mixing (see [13] for an overview). Let us recall some notions of weak dependence related to the mixing coefficients. Let \((Y_n)_{n\in\mathbb{Z}}\) be a random process and denote the corresponding \( \sigma \)-algebra as \( \mathcal{F}_{Y,s,t} = \sigma(Y_u; s \leq u \leq t) \).

**Definition 1 (Mixing coefficients)** For any integer \( n \geq 1 \), let

\[
\alpha(n) := \sup_{j \in \mathbb{Z}} \sup_{C \in \mathcal{F}_{Y,-\infty,j}, D \in \mathcal{F}_{Y,j+n,\infty}} |P(C \cap D) - P(C)P(D)|,
\]

\[
\beta(n) := \sup_{j \in \mathbb{Z}} E\left[ \sup_{C \in \mathcal{F}_{Y,-\infty,j}} |P(C|\mathcal{F}_{Y,-\infty,j}) - P(C)| \right],
\]

\[
\phi(n) := \sup_{j \in \mathbb{Z}} \sup_{C \in \mathcal{F}_{Y,-\infty,j}:P(C)>0, D \in \mathcal{F}_{Y,j+n,\infty}} |P(D|C) - P(D)|.
\]
The process \((Y_n)_{n \in \mathbb{Z}}\) is called strongly mixing if \(\alpha(n) \to 0\), \(\beta\)-mixing or absolutely regular if \(\beta(n) \to 0\), and \(\phi\)-mixing if \(\phi(n) \to 0\).

Note that, for a stationary time series, we can omit the \(\sup\) in the definition of the mixing notions and simply set, w.l.o.g., \(j = 0\). Also, \(\phi\)-mixing implies \(\beta\)-mixing that implies \(\alpha\)-mixing, but the converse does not hold in general.

A drawback of mixing is that, in general, a functional that depends on an infinite number of lags of a mixing process will not be mixing itself. This gave rise to the introduction of the notion of \(L_p\)-NED (that goes back to [37] in the case of \(p = 2\) for an underlying \(\phi\)-mixing process, even if not named NED there). It imposes additional structure on functionals of mixing processes enabling statistical inference for this larger class of processes. Alternatively, \(L_p\)-NED is also called \(p\)-approximation functional; see [9], [60].

**Definition 2 (\(L_p\)-NED, [2])** For \(p > 0\), a stationary process \((X_n)_{n \in \mathbb{Z}}\) is called \(L_p\)-NED on a process \((Z_n)_{n \in \mathbb{Z}}\) if, for \(k \geq 0\),

\[
\|X_0 - E[X_0|F_{Z,-k,+k}]\|_p \leq \nu(k),
\]

for non-negative constants \(\nu(k)\) such that \(\nu(k) \to 0\) as \(k \to \infty\).

Note that, in the econometrics literature, a specific terminology is used for the rate of \(\nu(k)\): If \(\nu(k) = O(k^{-\tau})\) for some \(\tau > 0\), one calls the process \(L_p\)-NED of size \((-\tau)\) and, if \(\nu(k) = O(e^{-\delta k})\) for some \(\delta > 0\), geometrically \(L_p\)-NED.

Finally, let us present some technical conditions on the marginal distribution function \(F_X\) of the stationary process \((X_n)_{n \in \mathbb{Z}}\) under study, and give the notation of the sample estimators for the two quantities of interest. We denote, whenever they exist, the probability density function (pdf) of \(F_X\) by \(f_X\), with mean \(\mu\), variance \(\sigma^2\), as well as, for any integer \(r \geq 1\), the \(r\)-th absolute centred moment, \(\mu(X, r) := E[|X| - \mu|^r]\). The quantile of order \(p\) of \(F_X\) is defined as \(q_X(p) := \inf\{x \in \mathbb{R} : F_X(x) \geq p\}\).

We impose four different types of conditions on the marginal distribution \(F_X\), as in the iid case (see [14]). First, we assume the existence of its finite \(2k\)-th moment for any integer \(k > 0\). Then, the continuity or \(l\)-fold differentiability of \(F_X\) (at a given point or neighbourhood) for any integer \(l > 0\), and the positivity of \(f_X\) (at a given point or neighbourhood). Those conditions are named as:

\[
(M_k) \quad E[|X_0|^{2k}] < \infty,
\]

\[
(C_0) \quad F_X \text{ is continuous},
\]

\[
(C'_l) \quad F_X \text{ is } l\text{-times differentiable},
\]

\[
(C''_l) \quad f_X \text{ is positive}.
\]

Given a sample \(X_1, \ldots, X_n\), with order statistics \(X_{(1)} \leq \cdots \leq X_{(n)}\), let \(q_n(p)\) denote the sample quantile of order \(p \in [0, 1]\), namely

\[
q_n(p) = X_{([np])},
\]

where \([x]\) = \(\min\{m \in \mathbb{Z} : m \geq x\}\) and \([x]\) are the rounded-up integer-part and the nearest-integer of a real number \(x \in \mathbb{R}\), respectively.

The \(r\)-th absolute centred sample moment is denoted by \(\hat{m}(X, n, r)\) and defined, for \(r \in \mathbb{N}\), by

\[
\hat{m}(X, n, r) := \frac{1}{n} \sum_{i=1}^{n} |X_i - \bar{X}_n|^r,
\]
Theorem 3 (bivariate FCLT) Consider a process \( X_n \) representing the empirical mean. Special cases of this latter estimator include the sample variance \((r = 2)\) and the sample mean absolute deviation around the sample mean \((r = 1)\).

Recall the standard notation \( u^T \) for the transpose of a vector \( u \) and \( \text{sgn}(x) := -\mathbb{I}_{x<0} + \mathbb{I}_{x>0} \) for the signum function. By \(|\cdot|\), we denote the euclidean norm, and the usual \( L_p \)-norm is denoted by \( \| \cdot \|_p := \mathbb{E}^{1/p}[|\cdot|^p] \). Moreover the notations \( \overset{d}{\rightarrow}, \overset{a.s.}{\rightarrow}, \overset{P}{\rightarrow} \) and \( D_{d,[0,1]} \) correspond to the convergence in distribution, almost surely, in probability and in distribution of a random vector in the d-dimensional Skorohod space \( D_d[0,1] \). Further, for real-valued functions \( f \) and \( g \), we write \( f(x) = O(g(x)) \) (as \( x \to \infty \)) if and only if there exists a positive constant \( M \) and a real number \( x_0 \) s.t. \( |f(x)| \leq Mg(x) \) for all \( x \geq x_0 \), and \( f(x) = o(g(x)) \) (as \( x \to \infty \)) if, for all \( \epsilon > 0 \), there exists a real number \( x_0 \) s.t., for all \( x \geq x_0 \), \( |f(x)| \leq \epsilon g(x) \). Analogously, for a sequence of rv’s \( X_n \) and constants \( a_n \), we denote by \( X_n = o_P(a_n) \) the convergence in probability to 0 of \( X_n/a_n \).

The structure of the paper is as follows. We present in Section 2 the main results on the bivariate FCLT for the sample quantile and the \( r \)-th absolute centred sample moment for functionals of \( \phi \)-mixing or absolutely regular processes. In Section 3, we apply our general results to the family of augmented GARCH\((p, q)\) processes and ARMA\((p, q)\) processes with mixing innovations. Section 4 gathers auxiliary results needed for the proofs of the main results, but also of interest on their own: the asymptotic representation of the \( r \)-th absolute centred sample moment and results on \( L_2 \)-NED of bounded and unbounded functionals. Finally, the proofs are given in Section 5.

2 The Bivariate FCLT

Let us state the main result. To ease its presentation, we introduce a trivariate normal random vector \((U, V, W)^T\), functional of a random process \( X \), with mean zero and the following covariance matrix:

\[
\begin{cases}
\text{Var}(U) = \text{Var}(X_0) + 2 \sum_{i=1}^{\infty} \text{Cov}(X_i, X_0) \\
\text{Var}(V) = \text{Var}(|X_0|^r) + 2 \sum_{i=1}^{\infty} \text{Cov}(|X_i|^r, |X_0|^r) \\
\text{Var}(W) = \text{Var}\left(\frac{p - \mathbb{I}_{X_0 \leq q_X(p)}}{f_X(q_X(p))}\right) + 2 \sum_{i=1}^{\infty} \text{Cov}\left(\frac{p - \mathbb{I}_{X_0 \leq q_X(p)}}{f_X(q_X(p))}, \frac{p - \mathbb{I}_{X_0 \leq q_X(p)}}{f_X(q_X(p))}\right) \\
\quad = \frac{p(1-p)}{f_X^2(q_X(p))} + 2 \sum_{i=1}^{\infty} \mathbb{E}\left[\mathbb{I}_{X_i \leq q_X(p)}\mathbb{I}_{X_0 \leq q_X(p)}\right] - p^2 \\
\text{Cov}(U, V) = \sum_{i\in\mathbb{Z}} \text{Cov}(|X_i|^r, X_0) = \sum_{i\in\mathbb{Z}} \text{Cov}(|X_0|^r, X_i) \\
\text{Cov}(U, W) = -\frac{1}{f_X(q_X(p))} \sum_{i\in\mathbb{Z}} \text{Cov}(\mathbb{I}_{X_i \leq q_X(p)} X_0) = -\frac{\sum_{i\in\mathbb{Z}} \text{Cov}(\mathbb{I}_{X_0 \leq q_X(p)}, X_i)}{f_X(q_X(p))} \\
\text{Cov}(V, W) = -\frac{1}{f_X(q_X(p))} \sum_{i\in\mathbb{Z}} \text{Cov}(|X_0|^r, \mathbb{I}_{X_i \leq q_X(p)}) = -\frac{\sum_{i\in\mathbb{Z}} \text{Cov}(|X_0|^r, \mathbb{I}_{X_0 \leq q_X(p)})}{f_X(q_X(p))}
\end{cases}
\]

Theorem 3 (bivariate FCLT) Consider a process \((X_n)_{n\in\mathbb{Z}}\) that can be represented as a functional \( X_n = f(Z_{n+k}) \) of a strictly stationary process \((Z_n)_{n\in\mathbb{Z}}\) and introduce the random vector

\[
T_n(r) = \left( q_n(p) - q_X(p) \right) \left( \frac{m(X, n, r) - m(X, r)}{\tilde{m}(X, r)} \right), \text{ with integer } r > 0.
\]
Assume for the marginals that the conditions \((M_r), (C_r^+)\) at \(q_X(p)\) and \((C_r')\) in a neighbourhood of \(q_X(p)\) and, for \(r = 1\), \((C_0)\) at \(\mu\), are satisfied.

In terms of dependence, suppose that:

(D1) \((X_n)_{n \in \mathbb{Z}}\) is \(L_2\)-NED with constants \(\nu(k) = O(k^{-\gamma})\), \(s > 6\), on a \(\phi\)-mixing process \((Z_n)_{n \in \mathbb{Z}}\) with mixing coefficient \(\beta(n) = O(n^{-\kappa})\), for some \(\kappa > 3\).

(D2) For \(r > 1\), \(|X_n^r|\) is \(L_2\)-NED with constants \(\nu(k) = O(k^{-\gamma})\), for \(\gamma > 2\).

Then, we have, for \(t \in [0, 1]\),

\[\sqrt{n} t T_{[nt],r}(X) \xrightarrow{D_{2}[0,1]} W_{\Gamma(r)}(t),\]

where \((W_{\Gamma(r)}(t))_{t \in [0,1]}\) is the 2-dimensional Brownian motion with covariance matrix \(\Gamma(r) \in \mathbb{R}^{2 \times 2}\) defined, for any \((s, t) \in [0, 1]^2\), by \(\text{Cov}(W_{\Gamma(r)}(t), W_{\Gamma(r)}(s)) = \min(s, t) \Gamma(r)\), where

\[
\begin{align*}
\Gamma_{11} &= \text{Var}(W), \\
\Gamma_{12} &= r^2 E[X_0^{r-1} \text{sgn}(X_0)^{\gamma}] \text{Var}(U) + \text{Var}(V) - 2r E[X_0^{r-1} \text{sgn}(X_0)^{\gamma}] \text{Cov}(U, V), \\
\Gamma_{21} &= r^2 E[X_0^{r-1} \text{sgn}(X_0)^{\gamma}] \text{Cov}(U, W) + \text{Cov}(V, W),
\end{align*}
\]

\((U, V, W)^T\) being the trivariate normal vector (functional of \(X\)) with mean zero and covariance given in \((Co)\), all series being absolute convergent.

Remark 4

1. The conditions of Theorem \(\mathbf{3}\) are those required to establish a univariate CLT for each component of \(T_{[nt],r}(X)\), namely the sample quantile and the \(r\)-th absolute centred sample moment. For the latter statistic, it requires first establishing a suitable representation, a new result on its own, presented in Proposition \(\mathbf{13}\). Requiring \((C_r')\) and \((C_r^+)\) in a neighbourhood of \(q_X(p)\) corresponds to the conditions for the CLT of the sample quantile of a stationary process, which is \(L_1\)-NED with constants \(\nu(k) = O(k^{-\beta - 3})\) on an absolutely regular process (thus also for a \(\phi\)-mixing process) with mixing rate \(O(n^{-\beta})\), for \(\beta > 3\) - see Corollary 1 in [60].

Further, the \(L_2\)-NED with constants \(\nu(k) = O(k^{-\beta + 3})\) on a \(\phi\)-mixing process with mixing rate \(O(n^{-\beta})\), for \(\beta > 3\), together with \((M_r)\) and the additional assumptions in the cases of \(r = 1\) ((\(C_0\)) at \(\mu\)) and \(r > 1\) (\(L_2\)-NED with constants \(\nu(k) = O(k^{-\gamma})\), \(\gamma > 2\), on \(|X_n|^r\) \(n \in \mathbb{N}\)), are sufficient conditions for the univariate CLT of the \(r\)-th centred sample moment (using Theorem 1.2 in [17]), which is a special case of Theorem 3.1 in [20]).

2. Note that we have a trade-off between more restrictive mixing conditions on the underlying process \((Z_n)_{n \in \mathbb{Z}}\) and more restrictive moment conditions on \(X\). We make this explicit in Corollary \(\mathbf{5}\).

3. For \(r > 1\), there can be an additional trade-off: In Corollary \(\mathbf{6}\) we show that sufficient conditions for the \(L_2\)-NED of \(|X_n|^r\) are a balance between more restrictive moment conditions than \((M_r)\) and less restrictive NED conditions on \((X_n)_{n \in \mathbb{Z}}\) (but more than \(L_2\)-NED).

4. The setting of Theorem \(\mathbf{3}\) includes, as special cases, the case of an underlying iid process \((Z_n)_{n \in \mathbb{Z}}\) as well as a causal representation of \((X_n)_{n \in \mathbb{Z}}\) with respect to \((Z_n)_{n \in \mathbb{Z}}\) - as used in many econometric time-series models; see Section \(\mathbf{3}\) for examples.

The result of Theorem \(\mathbf{3}\) holds also true if we consider an underlying absolutely regular process \((Z_n)_{n \in \mathbb{Z}}\), and slightly adapt the conditions:
**Corollary 5** Assume the setting and conditions of Theorem 3 with the following two changes:

- Take \( \beta = \min (s - 3, \kappa) \) and assume that the moment condition \( (M_\beta \eta) \) holds for an \( \epsilon > 0 \), instead of \( (M_r) \).
- \((D1^*)\) (instead of \((D1)\)): The process \( X \) is \( L_2\)-NED with constants \( \nu (k) = O(k^{-s}), s > 6 \), on an absolutely regular process \( (Z_n)_{n \in \mathbb{Z}} \) (instead of \( \phi \)-mixing) with the same order of mixing coefficient, i.e. \( \beta(n) = O(n^{-\kappa}) \) for some \( \kappa > 3 \).

Then, the FCLT holds for \( T_{n,r}(X) \) as in Theorem 3.

Moreover, we can provide sufficient conditions to reduce the \( L_2\)-NED of \( (|X_n|^r)_{n \in \mathbb{Z}} \) (condition \((D2)\) in Theorem 3 and Corollary 5 respectively) to \( L_q\)-NED of the process \((X_n)_{n \in \mathbb{Z}}\) itself, as follows:

**Corollary 6** Assume the setting and conditions of Theorem 3 or Corollary 5 respectively, replacing \((D2)\) with the following assumptions: For any choice of \( p, q \in (1, \infty) \) s.t. \( \frac{1}{p} + \frac{1}{q} = 1 \), suppose that

- \( \mathbb{E}[|X_0|^{2r-1}] < \infty \),
- \((X_n)_{n \in \mathbb{Z}}\) is \( L_q\)-NED with constants \( \nu (k) \) of the same order as in Theorem 3 or Corollary 5, i.e. \( O(k^{-s}), s > 6 \).

Then, the FCLT holds for \( T_{n,r}(X) \) as in Theorem 3 or Corollary 5 respectively.

Note the following two extreme cases for choices of \( p \) and \( q \) in Corollary 6. Choosing \( p = q = 2 \), we get the most restrictive moment condition \( (\mathbb{E}[|X_0|^{4r-2}] < \infty) \) in combination with \( L_2\)-NED of \((X_n)_{n \in \mathbb{Z}}\), while, when choosing \( p = \frac{2r}{2r-1} \), we get the least restrictive moment condition \( (\mathbb{E}[|X_0|^{2r}] < \infty) \) but needing \( L_{2r}\)-NED of \((X_n)_{n \in \mathbb{Z}}\) at the same time. Only for \( r = 1 \), these two cases coincide.

Let us now turn to two special cases of Theorem 3 presented in Corollaries 7 and 8 and given for sake of completeness. Choosing \( t = 1 \) in Theorem 3 provides the usual CLT, stated in Corollary 7.

**Corollary 7** (CLT) Under the same conditions as in Theorem 3, the joint behaviour of the sample quantile \( q_n(p) \) (for \( p \in (0, 1) \)) and the \( r \)-th absolute centred sample moment \( \hat{m}(X, n, r) \), is asymptotically bivariate normal:

\[
\sqrt{n} \left( \frac{q_n(p) - q_X(p)}{\hat{m}(X, n, r) - m(X, r)} \right) \xrightarrow{d_{n \to \infty}} \mathcal{N}(0, \Gamma^{(r)}),
\]

where the asymptotic covariance matrix \( \Gamma^{(r)} = \left( \Gamma^{(r)}_{ij} \right), 1 \leq i, j \leq 2 \) is defined in Theorem 3.

We can also recover the CLT between the sample quantile and the \( r \)-th absolute centred sample moment stated in the iid case; see [14], Theorem 4.1, which, of course, requires less restrictive conditions.

**Corollary 8** (Theorem 4.1 in [14]). Consider the series of iid rv’s \((X_n)_{n \in \mathbb{Z}}\) with parent rv \( X \). Assume that \( X \) satisfies \( (M_r) \) and both conditions \((C_0^r)\) and \((C_1^r)\) in a neighbourhood of \( q_X(p) \). Additionally, for \( r = 1 \), assume \((C_0)\) at \( \mu \). Then, the joint behaviour of the sample quantile \( q_n(p) \) (for \( p \in (0, 1) \)) and the \( r \)-th absolute centred sample moment \( \hat{m}(X, n, r) \), is asymptotically bivariate normal:

\[
\sqrt{n} \left( \frac{q_n(p) - q_X(p)}{\hat{m}(X, n, r) - m(X, r)} \right) \xrightarrow{d_{n \to \infty}} \mathcal{N}(0, \Gamma^{(r)}),
\]
where the asymptotic covariance matrix \( \Gamma^{(r)} = (\Gamma^{(r)}_{ij}, 1 \leq i, j \leq 2) \) simplifies to

\[
\begin{align*}
\Gamma^{(r)}_{11} &= \frac{p(1-p)}{f_X(q_X(p))}, \\
\Gamma^{(r)}_{12} &= \frac{1}{f_X(q_X(p))} \{ r \mathbb{E}[X_0^{-1} \mathrm{sgn}(X_0)^r] \sigma^2 + \operatorname{Var}(|X_0|^r) - 2r \mathbb{E}[X_0^{-1} \mathrm{sgn}(X_0)^r] \operatorname{Cov}(X_0, |X_0|^r) \}; \\
\Gamma^{(r)}_{21} &= \Gamma^{(r)}_{21} = \frac{1}{f_X(q_X(p))} \{ r \mathbb{E}[X_0^{-1} \mathrm{sgn}(X_0)^r] \operatorname{Cov}(I_{(X_0 \leq q_X(p))}, X_0) - \operatorname{Cov}(I_{(X_0 \leq q_X(p))}, |X_0|^r) \}.
\end{align*}
\]

**Idea of the proof** - Let us briefly sketch in three main steps the proof of Theorem 3 and, analogously, of Corollary 5 which is developed in Section 5.

- First, we represent the sample quantile and \( r \)-th absolute sample moment as sums of functionals of the process \( (X_n) \), so that \( T_{n,r}(X) \) has the following representation, based on \( f(X_j)_{j \in \mathbb{N}} \), for a measurable function \( f : \mathbb{R} \to \mathbb{R}^2 \), with \( \mathbb{E}[f(X_j)] = 0 \) and \( \|f(X_j)\|_2 < \infty \), \( \forall j \),

\[
T_{n,r}(X) = \frac{1}{n} \sum_{j=1}^{n} f(X_j) + o_P(1/\sqrt{n}) \quad \text{as } n \to \infty.
\]

For the sample quantile, we use its Bahadur representation, using a version given in [60] Theorem 1. For that, we check that the NED, moments and mixing of the process are satisfied. These conditions are needed for approximating the sample quantile by an iid sample (and showing that the rest is asymptotically negligible). This approximation idea (coupling technique) dates back to the 70’s with Sen (see e.g. [61] for a brief historical review on the Bahadur representation). Next, we prove, under some conditions, a corresponding asymptotic representation for the \( r \)-th absolute centred sample moment, extending results from [51] and [14]. Here, more work is involved as we consider not only the case when the mean is known but also an unknown mean. As such a representation is of interest on its own, we state it as a separate result, in Proposition [13].

- Next, we have to ensure that each of the components of \( f(X_j)_{j \in \mathbb{N}} \) fulfil the conditions needed for a multivariate FCLT. Those include moment conditions on \( f(X_j) \), for any \( j \), \( L_2 \)-NED of \( f(X_j) \), as well as summability conditions of the mixing coefficients and constants \( \nu(k) \). In particular, we show that the \( L_2 \)-NED condition on \( f(X_j) \) can be reduced to the \( L_2 \)-NED condition for the processes \( (X_j)_{j \in \mathbb{Z}} \) and \( (|X_j|^r)_{j \in \mathbb{Z}} \).

- Finally, combining these results, we show that all conditions are met to apply a multivariate FCLT given in [20] Theorem 3.2 (a multivariate version of Theorem 1.2 in [17]) that we present, adapted to our situation, in Theorem 20. Note that, for the FCLT to hold for the very general class of stationary processes considered in this study, we will have to separately analyse the cases of \( \phi \)-mixing (with moment condition \( (M_\phi) \)) and absolute regularity (with moment condition \( (M_\phi M_\phi^{-1}) \)) for \( (Z_n) \). In each of the two cases, we need to introduce a different set of conditions for the FCLT to hold for \( X_n = f(Z_n) \), making sure that the mixing conditions on \( (X_n) \) are inherited from those on \( (Z_n) \). The \( \phi \)-mixing case is done in the proof of Theorem 3 the absolute regularity one in the proof of Corollary 5

### 3 Application to augmented GARCH and ARMA processes

We focus in this section on two classes of processes widely used in application, which fall within the framework of Theorem 3. By doing so, we provide new results when considering joint standard esti-
mators for such processes. In Section 3.1 we consider the broad family of augmented GARCH(p, q) processes with iid innovations, while, in Section 3.2 ARMA processes with dependent (absolutely regular) innovations; we also discuss an example, the ARMA-GARCH process. For the interested reader, we provide in the Appendix further specific examples of processes that fall in one of the two mentioned families, with conditions on the moments and parameters of the process (for the presented results to hold) stated explicitly for convenience.

3.1 Bivariate FCLT for augmented GARCH(p,q) processes

Since the introduction of the ARCH and GARCH processes in the seminal papers by Engle, [24], and Bollerslev, [7], respectively, various GARCH modifications and extensions have been proposed and their statistical properties analysed (see e.g. [8] for an (G)ARCH glossary). Many such well-known GARCH processes can be seen as special cases of the class of augmented GARCH(p, q) processes, established by Duan in [23]. An augmented GARCH(p,q) process $X = (X_t)_{t \in \mathbb{Z}}$ satisfies, for integers $p \geq 1$ and $q \geq 0$,

$$X_t = \sigma_t \epsilon_t \quad \text{with} \quad \Lambda(\sigma_t^2) = \sum_{i=1}^{p} g_i(\epsilon_{t-i}) + \sum_{j=1}^{q} c_j(\epsilon_{t-j}) \Lambda(\sigma_{t-j}^2), \quad (4)$$

where $(\epsilon_t)$ is a series of iid rv’s with mean 0 and variance 1, $\sigma_t^2 = \text{Var}(X_t)$ and $\Lambda, g_i, c_j, i = 1, \ldots, p, j = 1, \ldots, q$, are real-valued measurable functions. Also, as in [40], we restrict the choice of $\Lambda$ to the so-called group of either polynomial GARCH(p,q) or exponential GARCH(p,q) processes (see Figure 1 in the Appendix for a representation of the hierarchy of these processes):

$$(Lee) \quad \Lambda(x) = x^\delta, \text{ for some } \delta > 0, \quad \text{or} \quad \Lambda(x) = \log(x).$$

Clearly, for a strictly stationary solution to (4) to exist, the functions $\Lambda, g_i, c_j$ as well as the innovation process $(\epsilon_t)_{t \in \mathbb{Z}}$ have to fulfill some regularity conditions (see e.g. [40], Lemma 1). Alike, for the bivariate FCLT to hold, certain conditions on the functions $g_i, c_j, i = 1, \ldots, p, j = 1, \ldots, q$, of the augmented GARCH(p,q) process of the (Lee) family are needed, namely: Positivity of the functions used, (A), and boundedness in $L_r$-norm for either the polynomial GARCH, $(P_r)$, or exponential GARCH, $(L_r)$, respectively, for a given integer $r > 0$,

$$(A) \quad g_i \geq 0, c_j \geq 0, i = 1, \ldots, p, j = 1, \ldots, q,$$

$$(P_r) \quad \sum_{i=1}^{p} \|g_i(\epsilon_0)\|_r < \infty, \quad \sum_{j=1}^{q} \|c_j(\epsilon_0)\|_r < 1,$$

$$(L_r) \quad \mathbb{E} \left[ \exp \left( 4r \sum_{i=1}^{p} |g_i(\epsilon_0)|^2 \right) \right] < \infty, \quad \sum_{j=1}^{q} |c_j(\epsilon_0)| < 1.$$

Note that Condition $(L_r)$ requires the $c_j$ to be bounded functions.

Remark 9 By construction, from (4), $\sigma_t$ and $\epsilon_t$ are independent and $\sigma_t$ is a functional of $(\epsilon_{t-j})_{j=1}^\infty$. Thus, the conditions on the moments, distribution and density, could be formulated in terms of $\epsilon_t$ only. At the same time, this might impose some conditions on the functions $g_i, c_j, i = 1, \ldots, p, j = 1, \ldots, q$ (which might not be covered by (A), (P_r) or (L_r)). Thus, we keep the conditions on $(X_t)_{t \in \mathbb{Z}}$ as such, even if they might not be minimal.

Now, let us explain why and under which circumstances Theorem 3 holds. It has been shown in the literature under which conditions the class of augmented GARCH(p, q) processes fulfills the $L_2$-NED.
More precisely, the conditions of geometric \( L_2 \)-NED on \((X_0)\) and \(|X_0|^r\) are satisfied, on the one hand in the polynomial case under \((M_r), (A)\) and \((P_{\text{max}(1,r/\delta)})\) via Corollary 2 in [40], on the other hand in the exponential case under \((M_r), (A)\) and \((L_r)\) via Corollary 3 in [40]. This can be directly used to reframe the general FCLT given in Theorem 3 for the class of augmented GARCH\((p,q)\) processes: \((L_\text{Lee}), (A)\) and \(P_{\text{max}(1,r/\delta)}\) for polynomial, \((L_r)\) for exponential GARCH models respectively, are sufficient conditions for \((D1)\) and \((D2)\) in Theorem 3 to hold. Thus, Theorem 3 translates as follows:

**Corollary 10** Consider an augmented GARCH\((p,q)\) process \(X\) as defined in [4], which satisfies:

- \((M_r)\), both conditions \((C_2^+), (C_1^+)\) at \(q_X(p)\), and \((C_0)\) at \(0\) for \(r = 1\)
- \((L_\text{Lee}), (A)\), and either \((P_{\text{max}(1,r/\delta)})\) for \(X\) belonging to the group of polynomial GARCH, or \((L_r)\) for the group of exponential GARCH

Then, for \(t \in [0,1]\), as \(n \to \infty\), we obtain the FCLT

\[
\sqrt{n} t \left( \frac{q[p](p) - qX(p)}{m(X, r)} \right) \to \mathcal{W}_{\Gamma(\delta)}(t),
\]

where the 2-dimensional Brownian motion, \((\mathcal{W}_{\Gamma(\delta)}(t))_{t \in [0,1]}\), has the same covariance matrix as in Theorem 3.

### 3.2 Bivariate FCLT for ARMA\((p,q)\) processes

Similarly to the family of GARCH processes, ARMA (AutoRegressive - Moving Average) processes are widely used in applications (e.g. for financial time series). While GARCH models specify a structure of the conditional variance of the process, ARMA processes specify the conditional mean.

Recall that a general ARMA\((p,q)\) process \(X = (X_t)_{t \in \mathbb{Z}}\), for integers \(p \geq 1\) and \(q \geq 0\), is a stationary process defined by

\[
\Phi(B)X_t = \Theta(B)\epsilon_t
\]

where \((\epsilon_t)\) is, in its most general form, a series of rv’s with mean 0 and finite variance, the backward operator \(B\) denotes the application \(B(X_t) = X_{t-1}\), and \(\Phi\) and \(\Theta\) are polynomials of order \(p\) and \(q\) respectively, defined by \(\Phi(z) = 1 + \phi_1z + \ldots + \phi_pz^p (\phi_p \neq 0)\) and \(\Theta(z) = 1 + \theta_1z + \ldots + \theta_qz^q (\theta_q \neq 0)\), such that \(\Phi\) and \(\Theta\) do not have any common root (to have a unique solution to (5)). A necessary and sufficient condition for \(X\) to be causal, i.e. \(X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, t \in \mathbb{Z}\), with \(\sum_{j=0}^{\infty} |\psi_j| < \infty\), is:

\[
\Phi(z) \neq 0, \forall z \in \mathbb{C}\\text{ s.t. } |z| \leq 1.
\]

As already mentioned, various specifications of ARMA processes exist; we refer to [15], [12] and the survey article [33] for further references and examples. While the simplest case assumes iid innovations \((\epsilon_t)\), mixing innovations have been introduced as they allow for broader applications. We can consider this general case thanks to the setup of Theorem 3 (and Corollary 5 respectively).

Note that the geometric \( L_2 \)-NED of ARMA\((p,q)\) processes with mixing innovations can be directly deduced from Theorem 3.1 in [49] under the causality condition \(6\). But, in contrast to the case of augmented GARCH\((p,q)\) processes, we do not have results on the \( L_2 \)-NED of \(|X_0|^r\). So we establish this property in Lemma 19 (see Section 4), showing that the causality condition \(6\) is still sufficient. With these informations at hand, we can replace conditions \((D1)\) and \((D2)\) of Theorem 3 by specific conditions for the class of ARMA\((p,q)\) processes with mixing innovations and state the following:
**Corollary 11** Consider a causal ARMA($p,q$) process $X$ as defined in [5], such that:

- Both conditions $(C'_2)$, $(C'_1)$ at $q_X(p)$ hold, as well as $(C_0)$ at 0 for $r = 1$
- $X$ has either $\phi$-mixing innovations (denoted $X \sim$ARMA-$\phi$), or absolutely regular innovations ($X \sim$ARMA-$\beta$) - in both cases with mixing coefficient of order $O(n^{-\kappa})$, for some $\kappa > 3$
- $X$ satisfies $(M_r)$ for $X$ ARMA-$\phi$, or $(M_{\epsilon+1/2})$, for any $\epsilon > 0$ for ARMA-$\beta$.

Then, for $t \in [0,1]$, as $n \to \infty$,

$$\sqrt{n} \left( \frac{q_{[nt]}(p) - q_X(p)}{n} - m(X, r) \right) \xrightarrow{D_2[0,1]} W_\Gamma(t),$$

where $(W_\Gamma(t))_{t \in [0,1]}$, the 2-dimensional Brownian motion, has the same covariance matrix as in Theorem 9.

Among ARMA processes with mixing innovations, the ARMA-GARCH process is a widespread example in applications (see e.g. [54, 55, 32]).

**Example 12 (ARMA($p,q$)-GARCH($r,s$) process (for $p \geq 1$, $q \geq 0$, $r \geq 1$, $s \geq 0$))**

A general ARMA($p,q$)-GARCH($r,s$) $(X_t)_{t \in \mathbb{Z}}$ is defined as follows:

$$X_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t, \text{ where } \epsilon_t = \eta_0 \sigma_t \text{ and } \sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i}^2,$$

for $(\eta_0)$ an iid series with mean 0 and variance 1.

Let us discuss the conditions of Corollary 17 applied to this class of processes.

- Causality: For an ARMA($p,q$)-GARCH($r,s$) to be causal, [6] needs to be fulfilled
- Conditions on the marginal distribution: $(C'_2)$, $(C'_1)$ at $q_X(p)$ as well as $(C_0)$ at 0 for $r = 1$
- Mixing innovations: The GARCH($r,s$) process is known to be absolutely regular with geometric rate as long as it is strictly stationary, $\gamma_0$ is absolutely continuous with Lebesgue density being strictly positive in a neighbourhood of 0, and $\mathbb{E}[|\eta_0|]$, for some $t \in (0, \infty)$; see [41] Theorem 8 (the original result going back to [17]).

A necessary condition for the strict stationarity of the GARCH($p,q$) process is known in the literature; see e.g. [10] Theorem 1.3. A sufficient condition, easier to verify in practice, is:

$$\mathbb{E}[|\gamma_0|] \sum_{i=1}^{r} \alpha_i + \sum_{i=1}^{s} \beta_i < 1 \text{ for } \mathbb{E}[\gamma_0] < \infty \text{ (see [7]).}$$

- Moment conditions: As the innovations are absolutely regular with geometric rate, the corresponding moment condition is $(M_{\epsilon+1/2})$ for any $\kappa, \epsilon > 0$. Necessary conditions for the existence of such moments depend on the specifications of the ARMA process and can be found e.g. in [45, Theorem 2.2]. In practice, a sufficient moment condition would be $(M_{r+1/2})$.

### 4 Auxiliary Results

In this section we present two different types of results. First, we provide in Proposition 13 an asymptotic representation of the $r$-th absolute centred sample moment $\bar{n}(X, n, r) = \frac{1}{n} \sum_{i=1}^{n} |X_i - X_n|^r$, for any integer $r \geq 1$. Then, in Lemmas 15, 17 and 18 we propose sufficient conditions such that the $L_{2-NED}$
property of \((X_n)_{n \in \mathbb{Z}}\) is inherited for certain bounded and unfunctionals \((f(X_n))_{n \in \mathbb{Z}}\). We use them to establish the \(L_2\)-NED of \((\mathbb{I}_{\{X_n \leq x(p)\}})_{n \in \mathbb{Z}}\) and \(|X_n|^r\)\(n \in \mathbb{Z}, r \in \mathbb{N}\), needed in the proof of Theorem 3. Additionally, Lemma [19] treats the specific case of the \(L_2\)-NED of \(|(X_n)|^r\)\(n \in \mathbb{Z}\) for any integer \(r\) for ARMA\((p,q)\) processes. Proofs of these results are deferred to Section 5.3.

4.1 Representation for the \(r\)-th absolute centred sample moment

Let us state the asymptotic relation between the \(r\)-th absolute centred sample moment with known and unknown mean, respectively. This enables us to compute the asymptotics of \(\hat{m}(X, n, r)\) \((\text{given that the needed moments exist})\).

**Proposition 13** Consider a stationary and ergodic process \((X_n)_{n \geq 0}\) that has ‘short-memory’, i.e. 
\[
\sum_{i=0}^{\infty} \text{Cov}(X_0, X_i) < \infty.
\]
Then, for any integer \(r \geq 1\), assuming that the \(r\)-th moment of \(X\) exists and \(\langle C_0 \rangle\) holds at \(\mu\) for \(r = 1\), we have the following asymptotics, as \(n \to \infty\),
\[
\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} |X_i - X_n|^r \right) = \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} |X_i - \mu|^r \right) - r\sqrt{n}(\bar{X}_n - \mu) E[(X_0 - \mu)^{r-1}\text{sgn}(X_0 - \mu)] + o_P(1). \tag{7}
\]

4.2 \(L_2\)-NED of functionals of \((X_n)_{n \in \mathbb{Z}}\)

Here, we question under which conditions some bounded and unbounded functionals, as e.g. \((\mathbb{I}_{\{X_n \leq x\}})_{n \in \mathbb{Z}}\) and \(|X_n|^r\)\(n \in \mathbb{Z}, r \in \mathbb{N}\), inherit the \(L_2\)-NED of the process \(X = (X_n)_{n \in \mathbb{Z}}\).

For bounded functionals, we can adapt the result proved for the \(L_1\)-NED case \((\text{see} \ [60], \text{Lemma 3.5})\) to the \(L_2\)-NED. For this, we introduce the ‘\(p\)-variation’ condition that dates back to [21], also used in [60, Definition 1.4]. As already noticed there, the \(p\)-variation condition is similar to the notion of \(p\)-continuity of [9, Definition 2.10].

**Definition 14** Let \((X_n)_{n \in \mathbb{Z}}\) be a stationary process. For \(p > 0\), a function \(g : \mathbb{R} \to \mathbb{R}\) satisfies the \(p\)-variation condition (with respect to the distribution of \(X_0\)), if there exists a constant \(c\) such that for any \(\epsilon > 0\)
\[
E\left[\sup_{x \in \mathbb{R} : |x - X_0| \leq \epsilon} |g(x) - g(X_0)|^p\right] \leq c \epsilon^p.
\]

Then, we can adapt Lemma 3.5 in [60] to the \(L_2\)-NED case.

**Lemma 15** Let \((X_n)_{n \in \mathbb{Z}}\) be \(L_2\)-NED with constants \(\nu(k), k \in \mathbb{N},\) on a stationary process \((Z_n)_{n \in \mathbb{Z}}\).

(i) Let \(g\) be a function bounded by \(K\) such that \(g\) satisfies the 2-variation condition with constant \(L\). Then,
\[
(g(X_n))_{n \in \mathbb{Z}} \text{ is } L_2 \text{-NED with constants } \sqrt{(L + 4K^2)}\nu(k), k \in \mathbb{N}. \tag{8}
\]

(ii) For the specific case of \(g\) being the indicator function \(g(x) := \mathbb{I}_{(x \leq t)} (t \in \mathbb{R})\), the result \((8)\) holds under the 1-variation condition (instead of the 2-variation).
Note that the result (ii) will be needed in the proof of Theorem 3.

**Remark 16** Recall that a sufficient condition for the indicator function to satisfy the $1$-variation condition with respect to the distribution of $X_0$, is the Lipschitz-continuity of the distribution function of $X_0$, as shown in [60, Example 1.5].

Turning to unbounded functions, rather than considering $p$-variation (or $p$-continuity), we assume a certain geometry on these functions, namely convexity.

**Lemma 17** Let a process $X = (X_n)_{n \in \mathbb{Z}}$ be $L_2$-NED with constants $\nu(k)$ on a stationary process $(Z_n)_{n \in \mathbb{Z}}$. Consider a positive, convex function $f$ with derivative $f'$ such that, on $\mathbb{R}^+$, $f \times f'$ is convex and positive, and $(f \times f')^p$ is convex with $p > 1$. Then, a sufficient condition for the $L_2$-NED of $(f(X_n))_{n \in \mathbb{Z}}$ with constants $\tilde{\nu}(k) = O(\nu^{1/2}(k))$ is:

$$\mathbb{E}[(f(|X_n|)f'(|X_n|))^p] < \infty \text{ and } X \text{ is } L_q - \text{NED for } q > 1 \text{ s.t. } \frac{1}{p} + \frac{1}{q} = 1, \text{ with } \tilde{\nu}(k) = O(\nu(k)).$$

Applying Lemma 17 for the function $f(\cdot) = |\cdot|^p$, $p \in \mathbb{Z}$, we obtain:

**Lemma 18** Let a process $X = (X_n)_{n \in \mathbb{Z}}$ be $L_2$-NED with constants $\nu(k)$ on a stationary process $(Z_n)_{n \in \mathbb{Z}}$. Then, for a given integer $r \geq 1$, a sufficient condition for the $L_2$-NED of $(|X_n|^r)_{n \in \mathbb{Z}}$ (with constants $\tilde{\nu}(k)$ of order $\tilde{\nu}(k) = O(\nu^{1/2}(k))$) is, for any choice of $p, q \in (1, \infty)$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$,

$$\mathbb{E}[|X_n|^{p(2r-1)}] < \infty, \text{ and } L_q - \text{NED of } X \text{ with constants of order } O(\nu(k)).$$

As an application, let us consider the class of ARMA($p, q$) processes, introduced in Section 3.2.

**Lemma 19** Let $X$ be a causal ARMA($p, q$) process as defined in (5). Then, for any integer $r > 0$, $(|X_0|^r)$ is geometrically $L_2$-NED if $\mathbb{E}[|X_0|^{2r}] < \infty$.

5 Proofs

The main theorem is proven in Section 5.1, while the corollaries in Section 5.2. Proofs of the auxiliary results (presented in Section 4) are given in Section 5.3.

5.1 Proof of Theorem 3

As described in Section 2, the proof consists of three main steps. The first one, more involved, is split into two parts.

**Step 1: Asymptotic representations for the two sample estimators**

1a) : A Bahadur representation for the sample quantile

As explained in Section 2, the main idea in the Bahadur representation is to approximate the sample quantile by an equivalent estimator based on iid rv, i.e. showing that the dependence is asymptotically negligible. The $L_2$-NED of $(X_n)$ allows to approximate these by functionals of finitely many $Z_n$, and the $p$-variation condition ensures that this also holds for $f(X_n)$. Coupling techniques are then used to show that short-range dependent variables have the same behavior as iid ones (for this, mixing conditions
are also necessary). The choice of mixing conditions in Theorem 3 on the underlying process comes from the use of the Bahadur representation for NED processes as given in [60, Theorem 1], for which we need to verify the following conditions.

(i) Let \( g(x, t) := \mathbb{I}_{(x \leq t)} \). It is straightforward to check that \( g \) is non-negative, bounded, measurable, and non-decreasing in the second variable. The function \( g \) also satisfies the variation condition uniformly in some neighbourhood of \( q_X(p) \) if it is Lipschitz-continuous. The latter follows from condition \((C'_2)\) in a neighbourhood of \( q_X(p) \).

(ii) The differentiability of \( \mathbb{E}[g(X, t)] = F_X(t) \) and positivity of its derivative at \( t = q_X(p) \) are given by condition \((C'_1)\) at \( q_X(p) \).

(iii) The condition

\[
|F_X(x) - F_X(q_X(p)) - f_X(q_X(p))(x - q_X(p))| = o \left( (x - q_X(p))^{3/2} \right) \quad \text{as } x \to q_X(p)
\]

is fulfilled as, by Assumption \((C'_2)\), \( F_X \) is twice differentiable in \( q_X(p) \).

(iv) As \((Z_n)_{n \in \mathbb{Z}}\) is stationary and \( \phi \)-mixing (Assumption \((D1)\)), \((X_n)_{n \in \mathbb{Z}}\), being a function of \((Z_n)_{n \in \mathbb{Z}}\), is also stationary and ergodic.

(v) Lastly, the process exactly fulfills the conditions on the mixing rate on the underlying process \((Z_n)_{n \in \mathbb{Z}}\) by assumption \((D1)\). Indeed taking \( \beta := \min(s - 3, \kappa) > 3 \), then, \( \beta(n) = O(n^{-\beta}) \) and \( \nu(k) = O(k^{-\beta+\delta}) \) as \( \beta > 3 \). Further, the \( L_1\)-NED is implied by the assumption of \( L_2\)-NED (Assumption \((D1)\)) at the exact same rate:

\[
\|X_n - \mathbb{E}[X_n | F_{n-k}^{n+k}]\|_1 \leq \|X_n - \mathbb{E}[X_n | F_{n-k}^{n+k}]\|_2 = O(k^{-(\beta+3)}).
\]

Thus, we can use the version of the Bahadur representation given in [60, Theorem 1], replacing (for our purposes) the exact remainder bound by \( o_P(1/\sqrt{n}) \), and write, as \( n \to \infty \),

\[
q_n(p) - q_X(p) + \frac{F_X(q_X(p)) - F_n(q_X(p))}{f_X(q_X(p))} = o_P(1/\sqrt{n}). \tag{10}
\]

1b) Representation of the \( \nu \)-th absolute centred sample moment

The representation is given in Proposition 13 (which proof can be found in Section 5.3). So, we simply need to check that the respective conditions (stationarity, finite 2\( \nu \)-th moment, ergodicity and short-memory) are fulfilled. Let us explain why these conditions hold. We have seen in \((iv)\) above that \((X_n)\) is stationary. The moment condition is fulfilled by assumption due to \((M_r)\) and, for \( r = 1 \), we have \((C_0)\) at \( \mu \) by assumption too. To prove the ergodicity and short-memory property, we use a classical CLT for functionals of \( \phi \)-mixing processes, [6 Theorem 21.1]. It means to check that its conditions are fulfilled. We have \( \mathbb{E}[|X_0|^2] < \infty \), by \((M_r), r \geq 1 \). Moreover, as, by assumption \((D1)\), we have \( L_2\)-NED with constants of the order \( O(k^{-s}), s > 6 \), it holds for those constants that

\[
\sum_{k=1}^{\infty} \nu(k) < \infty. \quad \text{Finally, } \sum_{n=1}^{\infty} \beta(n)^{1/2} < \infty, \quad \text{since via } (D1) \text{ the } \phi \text{-mixing rate is assumed to be of order } \beta(n) = O(n^{-\kappa}), \kappa > 3.
\]

Step 2: Establishing the conditions needed for each component \( f(X_i) \) in view of the FCLT

Following the representation \( \hat{m}(X, n, r) \) of \( \hat{m}(X, n, r) \), we introduce bidimensional tridimensional random vectors \( u_j = (u_{j,l}; l = 1, 2, 3) \), for \( j \in \mathbb{N} \) (anticipating their use in Step 3 for the FCLT of \( U_n(X) := \frac{1}{k} \sum_{j=1}^{k} u_j \)), defined by

\[
u_j = \begin{pmatrix} X_j |X_j|^r - m(X, r) \over p - \mathbb{I}_{|X_j| \leq q_X(p)} f_X(q_X(p)) \end{pmatrix}.
\]
The idea is then to apply a multivariate FCLT, which we adapt from [20, Theorem 3.1] / [17] (a multivariate version of Theorem 1.2) to our needs. Their result is not restricted to stationary processes, but, for us, this is sufficient. This is why our conditions, by stationarity, will hold uniformly in \( t \). It is also the reason why we do not need, in the definition of \( L^p\)-NED, a constant depending on the time \( t \) of the process. Let us present this FCLT here, so that it clearly states the required conditions that each random process \((u_{jl})_{j \in \mathbb{N}}\), for \( l = 1, 2, 3 \), respectively, has to satisfy, in view of its application.

**Theorem 20** Consider a \( d \)-dimensional stationary random process \((u_{j}, j \in \mathbb{N})\) where:

(a) Each of the \( d \) components is an \( L^2\)-NED process with constants \( \nu(k) = O(k^{-\beta}) \), \( \beta > 1/2 \), with respect to the same (unitvariate) process \((V_i)_{i \in \mathbb{Z}}\), which is either \( \alpha \)-mixing of order \( \alpha(n) = O(n^{-s}) \) for \( s > \bar{s}/(\bar{s} - 2), \bar{s} > 2 \), or \( \alpha \)-mixing of order \( \phi(n) = O(n^{-s}) \) for \( s > \bar{s}/(2\bar{s} - 2), \bar{s} \geq 2 \);

(b) For this choice of \( \bar{s} \), it must hold that \( |u_{jl}| \leq \infty, \forall j, l \);

(c) \( \text{Var}(\sum_{j=1}^{n} u_{jl})/n = \sigma_{n,l}^2/n \to 0, \forall l, n \to \infty \).

Then, the series \( \bar{X} = \sum_{j \in \mathbb{N}} \text{Cov}(u_0, u_j) \) converges (coordinate wise) absolutely and a FCLT holds for \( U_n := \frac{1}{n} \sum_{j=1}^{n} u_j \), i.e.

\[
\sqrt{n}U_{[nt]}^{[0,1]} D \to W(t), \quad \text{as } n \to \infty,
\]

where the convergence takes place in the \( d \)-dimensional Skorohod space \( \text{D}([0,1], \mathbb{R}^d) \) and \( (W(t), t \in [0,1]) \) is a \( d \)-dimensional Brownian motion with covariance matrix \( \bar{\Gamma} \), i.e. it has mean 0 and \( \text{Cov}(W(u), W(t)) = \min(u, t)\bar{\Gamma} \).

Choosing \( s = 2 \), let us now show that all the assumptions hold for each of the components of \((u_j)\):

- Condition (a): Let us first comment on the order of the \( L^2\)-NED constants.

For \( u_{j,1} = X_j \), by Assumption (D1), it is of order \( O(k^{-s}) \).

For \( u_{j,2} = I(X_j \leq \alpha_X(p)) \), by (D1) in conjunction with Lemma 15 ii), we have an order of \( O(k^{-s/2}) \).

Finally, for \( u_{j,3} = |X_j|^\gamma \), it is \( O(k^{-\bar{s}/2}) \) by Assumption (D2).

As \( s > 6 \) in (D1) and \( \gamma > 2 \) in (D2), the required order for Theorem 20 is fulfilled.

The mixing rate for \( u_{jl} \), for \( l = 1, 2, 3 \), is, by Assumption (D1), \( O(n^{-\kappa}) \), \( \kappa > 3 \). As we chose \( s = 2 \), the required rate in Theorem 20 is \( O(n^{-\bar{s}}) \) for \( s > \bar{s}/(2\bar{s} - 2) = 1 \). Thus \( \kappa > 3 \) fulfills this requirement.

- Condition (b): We treat each component of \( u_{jl} \), \( l = 1, 2, 3 \), separately. By assumption of \( (M_r), r \geq 1 \), it holds that \( \mathbb{E}[|X_n|^\kappa] < \infty \) (with \( s = 2 \), as chosen). Further, as \( I(X_j \leq \alpha_X(p)) \) is bounded, \( \mathbb{E}[I(X_j \leq \alpha_X(p))|\gamma] < \infty \) holds. Finally, \( \mathbb{E}[|X_n|^\gamma] < \infty \) holds using again \( (M_r), r \geq 1 \).

- Condition (c): For \( u_{j,1} \), the convergence of the variance of the partial sums was already shown in Step 1b) (using Theorem 21.1 of [6]). By the exact same argument, using once again Billingsley’s theorem, (c) holds for \( u_{j,2} \) and \( u_{j,3} \).

**Step 3: Multivariate FCLT**

Having checked the conditions for the FCLT (Theorem 20) in Step 2, we can now apply a trivariate FCLT for \((u_{jl})\). Using the Bahadur representation [10] of the sample quantile (ignoring the rest term for the moment), we can state:

\[
\sqrt{n} \bar{X}^{[nt]} - m(X, r) \xrightarrow{D} \bar{W}_r^{(r)}(t), \quad \text{as } n \to \infty,
\]

where \((\bar{W}_r^{(r)}(t), t \in [0,1])\) is the three-dimensional Brownian motion with covariance matrix \( \bar{\Gamma}^{(r)} \in \mathbb{R}^{3 \times 3} \), i.e. the components \( \bar{\Gamma}_{ij}^{(r)}, 1 \leq i, j \leq 3 \), satisfy the same dependence structure as for the random vector \((U, V, W)^T\) described in (C9), with all series being absolutely convergent. By the multivariate Slutsky theorem, we can add \( \begin{pmatrix} 0 \\ 0 \\ R_{[nt], p} \end{pmatrix} \) to the asymptotics in (11) without changing the resulting
distribution (since $\sqrt{n}R_{[nt],p} \xrightarrow{P} 0$ as $n \to \infty$). Hence, we obtain, as $n \to \infty$,
\[
\sqrt{n} \left( \frac{1}{[nt]} \sum_{j=1}^{[nt]} |X_j|^{r} - m(X,r) \right) + \sqrt{n} \left( 0 \right) = \sqrt{n} \left( \frac{1}{[nt]} \sum_{j=1}^{[nt]} |X_j|^{r} - m(X,r) \right)
\]
\[\xrightarrow{D_{[0,1]}} W_{\Gamma^{(r)}}(t).\]

Then, we apply to (12) the multivariate continuous mapping theorem using the function $f(x,y,z) \mapsto (ax + y, z)$ with $a = -r \mathbb{E}((X - \mu)^{r-1} \text{sgn}(X - \mu))$. Further, by Slutsky’s theorem, we can add to $ax + y$ a rest of $o_P(1/\sqrt{n})$ without changing the limiting distribution. So, we obtain, as $n \to \infty$,
\[
\sqrt{n} \left( a(\bar{X}_{[nt]} - \mu) + \frac{1}{[nt]} \sum_{j=1}^{[nt]} |X_j|^{r} - m(X,r) + o_P(1/\sqrt{n}) \right)
\]
\[= \sqrt{n} \left( m(X_{[nt]},r) - m(X,r) \right) \xrightarrow{D_{[0,1]}} W_{\Gamma^{(r)}}(t),\]
where $\Gamma^{(r)}$ follows from the specifications of $\bar{\Gamma}^{(r)}$ above the end the continuous mapping theorem. \(\square\)

### 5.2 Proofs of Corollaries

We start with the proof of Corollary 5 where we need to show why Theorem 3 also holds in the case of an underlying absolutely regular process. The proofs of corollaries 7, 8 and 10 are a direct consequence of Theorem 4, so can be omitted. The proof of Corollary 6 mainly consists of establishing sufficient conditions for the $L_2$-NED of $(|X_n|^r)_{n \in \mathbb{Z}}$, which will be done in (the proof of) Lemma 18 (proven in Section 5.3). We end with the proof of Corollary 11 for which the main work consists of proving the $L_2$-NED of $|X|^r$, $r > 0$, for ARMA processes.

**Proof of Corollary 5** The proof follows the one of Theorem 3 but we need to adapt (parts of) Steps 1b) and 2 to the setting of Corollary 5 i.e. for an underlying absolute regular process. Thus, we revisit these two steps.

**Step 1b)**: Representation of the $r$-th absolute centred sample moment - Conditions.

Given the representation given in Proposition 13, we need to check that the respective conditions, stationarity, finite 2$r$-th moment, ergodicity and short-memory, are fulfilled. Only the reasoning for the short-memory property differs from the one in Theorem 3. To prove the ergodicity and short-memory property, we verify that the conditions for a CLT of $X_n$ are fulfilled. Since absolute regularity implies strong mixing at the same rate, we consider the CLT for functionals of strongly mixing processes; see [X6, Theorem 18.6.2]. By choosing $\delta = 2 \frac{s + 2r - \beta}{\beta - 1}$ in that theorem, we check that the conditions stated there are fulfilled (recall that we defined in Corollary 5 $\beta := \min(s - 3, \kappa)$). We have $\mathbb{E}[|X_0|^{2+\delta}] = \mathbb{E}[|X_0|^{2+\frac{2+\delta}{1+\delta}}] < \infty$ by $(M_{r,\frac{2+\delta}{1+\delta}})$ (as $r \geq 1$). Moreover, $\sum_{k=1}^{\infty} ||X_0 - \mathbb{E}[X_0|\mathcal{F}_{k}^{+}]||_{2+\delta}/(1+\delta) = \sum_{k=1}^{\infty} ||X_0 - \mathbb{E}[X_0|\mathcal{F}_{k}^{+}]||_{2+\delta}/(1+\delta)$ holds as it is bounded by $\sum_{k=1}^{\infty} ||X_0 - \mathbb{E}[X_0|\mathcal{F}_{k}^{+}]||_{2}$ (since $\delta > 0$), which is finite by the assumption of $L_2$-NED with rate $O(k^{-\kappa})$, $s > 6$, i.e. $(D1^*)$. Finally, $\sum_{n=1}^{\infty} \beta(n)^{\delta/(2+\delta)} < \infty$ holds by construction (the choice of $\delta$ was made in a way that the sum is finite): $(D1^*)$ ensures that $\beta(n) = O(n^{-\kappa})$ and, by the choice of $\beta$ in the corollary, implies $\beta(n) = O(n^{-\beta})$. We assume, w.l.o.g., that $\beta(n) = n^{-\beta}$. We then have
\[
\sum_{n=1}^{\infty} \beta(n)^{\delta/(2+\delta)} = \sum_{n=1}^{\infty} \frac{n^{-\beta}}{\beta + \frac{2+\delta - \beta}{\beta - 1}} = 1 + \frac{\epsilon(\beta - 1)}{\beta + \epsilon}.
\]
As \( \beta := \min (s - 3, \kappa) > 3, \epsilon > 0 \), this quantity will always be bigger than 1 and hence the infinite sum remains summable. Hence Theorem 18.6.2 in [36] applies in this case for the process \((X_n)_{n \in \mathbb{Z}}\).

**Step 2: Conditions for applying the FCLT (Theorem 20)**

In Theorem [3], we used Theorem [20] as multivariate FCLT, which does not only cover the case of underlying \( \phi \)-mixing processes, but also strong mixing. It means that we can also use it here. Therefore, it comes back to verify Condition \((a^*)\), defined below, and Conditions \((b)\) and \((c)\) given in the proof of Theorem [3] Step 2.

\((a^*)\) L\(_2\)-NED process with constants \( \nu(k) = O(k^{-\eta}), \eta > 1/2 \) (changed from \( \beta \) to \( \eta \), to avoid notational confusion) on a univariate process \((V_t)_{t \in \mathbb{Z}}\) (the same for all components), which, in this case, is \( \alpha \)-mixing of order \( \alpha(n) = O(n^{-3}) \) for any \( \tilde{s} > s/(s - 2) \) for a \( s > 2 \).

Note that, in comparison to the proof of Theorem [3] only the first condition has been adapted. Nevertheless, as the underlying mixing property is different, the proof of conditions \((b)\) and \((c)\) have also to be adapted. Choosing \( s = 2 \frac{\beta + \epsilon}{\beta - 1} \) in Theorem [20], we can show that all the assumptions hold for each of the components of \((u_j)\).

- **Condition \((a^*)\):** The order of the \( L_2\)-NED constants being the same as for the \( \phi \)-mixing case, see \((D1^*)\) and \((D2)\), the same arguments as in the proof of Theorem [3] hold. The mixing rate for \((u_{j,l}), l = 1, \ldots, 3\), is, by assumption \((D1^*)\), \( O(n^{-\kappa}), \kappa > 3 \), which implies, by definition of \( \beta \), that it is also of order \( O(n^{-\beta}), \beta > 3 \). Since we chose \( s = 2 \frac{\beta + \epsilon}{\beta - 1} \), we have \( \frac{s}{s-2} = \beta \frac{1+\epsilon/\beta}{1+\epsilon} \) (as \( \epsilon > 0, \beta > 3 \)). The required rate in Theorem [20] is \( O(n^{-\tilde{s}}) \) for \( \tilde{s} > s/(s - 2) \) fulfills this requirement.

- **Condition \((b)\):** We treat each component of \((u_{j,l}), l = 1, 2, 3\), separately. By assumption of \((M_{r, \frac{s+\epsilon}{s-1}}), r \geq 1\), it holds that \( \mathbb{E}||X_n|^s| < \infty \). Further, as \( \mathbb{I}(X_j \leq qX(p)) \) is bounded, \( \mathbb{E}[\mathbb{I}(X_j \leq qX(p))^s]^s < \infty \). Finally, \( \mathbb{E}[|X_n|^s] < \infty \), using again the assumption of \((M_{r, \frac{s+\epsilon}{s-1}}), r \geq 1\).

- **Condition \((c)\):** For \((u_{j,1})\), the convergence of the variance of the partial sums was already shown in Step 1b (using [36] Theorem 18.6.2). By the same argument, with the same choice of \( \delta \) as for \( u_{j,1} \), it also follows for \( u_{j,2} \) and \( u_{j,3} \) (using once again Ibragimov’s theorem).

**Proof of Corollary 11.** Comparing the conditions of Corollary 11 with Theorem 3 or Corollary 5 respectively, we simply need to show why the causality (condition (6)), and \( X \sim (ARMA - \phi) \) or \( (ARMA - \beta) \), are sufficient for the \( L_2\)-NED of \((X_0)\) and \( |X_0^r| \) with constants \( \nu(k) = O(k^{-\gamma}), s > 6 \), and \( \nu(k) = O(k^{-\gamma}), \gamma > 2 \), respectively. For \( X_0 \), this follows directly by [49] Lemma 3.1] (by their result, a causal ARMA(\(\beta,q\)) is geometrically \( L_2\)-NED). For \(|X_0|^r\), the geometric \( L_2\)-NED has been exactly established through Lemma 19. Finally, as geometric \( L_2\)-NED implies a rate of \( \nu(k) = O(k^{-s}), s > 6 \), the necessary rate for the application of Theorem 3 and Corollary 5 respectively, is attained.

5.3 Proofs of Auxiliary Results

We start by establishing the asymptotics of the \( r \)-th absolute centred sample moment (Proposition 1.3). To do so, we need the following lemma, which extends Lemma 2.1 in [51] (case \( v = 1 \)) to any moment \( v \in \mathbb{N} \), as well as the iid case presented in Lemma A.1 in [14].
Lemma 21 Consider a stationary and ergodic process \((X_n, n \geq 0)\) with ‘short-memory’, i.e. \(\sum_{i=0}^{\infty} \text{Cov}(X_0, X_i) < \infty\). Then, for \(v = 1\) or \(2\), given that the 2nd moment of \(X_0\) exists, or, for any integer \(v > 2\), given that the \(v\)-th moment of \(X_0\) exists, it holds that, as \(n \to \infty\),

\[
\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^v (|X_i - \bar{X}_n| - |X_i - \mu|) = \left(\bar{X}_n - \mu\right) \times \mathbb{E}[(X_0 - \mu)^v \text{sgn}(\mu - X_0)] + o_P(1/\sqrt{n}).
\] (14)

Proof of Lemma 21 and of Proposition 13 The proofs of Lemma 21 and Proposition 13 follow the lines of their equivalents in the iid case; see the proof of Lemma A.1 and Proposition A.1 in [14]. In the dependent case, it needs to be adapted, using the stationarity, ergodicity and short-memory of the process. By these three properties, it follows that \(\sqrt{n} |\bar{X}_n - \mu|^v \overset{P}{\to} 0\) for any integer \(v \geq 1\). Further, we use the ergodicity of the process, instead of the strong law of large numbers, to conclude that

\[
\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^v \text{sgn}(\mu - X_i) \xrightarrow{a.s.} \frac{1}{n} \mathbb{E}[(X_0 - \mu)^v \text{sgn}(\mu - X_0)].
\]

Finally, for the proof of Proposition 13, we apply Lemma 21 instead of its counterpart in the iid case ([14] Lemma A.1).

\(\Box\)

Proof of Lemma 15

(i) First, let us observe that

\[
\mathbb{E}[(g(X_0) - E[g(X_0)|\mathcal{F}_k^+])^2] \leq \mathbb{E}[(g(X_0) - g(E[X_0|\mathcal{F}_k^+]))^2] = \mathbb{E}[(g(X_0) - g(E[X_0|\mathcal{F}_k^+]))^2 I_{|X_0 - E[X_0|\mathcal{F}_k^+]| \leq \sqrt{\nu(k)}}] + \mathbb{E}[(g(X_0) - g(E[X_0|\mathcal{F}_k^+]))^2 I_{|X_0 - E[X_0|\mathcal{F}_k^+]| > \sqrt{\nu(k)}}],
\]

where the first inequality follows from the fact that \(\mathbb{E}[(X - E[X|Z])^2] \leq \mathbb{E}[(X - h(Z))^2]\) for any measurable function \(h\), in particular for \(h = g\). For (15), we obtain by using the 2-variation condition,

\[
\mathbb{E}[(g(X_0) - g(E[X_0|\mathcal{F}_k^+]))^2 I_{|X_0 - E[X_0|\mathcal{F}_k^+]| \leq \sqrt{\nu(k)}}] \leq \mathbb{E}[\sup_{x \in \mathbb{R} : |x - X_0| \leq \sqrt{\nu(k)}} |g(X_0) - g(x)|^2] \leq L \nu(k).
\]

For (16), using the boundedness of \(g\) by \(K\), we have

\[
\mathbb{E}[(g(X_0) - g(E[X_0|\mathcal{F}_k^+]))^2 I_{|X_0 - E[X_0|\mathcal{F}_k^+]| > \sqrt{\nu(k)}}] \leq 4K^2 \mathbb{P}\left(|X_0 - E[X_0|\mathcal{F}_k^+]| \geq \sqrt{\nu(k)}\right).
\]

Then, by the extended version of the Markov inequality and the fact that, by assumption, \((X_0)\) is \(L_2\)-NED with constant \(\nu(k)\), we obtain

\[
\mathbb{P}\left(|X_0 - E[X_0|\mathcal{F}_k^+]| \geq \sqrt{\nu(k)}\right) \leq \frac{||X_0 - E[X_0|\mathcal{F}_k^+]|^2}{\nu(k)} \leq \nu(k),
\]

from which we deduce that (16) can be bounded by \(4K^2 \nu(k)\). Hence, we can conclude that

\[
\mathbb{E}[|g(X_0) - E[g(X_0)|\mathcal{F}_k^+]|^2] \leq (15) + (16) \leq (L + 4K^2) \nu(k)
\]
For the last step, notice that the composed function \( g \geq \) (where the second inequality follows by the convexity of \( f \)) and obtain an equivalent expression for the RHS of (19):

Consequently, we can apply Jensen’s inequality for conditional expectations to this compound function \( g \).

Denoting \( g \mid \), by the definition of conditional expectation as minimizer in \( \mathbb{L}_k \)

hence, after noticing that \( a = \alpha x + (1 - a) y \) with \( a \in [0, 1] \), we can bound the RHS of (17) as

\[
\mathbb{E} \left[ g(X_n) - g(\mathbb{E}[X_n|\mathcal{X}^k_{-k}]) \right] \leq \mathbb{E} \left[ g'(aX_n + (1 - a) \mathbb{E}[X_n|\mathcal{X}^k_{-k}]) \left( X_n - \mathbb{E}[X_n|\mathcal{X}^k_{-k}] \right) \right]
\]

hence, after noticing that \( g' = 2ff' \) is an increasing function and using the triangle inequality for the absolute value, it comes

\[
\mathbb{E} \left[ g(X_n) - g(\mathbb{E}[X_n|\mathcal{X}^k_{-k}]) \right] \leq \mathbb{E} \| g'(aX_n + (1 - a) \mathbb{E}[X_n|\mathcal{X}^k_{-k}]) \| (X_n - \mathbb{E}[X_n|\mathcal{X}^k_{-k}]) \]  \tag{18}

Now, applying Jensen’s inequality (as \( g' \) is convex on \( \mathbb{R}^+ \) by assumption), then, Hölder’s one, and finally the triangle inequality, the RHS of (18) can be bounded by

\[
\mathbb{E}[\|ag'(|X_n|) + (1 - a)g'(|\mathbb{E}[X_n|\mathcal{X}^k_{-k}]|)\| |X_n - \mathbb{E}[X_n|\mathcal{X}^k_{-k}]|] \leq \|ag'(|X_n|) + (1 - a)g'(|\mathbb{E}[X_n|\mathcal{X}^k_{-k}]|)\|_p \|X_n - \mathbb{E}[X_n|\mathcal{X}^k_{-k}]\|_q \leq \left( a\|g'(|X_n|)\|_p + (1 - a)\|g'(|\mathbb{E}[X_n|\mathcal{X}^k_{-k}]|)\|_p \right) \|X_n - \mathbb{E}[X_n|\mathcal{X}^k_{-k}]\|_q. \tag{19}
\]

For the last step, notice that the composed function \( g'(|\cdot|)^p \) is convex: As \( (g')^p \) is non-decreasing for \( x \geq 0 \), it holds by the convexity of the absolute value function:

\[
(g'(\alpha x + (1 - \alpha)y))^p \leq (g'(\alpha|x| + (1 - \alpha)|y|))^p \leq \alpha (g'(|x|))^p + (1 - \alpha) (g'(|y|))^p,
\]

where the second inequality follows by the convexity of \( (g')^p \) on \( \mathbb{R}^+ \), which also holds by assumption. Consequently, we can apply Jensen’s inequality for conditional expectations to this compound function and obtain an equivalent expression for the RHS of (19):

\[
(a\|g'(|X_n|)\|_p + (1 - a)\|g'(|\mathbb{E}[X_n|\mathcal{X}^k_{-k}]|)\|_p) \|X_n - \mathbb{E}[X_n|\mathcal{X}^k_{-k}]\|_q = \|g'(|X_n|)\|_p \|X_n - \mathbb{E}[X_n|\mathcal{X}^k_{-k}]\|_q.
\]

Thus, combining all these calculations and inequalities, we can conclude that

\[
\|f(X_n) - \mathbb{E}[f(X_n)|\mathcal{X}^k_{-k}]\|_2 \leq \sqrt{\|g'(|X_n|)\|_p \|X_n - \mathbb{E}[X_n|\mathcal{X}^k_{-k}]\|_q} \leq O(\nu^{1/2}(k)),
\]

where the last inequality holds using the assumptions \( L_q \)-NED of \( (X_n)_{n \in \mathbb{N}} \) and the \( L_p \)-boundedness of \( g'(|X_n|)\).
Proof of Lemma 18. Consider the function \( f(x) = |x|^r \). We have \( f'(x) = r|x|^{r-1} \text{sgn}(x) \) and \( f''(x) = r(r-1)|x|^{r-2} \). Thus, we can verify the conditions of Lemma 17 as \( f \) is positive and convex (\( f'' \geq 0 \)), \( f \times f' \) is positive for \( x \geq 0 \) and also convex for \( x \geq 0 \) (since \( f f'(x) = r x^{2r-1} \) and \( (f f')'' = r(2r-1)(2r-2)x^{2r-3} \) for \( r 
eq 1 \) (and 0 for \( r = 1 \)). Finally, \( (f f')'' \) is convex for \( x \geq 0 \), for a choice of \( p \). Indeed, as \( (f f')'' = r x^{p(2r-1)} - \frac{p+1}{2} (2p-1) \) for \( r 
eq \frac{p+1}{2} \) (and equal to 0 for \( r = \frac{p+1}{2} \)), which is, for all choices of \( p \in (1, \infty) \), positive for \( x \geq 0 \). □

Proof of Lemma 19. The proof consists of three steps. First, the process \( X \) being a causal ARMA\((p, q)\)-process, we can apply Theorem 3.1 from [49]. Choosing \( p = 2 \) therein, we can conclude that \( X \) is strong \( L_2\)-NED with rate \( \nu(k) = \rho^k \), for some \( 0 < \rho < 1 \), hence it is \( L_2\)-NED with the same rate (compare Definitions 1.1 and 1.2 in [49]). Second, since \( X \), equivalently \((X_0)\), is \( L_2\)-NED, we can apply Lemma 18. As it holds by assumption that \( |X_0|^{2r} < \infty \), we choose \( q = 2r \) and accordingly \( p = \frac{2r}{2r-1} \). Then, by Lemma 18, it holds that \( (|X_0|^r) \) is geometrically \( L_2\)-NED if \((X_0)\) is geometrically \( L_{2r}\)-NED. Finally, let us prove that this sufficient condition holds. For this, recall the representation of an ARMA process (which holds for an ARMA\((p, q)\) process satisfying (5), (6), see Lemma 3.1 in [49]):

\[
X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \quad \text{with} \quad |\psi_j| = O(\rho^j) \quad \text{for some} \quad 0 < \rho < 1.
\]

Then, we apply a standard truncation argument (as e.g. done in [40], Lemma 1 for augmented GARCH\((p, q)\) processes): Define the truncated variable \( h_{t,m} = \sum_{j=0}^{m} \psi_j \epsilon_{t-j} \), which, by construction, is \( F_{\epsilon,t-m,t+m} \)-measurable. Let \( h_{t,m}^* := X_t - h_{t,m} \). For any given integer \( r > 0 \), let us now verify the \( L_2r\)-NED of \((X_0):\)

\[
\|X_0 - \mathbb{E}[X_0 | F_{\epsilon,-m+1,m}]\|_r^2 = \|h_{0,m}^* - \mathbb{E}[h_{0,m}^* | F_{\epsilon,-m+1,m}]\|_r^2 \leq 2^{1+2r} \|h_{0,m}^*\|_r^2,
\]

where the first equality follows from the \( F_{\epsilon,-m+1,m} \)-measurability of \( h_{0,m} \) and the subsequent inequality because \( \mathbb{E}[|X - \mathbb{E}[X | Y]|^{2r}] \leq 2^{(1+2r)} \mathbb{E}[|X^r|] \) for any random variables \( X, Y \) and \( r \geq 1 \).

Thus, it is enough to prove the geometric \( L_{2r}\)-NED of \( h_{0,m}^* \) to conclude the proof. Since we can write

\[
\|h_{0,m}^*\|_r = \|\sum_{j=m+1}^{\infty} \psi_j \epsilon_{j-m}\|_r \leq \sum_{j=m+1}^{\infty} \|\psi_j \epsilon_{j-m}\|_r = \sum_{j=m+1}^{\infty} |\psi_j| \|\epsilon_{j-m}\|_r = \|\epsilon_0\|_r \sum_{j=m+1}^{\infty} |\psi_j|,
\]

where \( \epsilon_0 \) has a finite second moment by definition of the ARMA\((p, q)\) process, then, using (20), we obtain that, for a constant \( C > 0 \),

\[
\|h_{0,m}^*\|_r \leq \|\epsilon_0\|_r \sum_{j=m+1}^{\infty} |\psi_j| \leq C \sum_{j=m+1}^{\infty} \rho^j = C \rho^{m+1} \sum_{j=0}^{\infty} \rho^j = C \rho^{m+1} \frac{1}{1-\rho} = O(e^{-m \eta}),
\]

for \( \eta = -\log(\rho) \), hence the result. □

Financial disclosure

None reported.

Conflict of interest

The authors declare no potential conflict of interests.
References

[1] T. W. Anderson. The statistical analysis of time series. John Wiley & Sons, 1971.
[2] D. Andrews. Laws of large numbers for dependent non-identically distributed random variables. Econometric Theory, 4(3):458–467, 1988.
[3] A. Aue, I. Berkes, and L. Horváth. Strong approximation for the sums of squares of augmented GARCH sequences. Bernoulli, 12(4):583–608, 2006.
[4] R. Bahadur. A note on quantiles in large samples. The Annals of Mathematical Statistics, 37(3):577–580, 1966.
[5] I. Berkes, S. Hörmann, and L. Horváth. The functional central limit theorem for a family of GARCH observations with applications. Statistics & Probability Letters, 78(16):2725–2730, 2008.
[6] P. Billingsley. Convergence of probability measures. John Wiley & Sons, 1st edition, 1968.
[7] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3):307–327, 1986.
[8] T. Bollerslev. Glossary to ARCH (GARCH). CREATES Research Paper, 49, 2008.
[9] S. Borovkova, R. Burton, and H. Dehling. Limit theorems for functionals of mixing processes with applications to u-statistics and dimension estimation. Transactions of the American Mathematical Society, 353(11):4261–4318, 2001.
[10] P. Bougerol and N. Picard. Stationarity of GARCH processes and of some nonnegative time series. Journal of econometrics, 52(1-2):115–127, 1992.
[11] F. Boussama. Ergodicité, mélange et estimation dans les modeles GARCH. PhD thesis, Université Paris 7, 1998.
[12] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. Time series analysis: forecasting and control. John Wiley & Sons, 2015.
[13] R. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. Probability surveys, 2:107–144, 2005.
[14] M. Brätigam, M. Dacorogna, and M. Kratz. Pro-cyclicality beyond business cycles. Mathematical Finance, 33(2):308–341, 2023.
[15] P. J. Brockwell and R. A. Davis. Time series: theory and methods. Springer Science & Business Media, 1991.
[16] M. Carrasco and X. Chen. Mixing and moment properties of various GARCH and stochastic volatility models. Econometric Theory, 18(1):17–39, 2002.
[17] J. Davidson. Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. Journal of Econometrics, 106(2):243–269, 2002.
[18] R. Davis and T. Mikosch. The sample autocorrelations of heavy-tailed processes with applications to ARCH. The Annals of Statistics, 26(5):2049–2080, 1998.
[19] R. A. Davis, T. Mikosch, and B. Basrak. Sample ACF of multivariate stochastic recurrence equations with application to GARCH. Preprint, available at www. math. ku. dk/mikosch, 1999.
[20] R. M. De Jong and J. Davidson. The functional central limit theorem and weak convergence to stochastic integrals I: weakly dependent processes. Econometric Theory, 16(5):621–642, 2000.
[21] M. Denker and G. Keller. Rigorous statistical procedures for data from dynamical systems. Journal of Statistical Physics, 44(1-2):67–93, 1986.
[22] Z. Ding, C. Granger, and R. Engle. A long memory property of stock market returns and a new model. Journal of Empirical Finance, 1(1):83–106, 1993.
[23] J. Duan. Augmented GARCH (p, q) process and its diffusion limit. Journal of Econometrics, 79(1):97–127, 1997.

[24] R. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: Journal of the Econometric Society, 50(4):987–1007, 1982.

[25] R. Engle and V. Ng. Measuring and testing the impact of news on volatility. The Journal of Finance, 48(5):1749–1778, 1993.

[26] J. Geweke. Modeling the persistence of conditional variances: a comment. Econometric Reviews, 5:57–61, 1986.

[27] P. Giot and S. Laurent. Modelling daily value-at-risk using realized volatility and ARCH type models. Journal of empirical finance, 11(3):379–398, 2004.

[28] L. Giraitis, R. Leipus, and D. Surgailis. Recent advances in ARCH modelling. In G. Teyssi`ere and A.P. Kirman, editors, Long Memory in Economics, pages 3–38. Springer, 2007.

[29] L. Glosten, R. Jagannathan, and D. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance, 48(5):1779–1801, 1993.

[30] CH. Hesse. A Bahadur-type representation for empirical quantiles of a large class of stationary, possibly infinite-variance, linear processes. The Annals of Statistics, pages 1188–1202, 1990.

[31] M. Higgins and A. Bera. A class of nonlinear ARCH models. International Economic Review, 33(1):137–158, 1992.

[32] Y. Hoga. Confidence intervals for conditional tail risk measures in arma–garch models. Journal of Business & Economic Statistics, 37(4):613–624, 2019.

[33] S. H. Holan, R. Lund, and G. Davis. The ARMA alphabet soup: A tour of ARMA model variants. Statistics Surveys, 4:232–274, 2010.

[34] S. H¨ormann. Augmented GARCH sequences: Dependence structure and asymptotics. Bernoulli, 14(2):543–561, 2008.

[35] L. Horv´ath and P. Kokoszka. Sample autocovariances of long-memory time series. Bernoulli, 14(2):405–418, 2008.

[36] I. Ibragimov and Y. Linnik. Independent and stationary sequences of random variables. Wolters, Noordhoff Pub., 1971.

[37] I. A. Ibragimov. Some limit theorems for stationary processes. Theory of Probability & Its Applications, 7(4):349–382, 1962.

[38] J. Jeon and J. W. Taylor. Using CAViaR models with implied volatility for value-at-risk estimation. Journal of Forecasting, 32(1):62–74, 2013.

[39] R. Kulik and P. Soulier. Limit theorems for long-memory stochastic volatility models with infinite variance: partial sums and sample covariances. Advances in Applied Probability, 44(4):1113–1141, 2012.

[40] O. Lee. Functional central limit theorems for augmented GARCH (p, q) and FIGARCH processes. Journal of the Korean Statistical Society, 43(3):393–401, 2014.

[41] A. Lindner. Stationarity, mixing, distributional properties and moments of GARCH (p, q)–processes. In T. Mikosch, JP. Kreiß, R. Davis, and T. Andersen, editors, Handbook of financial time series, pages 43–69. Springer, 2009.

[42] S. Ling and W.K. Li. On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity. Journal of the American Statistical Association, 92(439):1184–1194, 1997.

[43] S. Ling and W.K. Li. Limiting distributions of maximum likelihood estimators for unstable autoregressive moving-average time series with general autoregressive heteroscedastic errors. Annals of Statistics, 26(1):84–125, 1998.

[44] S. Ling and M. McAleer. Necessary and sufficient moment conditions for the GARCH (r, s) and asymmetric power GARCH (r, s) models. Econometric theory, pages 722–729, 2002.
Examples of Augmented GARCH models

First, let us give a schematic overview of the nesting of some augmented GARCH(\(p,q\)) models in Figure 1 with a brief description of the acronyms, authors, and relations between these processes. Then, we explicitly state the conditions on the moments and parameters of these processes, for the bivariate asymptotics of Corollary 10 to be valid, in view of applications.

The restrictions on the parameters, if not specified differently, are \(\omega \geq 0, \alpha_i \geq 0, -1 \leq \gamma_i \leq 1, \beta_j \geq 0\) for \(i = 1, \ldots, p, j = 1, \ldots, q\).
• APGARCH: Asymmetric power GARCH, introduced by Ding et al. in [22]. One of the most general polynomial GARCH models.

• AGARCH: Asymmetric GARCH, defined also by Ding et al. in [22], choosing $\delta = 1$ in APGARCH.

• GJR-GARCH: This process is named after its three authors Glosten, Jagannathan and Runkle and was defined by them in [29]. For the parameters $\alpha_i^+, \alpha_i^-$ it holds that $\alpha_i^+ = \alpha_i (1 - \gamma_i)^2$ and $\gamma_i^+ = 4 \alpha_i \gamma_i$.

• GARCH: Choosing all $\gamma_i = 0$ in the AGARCH model (or $\gamma_i^+ = 0$ in the GJR-GARCH), gives back the well-known GARCH($p,q$) process by Bollerslev in [7].

• ARCH: Introduced by Engle in [24]. We recover it by setting all $\gamma_i = \beta_j = 0, \forall i, j$.

• TGARCH: Choosing $\delta = 1/2$ in the APGARCH model leads us the so called threshold GARCH (TGARCH) by Zakoian in [63]. For the parameters $\alpha_i^+, \alpha_i^-$ it holds that $\alpha_i^+ = \alpha_i (1 - \gamma_i)$, $\alpha_i^- = \alpha_i (1 + \gamma_i)$.

• TSGARCH: Choosing $\gamma_i = 0$ in the TGARCH model we get, as a subcase, the TSGARCH model, named after its authors, i.e. Taylor, [56], and Schwert, [50].

• PGARCH: Another subfamily of the APGARCH processes is the Power-GARCH (PGARCH), also called sometimes NGARCH (i.e. non-linear GARCH) due to Higgins and Bera in [31].

• VGARCH: The volatility GARCH (VGARCH) model by Engle and Ng in [25] is also a polynomial GARCH model but is not part of the APGARCH family.

• NGARCH: This non-linear asymmetric model is due to Engle and Ng in [25], and sometimes also called NAGARCH.

• MGARCH: This model is called multiplicative or logarithmic GARCH and goes back to independent suggestions, in slightly different formulations, of Geweke in [26], Pantula in [48] and Milhøj in [46].

• EGARCH: This model is called exponential GARCH, introduced by Nelson in [47].

Now, recalling the specific conditions mentioned in Corollary [10] note that the continuity and differentiability conditions, $(C_2^r), (C_1^r)$, each at $q_X(p)$, and $(C_0)$ at $r = 1$, remain the same for the whole class of augmented GARCH processes. Contrary to that, the moment condition $(M_r)$ imposes different restrictions on the parameters of the underlying process, depending on the given augmented GARCH($p,q$) process. To our knowledge, there exists no general result for the class of augmented GARCH($p,q$) describing the necessary conditions of the
process for a given moment to exist: E.g. in [34] the case of augmented GARCH(1,1) processes is considered, see Corollary 1 and Proposition 1 therein, and in [44] the GARCH(p, q) and APGARCH(p, q) see Theorem 2.1 and Theorem 3.2 therein. Sufficient but not necessary conditions, which are easier to verify in practice, follow from results in [28] as mentioned in [41], Proposition 2.

We focus here on solely explaining how, depending on the specifications [4] of the process, the conditions, $(P_{\text{max} (1, r/\delta)})$ for polynomial GARCH or $(L_r)$ for exponential GARCH respectively, translate differently in the various examples.

Table 1: Presentation of the volatility equation [4] and the corresponding specifications of functions $g_i, c_j$ for selected augmented GARCH models.

| Polynomial GARCH | standard formula for $\Lambda(\hat{\sigma}_t^2)$ | corresponding specifications of $g_i, c_j$ in [4] |
|------------------|-----------------------------------------------|-----------------------------------------------|
| APGARCH family   | $\sigma_t^{2g} = \omega + \sum_{i=1}^{p} \alpha_i \left( |y_{t-i}| - \eta_i |y_{t-i}| \right)^2 + \sum_{j=1}^{q} \beta_j \sigma_t^{2j-} \quad g_i = \omega/p$ and $c_j = \alpha_j \left( |e_{t-j}| - \gamma_j \epsilon_{t-j} \right)^2 + \beta_j$ |
| AGARCH           | $\sigma_t^{2g} = \omega + \sum_{i=1}^{p} \alpha_i \left( |y_{t-i}| - \eta_i |y_{t-i}| \right)^2 + \sum_{j=1}^{q} \beta_j \sigma_t^{2j-} \quad c_j = \alpha_j \left( |e_{t-j}| - \gamma_j \epsilon_{t-j} \right)^2 + \beta_j$ |
| GJR-GARCH        | $\sigma_t^{2g} = \omega + \sum_{i=1}^{p} \alpha_i \left( |y_{t-i}| - \eta_i |y_{t-i}| \right)^2 + \sum_{j=1}^{q} \beta_j \sigma_t^{2j-} \quad c_j = \beta_j + \alpha_j \epsilon_{t-j}^2 + \gamma_j^* \max(0, -\epsilon_{t-j})^2$ |
| GARCH            | $\sigma_t^{2g} = \omega + \sum_{i=1}^{p} \alpha_i \sigma_t^{2i-} + \sum_{j=1}^{q} \beta_j \sigma_t^{2j-} \quad c_j = \alpha_j \epsilon_{t-j}^2 + \beta_j$ |
| AR/ARCH          | $\sigma_t^{2g} = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 \quad c_j = \alpha_j \epsilon_{t-j}^2$ |
| TGARCH           | $\sigma_t = \omega + \sum_{i=1}^{p} \left( \alpha_i \max(y_{t-i}, 0) - \alpha_i^- \min(y_{t-i}, 0) \right) + \sum_{j=1}^{q} \beta_j \sigma_t \quad c_j = \alpha_j e_{t-j}^\epsilon - \gamma_j \epsilon_{t-j} + \beta_j$ |
| TSGARCH          | $\sigma_t = \omega + \sum_{i=1}^{p} \alpha_i |y_{t-i}| + \sum_{j=1}^{q} \beta_j \sigma_t \quad c_j = \alpha_j e_{t-j}^\epsilon + \beta_j$ |
| PGARCH           | $\sigma_t^{4g} = \omega + \sum_{i=1}^{p} \alpha_i |y_{t-i}|^4 + \sum_{j=1}^{q} \beta_j \sigma_t^{4j-} \quad c_j = \alpha_j e_{t-j}^\epsilon + \beta_j$ |
| VGARCH           | $\sigma_t^{2g} = \omega + \sum_{i=1}^{p} \alpha_i (\epsilon_{t-i} - \gamma_i^*)^2 + \sum_{j=1}^{q} \beta_j \sigma_t^{2j-} \quad g_i = \omega/p + \alpha_i (\epsilon_{t-i} - \gamma_i^*)^2$ and $c_j = \beta_j$ |
| NGARCH           | $\sigma_t^{2g} = \omega + \sum_{i=1}^{p} \alpha_i (y_{t-i} + \gamma_i \sigma_{t-i})^2 + \sum_{j=1}^{q} \beta_j \sigma_t^{2j-} \quad g_i = \omega/p$ and $c_j = \alpha_j (\epsilon_{t-i} + \gamma_i^\epsilon)^2 + \beta_j$ |

Exponential GARCH

|                      | standard formula for $\Lambda(\hat{\sigma}_t^2)$ | corresponding specifications of $g_i, c_j$ in [4] |
|----------------------|-----------------------------------------------|-----------------------------------------------|
| MGARCH               | $\log(\sigma_t^2) = \omega + \sum_{i=1}^{p} \alpha_i \log(\epsilon_{t-i}^\epsilon) + \sum_{j=1}^{q} \beta_j \log(\sigma_t^{2j-}) \quad g_i = \omega/p + \alpha_i \log(\epsilon_{t-i}^\epsilon)$ |
| EGARCH               | $\log(\sigma_t^2) = \omega + \sum_{i=1}^{p} \alpha_i (\epsilon_{t-i} - \mathbb{E}[\epsilon_{t-i}]) + \gamma_i \epsilon_{t-i} + \sum_{j=1}^{q} \beta_j \log(\sigma_t^{2j-}) \quad g_i = \omega/p + \alpha_i (\epsilon_{t-i} - \mathbb{E}[\epsilon_{t-i}]) + \gamma_i \epsilon_{t-i}$ |

For this, we introduce in Table [1] a non-exhaustive selection of different augmented GARCH(p,q) models, providing for each the corresponding volatility equation, [4], and the specifications of the functions $g_i$ and $c_j$. We consider 10 models that belong to the group of polynomial GARCH ($\Lambda(x) = x^p$) and two examples of exponential GARCH ($\Lambda(x) = \log(x)$). Note that in Table [1] the specification of $g_i$ is the same for the whole APGARCH family (only the $c_j$ change), whereas for the two exponential GARCH models, it is the reverse. The general restrictions on the parameters are as follows: $\omega > 0$, $\alpha_i \geq 0$, $-1 \leq \gamma_i \leq 1$, $\beta_j \geq 0$ for $i = 1, ..., p$, $j = 1, ..., q$. Further, the parameters in the GJR-GARCH (TGARCH) are denoted with an asterix (with a plus or minus) as they are not the same as in the other models.

In Tables [2] and [3] we present how the conditions $(P_{\text{max} (1, r/\delta)})$ or $(L_r)$ translate for each model. Table [2] treats the specific case of an augmented GARCH(p,q) process with $p = q = 1$, whereas Table [3] treats the general case for
for any $p \geq 1, q \geq 0$.

Table 2: Conditions ($P_{\max(1,r/\delta)}$) or ($L_r$) respectively translated for different augmented GARCH(1,1) models. Left column for the general $r$-th absolute centred sample moment, middle for the MAD ($r = 1$) and right for the variance ($r = 2$).

| augmented GARCH (1, 1) | $r \in \mathbb{N}$ | $r = 1$ | $r = 2$ |
|------------------------|---------------------|---------|---------|
| APGARCH                | $E[|\alpha_1 (|\epsilon_t - 1|)^{2\delta} + \beta_1 |\epsilon_t - 1|] < 1$ | $\alpha_1 E[|\epsilon_t - 1|^{2\delta}] + \beta_1 < 1$ | $E[|\alpha_1 (|\epsilon_t - 1|)^{2\delta} + \beta_1 |\epsilon_t - 1|] < 1$ |
| AGARCH                 | $E[|\alpha_1 (|\epsilon_t - 1|)^2 + \beta_1 |\epsilon_t - 1|] < 1$ | $\alpha_1 E[|\epsilon_t - 1|^{2\delta}] + \beta_1 < 1$ | $E[|\alpha_1 (|\epsilon_t - 1|)^2 + \beta_1 |\epsilon_t - 1|] < 1$ |
| GJR-GARCH              | $E[|\alpha_j \epsilon_t^2 + \beta_1 + \gamma_1^* \max(0, -\epsilon_t^2)|\epsilon_t - 1|] < 1$ | $\alpha_j + \beta_1 + \gamma_1^* E[\max(0, -\epsilon_t^2)] < 1$ | $E[|\alpha_j \epsilon_t^2 + \beta_1 + \gamma_1^* \max(0, -\epsilon_t^2)|\epsilon_t - 1|] < 1$ |
| GARCH                  | $E[|\alpha_1 \epsilon_t^2 + \beta_1 |\epsilon_t - 1|] < 1$ | $\alpha_1 + \beta_1 < 1$ | $\alpha_1^2 E[\epsilon_t^4] + \alpha_1 \beta_1 + \beta_1^2 < 1$ |
| ARCH                   | $\alpha_j E[\epsilon_t^2] < 1$ | $\alpha_1 < 1$ | $\alpha_1^2 E[\epsilon_t^4] < 1$ |
| TGARCH                 | $E[|\alpha_1 |\epsilon_t - 1| - \alpha_1 \gamma_1 |\epsilon_t - 1| + \beta_1 |\epsilon_t - 1|] < 1$ | $\alpha_1 E[|\epsilon_t - 1|] + \beta_1 < 1$ | $E[|\alpha_1 |\epsilon_t - 1| - \alpha_1 \gamma_1 |\epsilon_t - 1| + \beta_1 |\epsilon_t - 1|] < 1$ |
| TSGARCH                | $E[|\alpha_1 |\epsilon_t - 1| + \beta_1 |\epsilon_t - 1|] < 1$ | $\alpha_1 E[|\epsilon_t - 1|] + \beta_1 < 1$ | $E[|\alpha_1 |\epsilon_t - 1| + \beta_1 |\epsilon_t - 1|] < 1$ |
| PGARCH                 | $E[|\alpha_1 |\epsilon_0| + \beta_1 |\epsilon_0|] < 1$ | $\alpha_1 + 2\alpha_1 \beta_1 E[|\epsilon_0|] + \beta_1^2 < 1$ | $E[|\alpha_1 |\epsilon_0| + \beta_1 |\epsilon_0|] < 1$ |
| VGARCH                 | for any $r \in \mathbb{N}$: $\beta_1 < 1$ | $\alpha_1^{(1 + \gamma_1^2)} + \beta_1 < 1$ | $E[|\alpha_1 |\epsilon_0| + \beta_1 |\epsilon_0|] < 1$ |
| NGARCH                 | $E[|\alpha_1 |\epsilon_0| + \gamma_1 |\epsilon_0| + \beta_1 |\epsilon_0|] < 1$ | $\alpha_1^{(1 + \gamma_1^2)} + \beta_1 < 1$ | $E[|\alpha_1 |\epsilon_0| + \gamma_1 |\epsilon_0| + \beta_1 |\epsilon_0|] < 1$ |
| MGARCH                 | for any $r \in \mathbb{N}$: $E[\exp(4r|\omega/p + \alpha_1 \log(|\epsilon_0^2|)^2)] < \infty$ and $|\beta_1| < 1$ | $E[|\alpha_1 |\epsilon_0| + \gamma_1 |\epsilon_0| + \beta_1 |\epsilon_0|] < 1$ |
| EGARCH                 | for any $r \in \mathbb{N}$: $E[\exp(4r|\omega/p + \alpha_1 (|\epsilon_0| - E[|\epsilon_0|]) + \gamma_1 |\epsilon_0|)] < \infty$ and $|\beta_1| < 1$ |

In Table 2, we consider in the first column the conditions for the general $r$-th absolute centred sample moment, $r \in \mathbb{N}$. We also specifically look at the standard cases of the sample MAD ($r = 1$) and the sample variance ($r = 2$) as measure of dispersion estimators respectively, presented in the second and third column.

For the selected polynomial GARCH models, the requirement $\sum_{j=1}^{p} \|\beta_j(\epsilon_0)\|_{max(1,r/\delta)} < \infty$ in condition ($P_{\max(1,r/\delta)}$) will always be fulfilled. Thus, we only need to analyse the condition

$$\sum_{j=1}^{q} \|c_j(\epsilon_0)\|_{max(1,r/\delta)} < 1.$$

Lastly, we present in Table 3 how the conditions ($P_{\max(1,r/\delta)}$) or ($L_r$) respectively translate for those augmented GARCH($p,q$) processes - this is the generalization of Table 2. As, in contrast to Table 2, we do not gain any insight by considering the choices of $r = 1$ or $r = 2$, we only present the general case, $r \in \mathbb{N}$.

When $p \neq q$, we need to consider coefficients $\alpha_j, \beta_j, \gamma_j$ for $j = 1, \ldots, \max(p,q)$. In case they are not defined, we set them equal to 0.

Note that, in Table 2 (and also Table 3), the restrictions on the parameter space, given by ($P_{\max(1,r/\delta)}$) or ($L_r$) respectively, are the same as the conditions for univariate FCLTs of the process $X_t^2$ itself (see [5], [34]). For $r = 1$, they coincide with the conditions for e.g. $\beta$-mixing with exponential decay (see [16]).
Table 3: Conditions \( (P_{\max (1, r/\delta)}) \) or \( (L_r) \) respectively translated for different augmented GARCH\((p,q)\) models for the general \( r \)-th absolute centred sample moment, \( r \in \mathbb{N} \).

| augmented GARCH \((p,q)\) | \( r \in \mathbb{N} \) |
|---------------------------|------------------|
| APGARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |\alpha_j (|\epsilon_0| - \gamma_j \epsilon_{t-j})^2 + \beta_j |^{r/\delta} < 1 \) |
| AGARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |(\alpha_j |\epsilon_0| - \gamma_j \epsilon_{t-j})^2 + \beta_j |^{1/r} < 1 \) |
| GJR-GARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |\alpha_j^* \epsilon_0^2 + \beta_j + \gamma_j^* \max(0, -\epsilon_0^2)|^{1/r} < 1 \) |
| GARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |(\alpha_j^2 + \beta_j)^{1/r} < 1 \) |
| ARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} \alpha_j (|\epsilon_0|)^{1/r} < 1 \) |
| TARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |(\alpha_j \epsilon_{t-j} - \alpha_j \gamma_j \epsilon_{t-j} + \beta_j |)^{1/r} < 1 \) |
| TSGARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |(\alpha_j \epsilon_{t-j} + \beta_j |)^{1/r} < 1 \) |
| PGARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |\alpha_j |\epsilon_0| + \beta_j |^{2r/(2r)} < 1 \) |
| VGARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |\alpha_j |\epsilon_0| + \gamma_j |^{1/r} < 1 \) |
| NGARCH \( \sum_{j=1}^{\max(p,q)} \mathbb{E} |(\alpha_j |\epsilon_0| + \gamma_j)^2 + \beta_j |^{1/r} < 1 \) |
| MGARCH \( \mathbb{E} \exp(4\sum_{i=1}^{p} |\omega/p + \alpha_i \log(\epsilon_0^2)|^2) < \infty \) and \( \sum_{j=1}^{\max(p,q)} |\beta_j| < 1 \) |
| EGARCH \( \mathbb{E} \exp(4\sum_{i=1}^{p} |\omega/p + \alpha_i (|\epsilon_0| - \mathbb{E} |\epsilon_0|)| + \gamma_i |\epsilon_0|)^2 < \infty \) and \( \sum_{j=1}^{\max(p,q)} |\beta_j| < 1 \) |