ASYMPTOTICS OF THE QMLE FOR THE GENERAL ARCH(q) MODEL

Dennis Kristensen
Dept. of Economics, University of Wisconsin.
E-mail: dkristen@ssc.wisc.edu

AND

Anders Rahbek
Dept. of Applied Mathematics and Statistics, University of Copenhagen.
Tel: +45 3532 0682
E-mail: rahbek@stat.ku.dk

THIS VERSION: APRIL 25, 2005
PRELIMINARY - COMMENTS ARE WELCOME
Corresponding author:
Anders Rahbek

Department of Applied Mathematics and Statistics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark.
E-mail: rahbek@stat.ku.dk
Abstract: The asymptotic properties of the quasi-maximum likelihood estimator (QMLE) for the general ARCH(q) model are derived under weak regularity conditions. To show consistency, we require the ARCH-process to be geometrically ergodic, the conditional variance function to have a log-moment, and 2nd moment of the rescaled errors. We allow for a non-compact parameter space. Asymptotic normality of the estimator is established under the additional assumption that certain ratios involving the conditional variance function are suitably bounded, and that the rescaled errors have little more than 4th moment. We verify our general conditions for a range of specific ARCH-models.
ARCH/GARCH-models are widely used in empirical finance to explicitly model the behaviour of the conditional variance of time series, see Bera and Higgins (1993) for an overview. This is of great relevance when explaining and modelling risk and uncertainty of financial time series such as stock returns, exchange rates and interest rates. The estimation of these models are mostly done by conditional maximum-likelihood estimation under the assumption of Normal errors. This estimator proves to be robust to departures from Normality and is therefore normally referred to as the quasi-maximum-likelihood estimator (QMLE). This feature of the estimator is important since empirical evidence indicates that the errors in most cases are not Normally distributed.

A large number of different ARCH models can be found in the literature, each having been proposed to capture certain characteristics of the data. While the number of models has been growing over the years, rigorous results concerning the behaviour of the QMLE for the individual model are fairly limited. Up until recently, the only ARCH model for which the asymptotic properties of the QMLE had been fully investigated was the linear ARCH-model as proposed by Engle (1982) and extended by Bollerslev (1986). Results for this specific model can be found in e.g. Lee and Hansen (1994), Lumsdaine (1996), Berkes et al (2003), Weiss (1986), Jensen and Rahbek (2004a,b). As a first step to extend these results to the general ARCH-model, Kristensen and Rahbek (2005) showed strong consistency and asymptotic normality of the QMLE for a class of ARCH-models.

In this paper, we generalize these results to the fully general ARCH-model. Under weak regularity conditions, we show consistency and asymptotic normality of the QMLE. To demonstrate the usefulness of our general results, we verify that the imposed conditions hold for most ARCH-models which have been proposed in the literature. Our results confirm the commonly held belief that the QMLE is well-behaved as long as the ARCH model has a stationary solution with suitable moments.\footnote{The stationarity and moment assumptions are in some cases actually not needed, c.f. Jensen and Rahbek (2004a,b)}

Our strategy of proof borrows from Kristensen and Rahbek (2005), but involves some non-trivial extensions of their results. As they do, we rely on the concept of geometrically ergodic Markov chains where we assume that the observed ARCH-process can be imbedded within a such. This allows us to draw upon a number of strong results found in the Markov chain literature as exposed in Meyn and Tweedie (1993). In particular, we are given a Law
of Large Numbers (LLN) and a Central Limit Theorem (CLT) which hold irrespective of the starting value of the process. Thus, we do not need to assume that we have observed the stationary version of the process in contrast to for example Jeantheau (1998), or verify that the observed likelihood function converges uniformly towards the stationary version as done in e.g. Lee and Hansen (1994). Geometric ergodicity may appear to be a strong assumption to impose on the ARCH-process. However, in Section 5 we verify that this property actually holds for most of the ARCH specifications found in the literature under weak regularity conditions on the rescaled errors and weak parameter restrictions.

As an additional feature compared to most papers concerned with asymptotics of estimators in non-linear time series models, we allow for a non-compact parameter space. This means we are able to define the parameter space in a more unrestrictive manner than found in most results on estimation in nonlinear models.

Throughout the paper we will use the following notation: For any function \( f : \mathbb{R}^p \times \Theta \mapsto \mathbb{R}^q \), we shall write the partial derivative w.r.t. \( \theta \) as \( Df(x, \theta) = (D_j(f_i(x, \theta)))_{i,j} \) where \( D_jf_i = \left( \frac{\partial f_i(x, \theta)}{\partial j} \right) \); the second order partial derivatives are written as \( D^2 f(x, \theta) = (D_{jk}f_i(x, \theta))_{i,j,k} \) where \( D_{jk}f_i = \left( \frac{\partial^2 f_i(x, \theta)}{\partial j \partial k} \right) \). We will write \( \log^+(x) = \max \{ \log(x), 0 \} \) and \( \log^-(x) = \max \{-\log(x), 0\} \). Also, \( C \) and \( C_i, i = 1, 2, \ldots \), will denote generic constants, and \( \text{int}A \) and \( \overline{A} \) the interior and closure respectively of any set \( A \). For any process \( X_t \), use \( \{X_t\} \) to denote the infinite sequence \( X_1, X_2, \ldots \).

2. The QMLE of the General ARCH Model

Assume that we have observed \( \{y_t\}_{t=1}^n \) with \( y_t \in \mathbb{R} \) from the following the ARCH model,

\[
y_t = h_t^{1/2} z_t, \tag{1}
\]

\[
h_t = H(x_{t-1}; \theta), \tag{2}
\]

where \( x_{t-1} = (y_{t-1}, \ldots, y_{t-q}) \in \mathbb{R}^q, \theta \in \Theta \subseteq \mathbb{R}^d, \) and \( H : \mathbb{R}^q \times \Theta \mapsto \mathbb{R}_+ \). The function \( H \) is assumed to be known up to the parameter \( \theta \) which we then wish to estimate based on the \( n \) observations. In Section 5, we consider specific choices of \( H \) as proposed in the literature. We take the starting value \( x_0 \) as fixed in the following, but our results can be extended to allow for \( x_0 \) being random as long as its distribution does not depend on \( \theta \).

For the asymptotic analysis, we assume that for \( \theta \) at the true value \( \theta_0 \) there exists a stationary solution to equations (1)-(2) and denote this \( \{y_t^*\} \); the associated stationary vector process of lagged values is denoted \( \{x_t^*\} \). However, the observed process \( \{y_t\} \) may
be non-stationary as \( x_0 \) may not have been drawn from the stationary distribution. To emphasize this, let in the following \( h_{0t} = H(x_{t-1}; \theta_0) \) and \( h^*_{0t} = H(x^*_{t-1}; \theta_0) \) denote the true conditional variance process corresponding to \( \{x_t\} \) and \( \{x^*_t\} \) respectively. Correspondingly, write \( h_t = H(x_{t-1}; \theta) \) and \( h^*_t = H(x^*_{t-1}; \theta) \) for any \( \theta \neq \theta_0 \).

We here strengthen the assumption of the existence of a stationary solution to the one that \( \{x_t\} \) is a geometrically ergodic Markov chain: Under the assumption that the innovation process \( f_z \) is i.i.d., \( \{x_t\} \) is obviously a Markov chain with transition probability \( P^t(x, A) = P \{x_t \in A \mid x_0 = x\} \). We then say that \( \{x_t\} \) is \( V \)-geometrically ergodic if \( \|P^t(x, \cdot) - \pi\|_V \leq \rho^t M(x), x \in \mathbb{R}^d \), for some \( \rho < 1 \) and some functions \( M \) and \( V \geq 1 \), where \( \|\cdot\|_V \) is the total variation norm for any probability measure \( \nu \),

\[
\|\nu\|_V = \sup \left\{ \left| \int_{\mathbb{R}^d} f(x) \nu(dx) \right| : |f| \leq V \right\}.
\]

This property in turn gives us a LLN and a CLT which holds irrespective of the starting condition \( x_0 \), c.f. Meyn and Tweedie (1993, Theorem 17.0.1). Geometric ergodicity implies \( \beta \)-mixing with geometrically decaying mixing coefficients, c.f. Davydov (1973).

The estimator of interest is \( \hat{\theta} \) defined by

\[
\hat{\theta} = \arg \inf_{\theta \in \Theta} L_n(\theta),
\]

where

\[
L_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} l(y_t|x_{t-1}; \theta), \quad l(y_t|x_{t-1}; \theta) = \log(h_t) + \frac{y_t^2}{h_t}.
\]

If \( z_t \sim \text{i.i.d.} N(0, 1) \), \( \hat{\theta} \) is the actual MLE, but we will not impose this assumption, and \( \hat{\theta} \) is therefore a QMLE. We furthermore define the associated moment function \( L(\theta) = E \left[l(y_t^*|x^*_{t-1}; \theta)\right] \), which takes values on the extended real line \( \mathbb{R} \cup \{-\infty, +\infty\} \). Under the conditions given in the following section \( L(\theta) \) is well-defined.

In the next two sections, we give conditions for the QMLE to be consistent and asymptotically normally distributed respectively.

3. Consistency

Our first main result is that the QMLE \( \hat{\theta} \) is strongly consistent. In order to allow for non-compact \( \Theta \) when proving consistency, we adopt the strategy of Kristensen and Rahbek (2005). The idea is to find a compact set \( C_0 \subseteq \Theta \) and a lower bound for \( H(x; \theta) \) outside of this set, and use these to verify that \( \inf_{\theta \in \Theta \setminus C_0} L_n(\theta) \geq L(\theta_0) + \varepsilon \) a.s. as \( n \to \infty \).
for some $\varepsilon > 0$. If $L(\theta_0) < +\infty$, it then follows that $\hat{\theta} \in C_0$ a.s. as $n \to \infty$, and we can apply a general consistency result for geometrically ergodic Markov chains as stated in Kristensen and Rahbek (2005, Proposition 2) on the compact set $C_0$ under certain regularity conditions. These regularity conditions are verified in the lemmas found in Appendix B under the following set of conditions on the ARCH model:

C.1 The innovations $\{z_t\}$ are i.i.d. $(0,1)$.

C.2 For the true parameter $\theta_0 \in \Theta$, $\{x_t\}$ is $V$-geometrically ergodic.

C.3 $E\left[ \log^+ (H (x_t^*; \theta_0)) \right] < \infty$ and there exists a function $H_1 : \mathbb{R}^q \mapsto \mathbb{R}_+$ such that $H (x; \theta) \geq H_1(x), \theta \in \Theta$ and $E\left[ \log^+ (H_1 (x_{t-1}^*)) \right] < \infty$.

C.4 Either $\Theta$ is compact, or there exists a compact set $C_0 \subset \Theta$ and a function $H_2 : \mathbb{R}^q \mapsto \mathbb{R}_+$ such that

1. $H (x; \theta) \geq H_2(x)$ for all $(x, \theta) \in \mathbb{R}^q \times \Theta \setminus C_0$.

2. $E\left[ \log H_2 (x_{t-1}^*) \right] > L(\theta_0)$.

C.5 $P\left( H (x_{t-1}^*; \theta) \neq H (x_{t-1}^*; \theta_0) \right) > 0$ for any $\theta \neq \theta_0$.

The assumptions made in (C.1) about the rescaled errors are fairly standard. The i.i.d. assumption is needed to obtain the Markov property of $\{x_t\}$. We conjecture that our results will go through while allowing for dependence in $\{z_t\}$, for example assuming it to be a stationary martingale difference sequence, c.f. Lee and Hansen (1994). This would however invalidate the Markov property, and parts of our proofs would have to be changed since these rely on this. The Markov property could still be maintained while allowing for limited dependence in the $\{z_t\}$ process though. For example, by assuming that $\{(z_t, \ldots, z_{t-p})\}$ is a geometrically ergodic Markov chain for some finite $p \geq 1$, and satisfying $E[z_t|\mathcal{F}_{t-1}] = 0$ and $E[z_t^2|\mathcal{F}_{t-1}] = 1$ where $\mathcal{F}_{t-1} = \mathcal{F}(z_{t-1}, \ldots, z_{t-p})$. For simplicity, we maintain the i.i.d. assumption however.

Condition (C.2) ensures that the Markov chain $\{x_t\}$ is geometrically ergodic; in particular it implies that there exists a stationary solution $\{x_t^*\}$. Explicit conditions for which (C.4) holds in some leading examples are given in Section 5. The advantage of geometric ergodicity over the alternative assumption of the existence of a strictly stationarity solution is that the LLN and CLT hold irrespective of the initial condition. This in turn means we
do not have to check that the quasi-likelihood for the observed, potentially non-stationary, process converges to the stationary version of the quasi-likelihood at a sufficiently high rate; this approach is pursued in Lee and Hansen (1994).

(C.3) ensures that the conditional variance \( h_t = H(x_{t-1}; \theta) \) is well-defined and restricts it such that \( L(\theta_0) < \infty \) and \( L(\theta) > -\infty \) for all \( \theta \in \Theta \). Observe that we allow for the possibility of \( L(\theta) = +\infty \) for \( \theta \neq \theta_0 \).

(C.4) allows for a non-compact parameter space \( \Theta \) by imposing a uniform lower bound on the conditional variance function outside of a compact set \( \mathcal{C}_0 \) of the parameter space containing \( \theta_0 \). The restriction that \( H(x; \theta) \) is continuous in \( \theta \) can be weakened to lower semicontinuity. However, the models found in the literature all satisfy the continuity assumption. If \( \Theta \) is chosen to be compact, (C.4) is redundant since one can choose \( \mathcal{C}_0 = \Theta \).

Condition (C.5) ensures identification.

**Theorem 1.** Under (C.1)-(C.5), the QMLE defined by (3) satisfies \( \hat{\theta} \xrightarrow{a.s.} \theta_0 \).

### 4. Asymptotic Normality

In this section we show that the QMLE is asymptotically normally distributed. The proof of asymptotic normality is standard, based on a Taylor expansion of the score,

\[
S_n(\theta) = n^{-1} \sum_{i=1}^{n} s(y_t|x_{t-1}; \theta), \quad s(y_t|x_{t-1}; \theta) = Dl(y_t|x_{t-1}; \theta),
\]

around \( \theta_0 \); see e.g. Basawa, Feigin and Heyde (1976). In particular, we show in Lemma 6 that \( \sqrt{n} S_n(\theta_0) \xrightarrow{D} N(0, \Omega(\theta_0)) \) where

\[
\Omega(\theta) = E \left[ s(y^*_t|x^*_t; \theta) s(y^*_t|x^*_t; \theta)^\top \right] = \kappa_4 E \left[ \frac{Dh_t^* Dh_t^{*\top}}{h_t^2} \right],
\]

and furthermore in Lemma 5 that uniformly over \( \theta \) in a neighborhood \( \mathcal{N}_0 \) of \( \theta_0 \),

\[
\Lambda_n(\theta) = DS_n(\theta) \xrightarrow{a.s.} \Lambda(\theta) = E \left[ \frac{Dh_t^* Dh_t^{*\top}}{h_t^2} \right].
\]

To establish these results, we assume in addition to (C.1)-(C.5) that:

**N.1** \( \kappa_4 = E[(z_t^2 - 1)^2] \in (0, \infty) \)

**N.2** There exists a neighbourhood \( \mathcal{N}_0 \) of \( \theta_0 \) with \( \mathcal{N}_0 \subseteq \Theta \) and functions \( b_i : \mathbb{R}^{q+1} \mapsto \mathbb{R}_+ \), \( i = 1, 2 \), such that \( \theta \mapsto H(x; \theta) \) is twice continuously differentiable for any \( (x, \theta) \in \mathbb{R}^q \times \mathcal{N}_0 \) with (i) \( E \left[ \|s(y^*_t|x^*_{t-1}; \theta)\|^{2+\delta} \right] < \infty \) for some \( \delta > 0 \), (ii) \( \|Ds(y|x; \theta)\|^2 \leq b_1(y|x) \) and (iii) \( \|s(y|x; \theta)\| \leq b_2(y|x), \theta \in \mathcal{N}_0 \), where \( E \left[ b_1(y^*_t|x^*_t-1) \right] < \infty \).
Condition (N.1) is needed as part of the proof showing that the variance of the quasi-score, defined below, exists and is non-degenerate. Under (N.2), \( \theta_0 \in \text{int}\Theta \), and the squared quasi-score function evaluated at \( \theta = \theta_0 \) has a finite \((2 + \delta)\)th moment; this ensures that we can employ a CLT for \( V \)-geometrically ergodic Markov chains. The bound in (ii) ensures that (7) holds, while (iii) is used to prove consistency of the asymptotic variance estimator. So indeed (iii) can be left out if one is only interested in establishing asymptotic normality.

The following simple condition (N.2') implies (N.2), c.f. Lemma 7. This can be used for most of the standard models.

**N.2'** \( E[z_t^{4+\delta}] < \infty \) for some \( \delta > 0 \) and

\[
EH (x; \theta_0) / H (x; \theta) \leq m_0 (x), \quad \| D^i H (x; \theta) \| / H (x; \theta) \leq m_i (x), \quad i = 1, 2,
\]

where

\[
E [m_1^{2+\delta} (x_t^*)] < \infty, \quad E [m_2^2 (x_t^*)] < \infty, \quad E [m_0 (x_t^*) \{m_3 (x_t^*) + 1\}] < \infty.
\]

**Theorem 2.** Under (C.1)-(C.5) and (N.1)-(N.2),

\[
\sqrt{n}(\hat{\theta} - \theta_0) \overset{D}{\to} N (0, \Lambda^{-1}(\theta_0) \Omega(\theta_0) \Lambda^{-1}(\theta_0))
\]

where \( \Omega(\theta_0) \) and \( \Lambda(\theta_0) \) given in (6) and (7) both are positive definite. Furthermore,

\[
\hat{\Omega} = \frac{1}{n} \sum_{t=1}^{n} s(y_t|x_{t-1}; \hat{\theta}) s(y_t|x_{t-1}; \hat{\theta})',
\]

\[
\hat{\Lambda} = \frac{1}{n} \sum_{t=1}^{n} Ds(y_t|x_{t-1}; \hat{\theta})
\]

are strongly consistent estimators of \( \Omega(\theta_0) \) and \( \Lambda(\theta_0) \) respectively.

**Remark 1.** In the case where \( z_t \sim i.i.d. N(0, 1) \), the result collapses to the standard MLE result with \( \kappa_4 = 2 \) and \( \Lambda(\theta_0) \) being the information matrix multiplied by 2.

5. **Examples**

In this section we consider a number of different ARCH-models proposed in the literature, and demonstrate that they under suitable conditions satisfy (C.2)-(C.5) and (N.2). The remaining assumptions concern the error structure ((C.1) and (N.1)). We will impose the following slightly stronger condition on the error distribution:
The innovations \( \{ z_t \} \) satisfy (C.1) and (N.1) with a marginal distribution given by a lower semicontinuous density \( f \) w.r.t. the Lebesgue measure, and \( E[z_t^{1+\delta}] < \infty \).

This assumption is useful when establishing \( V \)-geometric ergodicity of \( \{ x_t \} \) and identification of \( \theta_0 \). For geometric ergodicity, (C.Z) is used to verify so-called irreducibility of the Markov chain w.r.t. the Lebesgue measure. That is, \( \{ x_t \} \) does not degenerate to some lower-dimensional subspace of \( \mathbb{R}^q \), and its transition density is absolutely continuous w.r.t. the Lebesgue measure. This in turn can be used to verify the identification of \( \theta_0 \).

In what follows, we will throughout implicitly assume that (C.Z) holds and that \( \theta_0 \in \text{int}\Theta \). In the specific examples, we propose drift functions \( V(x) \). These are chosen as "mild" as possible such that provided \( V \)-geometric ergodicity is established for these, then the conclusions of Theorem 1 and 2, that is consistency and normality, hold. When actually showing geometric ergodicity for the examples, we then often choose a drift function \( V_d \) which dominates the "mild" or minimal one, \( V_d \geq V \), as \( V \) itself is seldomly operational.

**Linear ARCH.** The first example is the linear specification considered in Kristensen and Rahbek (2005) given by

\[
H(x; \theta) = \omega + \sum_{i=1}^{q} \sum_{j=1}^{p} \alpha_{ij} g_j(x_i),
\]

for functions \( g_j : \mathbb{R} \to [0,\infty) \). We choose the parameter space as

\[
\Theta = \left\{ \theta = (\omega, \alpha) \in \mathbb{R}^{1+q+p} | \omega \leq \omega, 0 \leq \alpha_{ij}, i = 1, \ldots, q, j = 1, \ldots, p \right\}.
\]

**Corollary 1.** Assume that (i) (C.2) holds with \( V(x) \geq 1 + \log \left( 1 + \sum_{i=1}^{q} \sum_{j=1}^{p} g_j(x_i) \right) \), and (ii) \( P \left( \sum_{i=1}^{q} \sum_{j=1}^{p} \beta_{ij} g_j(y_{t-i}) = c \right) = 0 \) for any \( \beta \in \mathbb{R}^{q+p} \) and \( c \in \mathbb{R} \). Then the QMLE is consistent and asymptotically normally distributed as given in Theorem 2.

A sufficient condition for (i) to hold is that \( g_j(x_i) \leq C_{1,j} + C_{2,j} x_i^2 \) and \( \sum_{i=1}^{q} \sum_{j=1}^{p} \alpha_{ij} C_{2,j} < 1 \).

Our conditions are slightly weaker than the ones in Kristensen and Rahbek (2005).

**Asymmetric Power ARCH (Ding et al, 1993; Hentschel, 1995).** The conditional variance is given by

\[
h_t^{\mu/2} = \omega + \sum_{i=1}^{q} \alpha_i (|y_{t-i} - b_i| + \gamma_i (y_{t-i} - b_i))^\mu.
\]

10
We define

$$\Theta = \{ \theta = (\omega, \mu, \alpha, \gamma) \in \mathbb{R}^{2+2q} | \omega \leq \omega, \mu \leq \mu, 0 \leq \alpha_i, -1 + \varepsilon \leq \gamma_i \leq 1 - \varepsilon, i = 1, ..., q \}$$

with \( \omega < 1 \), and \( \mu, \varepsilon > 0 \). The parameter \( b = (b_1, ..., b_q)' \in \mathbb{R}^q \) is assumed to be known. The reason for this is that \( h_t \) is not differentiable w.r.t. \( b \). To allow for \( b \) unknown and as such to be estimated from data, a different approach has to be pursued, see e.g. Pakes and Pollard (1989).

**Corollary 2.** Assume that (i) (C.2) and (C.5) hold with \( E [\log (1 + |y_t^\delta|)] < \infty \). Then the QMLE is consistent. If furthermore (ii) \( E [|y_t^\delta|] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2 hold.

A sufficient condition for (i)-(ii) to hold is

\[
q = 1 : \alpha \left\{ (1 + \gamma)\mu E \left[ |z_{t-1}|^\mu 1_{(z_{t-1}>0)} \right] + (1 - \gamma)\mu E \left[ |z_{t-1}|^\mu 1_{(z_{t-1}>0)} \right] \right\} < 1, \tag{10}
\]

\[
q > 1 : E \left[ |z_{t-1}|^\alpha \right] \sum_{i=1}^q \alpha_i \max \{ 1 + \gamma_i, 1 - \gamma_i \} < 1. \tag{11}
\]

**Asymmetric ARCH (Glosten, Jagannathan and Runkle, 1993).** The conditional variance is given by

$$h_t = \omega + \sum_{i=1}^q \left\{ \alpha_{1i} 1_{(y_{t-1}<0)} |y_{t-1}|^\mu + \alpha_{2i} 1_{(y_{t-1}>0)} |y_{t-1}|^\mu \right\}.$$

We define

$$\Theta = \{ \theta = (\omega, \mu, \alpha) \in \mathbb{R}^{2+q} | \omega \leq \omega, \mu \leq \mu, 0 \leq \alpha_i, i = 1, ..., q \}$$

with \( \omega < 1 \) and \( \mu > 0 \).

**Corollary 3.** Assume that (i) (C.2) and (C.5) hold with \( E [\log (1 + |y_t^\delta|)] < \infty \). Then the QMLE is consistent. If furthermore (ii) \( E [|y_t^\delta|] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2 hold.

A sufficient condition for (i)-(ii) to hold is that: \( \bar{\mu} < 2 \) or \( \bar{\mu} = 2 \) and

\[
q = 1 : \alpha_1 E \left[ \max (0, z_t)^2 \right] + \alpha_2 E \left[ \max (0, -z_t)^2 \right] < 1,
\]

\[
q > 1 : \sum_{i=1}^q \alpha_{1i} + \alpha_{2i} < 1.
\]
Power ARCH (Engle and Bollerslev, 1986; Schwert, 1989a,b; Taylor, 1986)

\[ h_t = \omega + \sum_{i=1}^{q} \alpha_i |y_{t-i}|^{\mu}. \]

Define

\[ \Theta = \{ \theta = (\omega, \mu, \alpha) \in \mathbb{R}^{2+q} | \omega \leq \omega, \mu \leq \bar{\mu}, 0 \leq \alpha_i, i = 1, \ldots, q \} \]

with \( \omega < 1 \) and \( \bar{\mu} > 0 \).

**Corollary 4.** Assume that (i) (C.2) and (C.5) hold with \( E \left[ \log (1 + |y_t^*|^\bar{\mu}) \right] < \infty \). Then the QMLE is consistent. If furthermore, (ii) \( E \left[ |y_t^*|^\delta \right] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2 hold.

A sufficient condition for (i)-(ii) to hold is that \( \bar{\mu} < 2 \) or \( \bar{\mu} = 2 \) and \( \sum_{i=1}^{q} \alpha_i < 1 \).

CDF-ARCH (Engle and Bollerslev, 1986): \( h_t = \omega + \sum_{i=1}^{q} \alpha_i \left( y_{2i}^2 - 1 \right) \). Define

\[ \Theta = \{ \theta = (\omega, \mu, \alpha) \in \mathbb{R}^{2+q} | \omega \leq \omega, \mu \leq \bar{\mu}, 0 \leq \alpha_i, i = 1, \ldots, q \} \]

with \( \omega < 1 \) and \( \bar{\mu} > 0 \).

**Corollary 5.** Assume that (i) (C.2) and (C.5) hold with \( V(x) \geq 1 \). Then the QMLE is consistent. If furthermore, (ii) \( E \left[ |y_t^*|^\delta \right] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2.

(i)-(ii) hold for any choice of parameter values.

NARCH (Bera and Higgins, 1992). \( h_t^\mu = \omega + \sum_{i=1}^{q} \alpha_i y_{i-1}^{2\mu} \). Define

\[ \Theta = \{ \theta = (\omega, \mu, \alpha) \in \mathbb{R}^{2+q} | \omega \leq \omega, \mu \leq \bar{\mu}, 0 \leq \alpha_i, i = 1, \ldots, q \} \]

with \( \omega < 1 \), \( \bar{\alpha}_i > 0 \) and \( \bar{\mu} > 0 \).

**Corollary 6.** Assume that (i) (C.2) and (C.5) hold with \( E \left[ \log (1 + |y_t^*|^{2\bar{\mu}}) \right] < \infty \). Then the QMLE is consistent. If furthermore, (ii) \( E \left[ |y_t^*|^\delta \right] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2 hold.

A sufficient condition for (i)-(ii) to hold is \( E \left[ |z_t|^\mu \right] \sum_{i=1}^{q} \alpha_i < 1 \).
VARCH (Engle, 1990; Engle and Ng, 1993). \( h_t = \omega + \sum_{i=1}^{q} \alpha_i (y_{t-i} + \gamma_i)^2 \). Define
\[
\Theta = \left\{ \theta = (\omega, \alpha, \gamma) \in \mathbb{R}^{1+2q} | \omega \leq \omega, 0 \leq \alpha_i, \gamma_i \in \mathbb{R}, i = 1, \ldots, q \right\},
\]
with \( \omega < 1 \).

Corollary 7. Assume that (i) (C.2) and (C.5) hold with \( E \left[ \log (1 + |y_t^*|^2) \right] < \infty \). Then the QMLE is consistent. If furthermore, (ii) \( E \left[ |y_t^*|^\delta \right] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2 hold.

A sufficient condition for (i)-(ii) to hold is that \( \sum_{i=1}^{q} \alpha_{0,i} < 1 \).

TSARCH/TARCH (Taylor, 1986; Schwert, 1989; Zakoian, 1994).
\[
\sqrt{h_t} = \omega + \sum_{i=1}^{q} \left\{ \alpha_{1,i} |y_{t-i}| + \alpha_{2,i} \max (y_{t-i}, 0) \right\}.
\]
Define
\[
\Theta = \left\{ \theta = (\omega, \alpha_1, \alpha_2) \in \mathbb{R}^{1+2q} | \omega \leq \omega, 0 \leq \alpha_{i,j} \leq \bar{\alpha}_{i,j}, i = 1, \ldots, q, j = 1, 2 \right\},
\]
with \( \omega < 1 \) and \( \bar{\alpha}_{i,j} > 0 \).

Corollary 8. Assume that (i) (C.2) holds with \( E \left[ \log (1 + |y_t^*|) \right] < \infty \). Then the QMLE is consistent. If furthermore, (ii) \( E \left[ |y_t^*|^\delta \right] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2 hold.

A sufficient condition for (i)-(ii) to hold is:
\[
q = 1 : \alpha_1 E \left[ |z_{t-1}| \right] + \alpha_2 E \left[ |z_{t-1}| 1(z_{t-1} > 0) \right] < 1
\]
\[
q > 1 : \sum_{i=1}^{q} \left\{ \alpha_{1,i} + \alpha_{2,i} \right\} < 1
\]

QARCH (Sentana, 1991). \( \sqrt{h_t} = \omega + \sum_{i=1}^{q} \gamma_i y_{t-i} + \sum_{i=1}^{q} \alpha_{ii} y_{t-i}^2 + 2 \sum_{i=1}^{q} \sum_{j=i+1}^{q} \alpha_{ij} y_{t-i} y_{t-j} \). Define
\[
\Theta = \left\{ \theta = (\omega, \alpha_1, \alpha_2) \in \mathbb{R}^{1+2q} | \omega \leq \omega, \ldots \right\},
\]
with \( \omega < 1 \).
Corollary 9. Assume that (i) (C.2) holds with \( E \left[ \log (1 + |y_t^*|^2) \right] < \infty \). Then the QMLE is consistent. If furthermore, (ii) \( E \left[ |y_t^*|^{\delta} \right] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2 hold.

A sufficient condition for (i)-(ii) to hold is that \( \sum_{i=1}^{q} \alpha_{0,i} \).

Log-ARCH (Geweke, 1986; Pantula, 1986). \( \log (h_t) = \omega + \sum_{i=1}^{q} \alpha_i \log (y_{t-i}^2) \). Define

\[ \Theta = \{ \theta = (\omega, \alpha) \in \mathbb{R}^{2+q} \mid \omega \leq \omega, \alpha_i \leq \alpha_i, i = 1, \ldots, q \} \]

and the characteristic polynomial,

\[ P(z; \theta) = 1 - \sum_{i=1}^{q} \alpha_i z^i, \]

and let \( \rho(\theta) = \max_{i=1,\ldots,q} |\lambda_i(\theta)| \) where \( \lambda_i(\theta) \) is the \( i \)th root of \( P \).

Corollary 10. Assume that (i) (C.2) holds with \( E [\| \log (y_t^*2) \|] < \infty \). Then the QMLE is consistent. If furthermore, (ii) \( E \left[ h_t^{\delta/2} \right] < \infty \), for some \( \delta > 0 \), then the conclusions of Theorem 2 hold.

A necessary and sufficient condition for (i) to hold is that \( E [\| \log (z_t^2) \|] < \infty \) and \( \rho(\theta_0) < 1 \). A necessary and sufficient condition for (ii) to hold is \( E \left[ z_t^{-\delta} \right] < \infty \) for some \( \delta > 0 \).

6. Conclusion

We have under relatively weak conditions shown consistency and asymptotic normality for the QMLE of the general ARCH model. Several extensions are of interest; for example, GARCH models, and multivariate extensions of such.

The asymptotic properties of the QMLE for the linear GARCH(1,1)-model is well-established, see Lee and Hansen (1994), Lumsdaine (1996), Jensen and Rahbek (2004a,b), but the nonlinear case has not been fully explored yet; the only study in this area known to the authors is Straum and Mikosch (2004).

Comtes and Lieberman (2003) derive the asymptotics of the QMLE for a specific class of multivariate GARCH-models, but do this under very strong moment conditions. Ling and McAleer (2003) restrict attention to another multivariate GARCH-model where there is constant cross-correlation, which allows for a relatively simple analysis.
7. Acknowledgements

Rahbek acknowledges continuing financial support from the Danish Social Sciences Research Council. Kristensen received funding from the Shoemaker Foundation during this research.
A. Proofs: Section 3 and 4

Proof of Theorem 1. We first establish that, \( \lim_{n \to \infty} \hat{\theta} \in \mathcal{C}_0 \) a.s. where \( \mathcal{C}_0 \) is the compact set given in (C.4) with \( \theta_0 \in \mathcal{C}_0 \). Observe first that by the strong LLN,

\[
\lim_{n \to \infty} L_n(\hat{\theta}) = \lim_{n \to \infty} \left[ \inf_{\theta \in \Theta} L_n(\theta) \right] \geq \lim_{n \to \infty} L_n(\theta_0) = L(\theta_0) \text{ a.s.} \tag{12}
\]

where \( L(\theta_0) < \infty \) by Lemma 4. Next, applying Lemma 3 and the definition of \( l(y_t|x_{t-1}) \) therein, together with the strong LLN gives,

\[
\lim_{n \to \infty} \left[ \inf_{\theta \in \Theta \backslash \mathcal{C}_0} L_n(\theta) \right] \geq \lim_{n \to \infty} \left[ \frac{1}{n} \sum_{i=1}^{n} l(y_{t_i}|x_{t_i-1}) \right] > L(\theta_0) \text{ a.s.} \tag{13}
\]

Combining (12) and (13), we obtain that \( \lim_{n \to \infty} \hat{\theta}_n \in \mathcal{C}_0 \) a.s. Hence we can restrict attention to \( \mathcal{C}_0 \) and redefine \( \hat{\theta} \) as \( \hat{\theta} = \arg \min_{\theta \in \mathcal{C}_0} L_n(\theta) \). The set \( \mathcal{C}_0 \) is compact and we can apply the general consistency result stated in Rahbek and Kristensen (2005, Proposition X); we check that the conditions given there are satisfied: The function \( l(y_t|x_{t-1}; \theta) = \log (h_t) + y_t^2/h_t \) is continuous in \( \theta \) for all \( (y_t, x_{t-1}) \); by Lemma 2, \( L(\theta_0) > -\infty \), so the moment is well-defined; identification of \( \theta_0 \) holds by Lemma 4. Finally, for any compact subset \( \mathcal{C} \subseteq \mathcal{C}_0 \), \( E[\inf_{\theta \in \mathcal{C}} l(y_t^*|x_{t-1}^*)] > -\infty \) by Lemma 2. We are therefore able to conclude that \( P(\lim_{n \to \infty} \hat{\theta}_n = \theta_0) = 1 \) as desired. \( \square \)

Proof of Theorem 2. The result is obtained by the following standard Taylor expansion of \( S_n(\hat{\theta}_n) \) in (5). Note first that by (N.2), \( S_n(\theta) \) is well-defined for all \( \theta \in \Theta \), and there exists a neighbourhood \( \mathcal{N}_0 \subseteq \Theta \) (which we assume is convex and bounded without loss of generality) such that \( \theta_0 \in \mathcal{N}_0 \). As \( n \to \infty \), \( \hat{\theta}_n \in \text{int}(\mathcal{N}_0) \) a.s. by Theorem 1, such that \( S_n(\hat{\theta}) = 0 \) by the definition of \( \hat{\theta}_n \) and hence,

\[
0 = S_n(\hat{\theta}) = S_n(\theta_0) + \Lambda_n(\hat{\theta} - \theta_0)
\]

where \( \Lambda_n(\theta) \) is given in (7) and \( \hat{\theta} = (\hat{\theta}_1, ..., \hat{\theta}_d)' \in \mathcal{N}_0 \) since \( \hat{\theta}_i \in [\hat{\theta}_i, \theta_{0,i}] \). Lemma 5 and 6 then imply,

\[
\sqrt{n}(\hat{\theta} - \theta_0) = \Lambda_n^{-1}(\theta_0)\sqrt{n}S_n(\theta_0) \xrightarrow{D} N(0, \Lambda^{-1}(\theta_0) \Omega(\theta_0) \Lambda^{-1}(\theta_0)).
\]

To prove that \( \hat{\Omega} \xrightarrow{a.s.} \Omega(\theta_0) \), we use the following inequality,

\[
||\hat{\Omega} - \Omega(\theta_0)|| \leq ||\hat{\Omega}_n - \Omega(\hat{\theta})|| + ||\Omega(\hat{\theta}) - \Omega(\theta_0)|| \leq \sup_{\theta \in \mathcal{N}_0} ||\Omega_n(\theta) - \Omega(\theta)|| + ||\Omega(\hat{\theta}) - \Omega(\theta_0)||.
\]
By Lemma 5, (i), \( \sup_{\theta \in \mathcal{N}_0} \| \Omega_n (\theta) - \Omega (\theta) \| \overset{a.s.}{\to} 0 \) with \( \Omega (\theta) \) continuous. Thus, \( \hat{\theta} \overset{a.s.}{\to} \theta_0 \) implies \( \| \Omega (\hat{\theta}_n) - \Omega (\theta_0) \| \overset{a.s.}{\to} 0 \). Similarly, \( \hat{\lambda} \overset{a.s.}{\to} \Lambda (\theta_0) \), by Lemma 5, (ii).

A.1. Lemmas

**Lemma 1.** Under (C.1)-(C.2):

1. \( \{Z_t\} \) defined by \( Z_t = (x_{t-1}, y_t) \) is \( V_Z \)-geometrically ergodic for any function \( V_Z \) satisfying
   \[
   \int V_Z (x, y) P (x, dy) \leq c V (x),
   \]
   for some constant \( c > 0 \), where \( P (x, dy) = P_z \left( H_0^{1/2} (x) dy \right) \) and \( P_z \) is the marginal distribution of \( z_t \). This holds in particular for \( V_Z (x, y) = 1 \).

2. \( \{Z_t\} \) defined by \( Z_t = (x_{t-1}, z_t) \) is \( V_Z \)-geometrically ergodic for any function \( V_Z \) satisfying
   \[
   \int V_Z (x, z) P_z (dz) \leq c V (x).
   \]
   This holds in particular for \( V_Z (x, y) = 1 \).

**Proof.** We apply Meitz and Saikkonen (2004, Proposition 1). In 1. we have

\[
\begin{align*}
y_t &= F_y (x_{t-1}, z_t), \\
x_{t-1} &= F_x (x_{t-2}, y_{t-1}),
\end{align*}
\]

where \( F_y (x_{t-1}, z_t) = H_0^{1/2} (x_{t-1}) z_t, F_x,1 (x_{t-2}, y_{t-1}) = y_{t-1} \), and \( F_{x,i} (x_{t-2}, y_{t-1}) = x_{i-1,t-2} \), \( i = 2, \ldots, q \). In 2.,

\[
\begin{align*}
z_t &= F_y (x_{t-1}, z_t), \\
x_{t-1} &= F_x (x_{t-2}, z_{t-1}),
\end{align*}
\]

where \( F_y (x_{t-1}, z_t) = z_t, F_x,1 (x_{t-2}, z_{t-1}) = H^{1/2} (x_{t-2}) z_{t-1} \), and \( F_{x,i} (x_{t-2}, y_{t-1}) = x_{i-1,t-2} \), \( i = 2, \ldots, q \).

**Lemma 2.** Under (C.1)-(C.5), \( L (\theta) = E \left[ l (y^*; x_{t-1}^*; \theta) \right] > -\infty \), \( \theta \in \Theta \). Furthermore, it holds that for any set \( K \subseteq C_0 \), \( E \left[ \inf_{\theta \in K} l (y^*; x_{t-1}^*; \theta) \right] > -\infty \).
Proof. We have \( l (y|x; \theta) \geq \log H (x; \theta) \geq \log H_1 (x) \), where \( E [\log ^- (H (x^*; \theta))] < \infty \) by (C.3) implying \( L (\theta) > -\infty \) which therefore is well-defined for any \( \theta \in \Theta \). Since \( H_1 \) does not depend on \( \theta \in \Theta \), we are also able to conclude that \( E [\inf_{\theta \in \mathcal{K}} l (y^*_t|x^*_{t-1}; \theta)] > -\infty \) for any set \( \mathcal{K} \subseteq \mathcal{C}_0 \).

\[ \square \]

Lemma 3. Under (C.1)-(C.5), there exists a function \( l \) such that \( \inf_{\theta \in \Theta \setminus \mathcal{C}_0} l (y_t|x_{t-1}; \theta) \geq l (y_t|x_{t-1}) \) a.s. with \( E [l (y^*_t|x^*_t)] > L (\theta_0) \), where \( \mathcal{C}_0 \) is given in (C.4).

Proof. Choose \( \mathcal{C}_0 \) as in (C.4) and define \( l (y|x) = \log (H_2 (x)) \) where \( H_2 \) is also given in (C.4). We then have for any \( \theta \in \Theta \setminus \mathcal{C}_0 \) that

\[ l (y_t|x_{t-1}; \theta) \geq \log (h_t) = \log (H (x_{t-1}; \theta)) \geq l (y_t|x_{t-1}). \]

Furthermore,

\[ E [l (y^*_t|x^*_t)] = E [\log (H_2 (x^*_t))] > L (\theta_0). \]

\[ \square \]

Lemma 4. Under (C.1)-(C.5), \( L (\theta_0) < \infty \) and \( L (\theta_0) \leq L (\theta) \) with equality if and only if \( \theta = \theta_0 \).

Proof. First observe that \( L (\theta_0) < \infty \): Since

\[ l (y^*_t|x^*_t; \theta_0) = \log (h^*_{0t}) + \frac{h^*_{0t}z^2_t}{h^*_{0t}} = \log (h^*_{0t}) + z^2_t, \]

it follows by (C.1) and (C.4), (ii), that

\[ E [\log (H (x^*_t; \theta_0))] \leq E [\log (H (x^*_t; \theta_0))] + 1 < \infty. \]

Let \( \theta \neq \theta_0 \) be given. Then by Lemma 2 either (i) \( L (\theta) = \infty \) or (ii) \( L (\theta) \in (-\infty, \infty) \). If (i) holds, \( L (\theta_0) < \infty = L (\theta) \). If (ii) holds, the following calculations are allowed:

\[ L (\theta) = E \left[ \log (h^*_{0t}) + \frac{y^*_{0t}z^2_t}{h^*_{0t}} \right] = E \left[ \log (h^*_{0t}) + \frac{h^*_{0t}z^2_t}{h^*_{0t}} \right] = E \left[ \log (h^*_{0t}) + \frac{h^*_{0t}}{h^*_t} \right], \]

where we made use of (C.1). From the last equality,

\[ L (\theta) - L (\theta_0) = E \left[ \log \left( \frac{h^*_{0t}}{h^*_t} \right) + \frac{h^*_{0t}}{h^*_t} \right] - 1 \geq 0 \]

with equality if and only if \( H (x^*_{t-1}; \theta_0) = h^*_{0t} = h^*_t = H (x^*_{t-1}; \theta) \) a.s.. This in turn only holds if \( \theta = \theta_0 \) by (C.5) such that \( L (\theta) = L (\theta_0) \) if and only if \( \theta = \theta_0 \).
Lemma 5. Under (C.1)-(C.5) and (N.1)-(N.2),

(i) The moment function $\theta \mapsto \Omega (\theta)$ given in (6) is well-defined and continuous for $\theta \in \mathcal{N}_0$, and $\sup_{\theta \in \mathcal{N}_0} \| \Omega_n (\theta) - \Omega (\theta) \| \xrightarrow{a.s.} 0$.

(ii) The moment function $\theta \mapsto \Lambda (\theta)$ given in (7) is well-defined and continuous for $\theta \in \mathcal{N}_0$, and $\sup_{\theta \in \mathcal{N}_0} \| \Lambda_n (\theta) - \Lambda (\theta) \| \xrightarrow{a.s.} 0$.

(iii) The matrix $\Lambda (\theta_0) = E \left[ Dh_{0t}^* (Dh_{0t}^*)_t h_{0t}^{*-2} \right]$ is non-singular.

Proof. First observe that $\theta \mapsto H (x_{t-1}; \theta)$ is twice continuously differentiable, so by the chain rule so is $\theta \mapsto l (y_t | x_{t-1}; \theta)$. To prove (i)-(ii), apply the uniform law of large numbers stated in Kristensen and Rahbek (2005, Proposition X) with $\mathcal{K} = \mathcal{N}_0$ (which we can assume is compact without loss of generality) and $f = s$ and $Ds$ respectively. The conditions stated there hold by the definition of $\mathcal{N}_0$, and the bounds in (N.2).

Finally, from (15) and (C.1),

$$
\Lambda (\theta_0) = E \left[ Ds \left( y_t^* | x_{t-1}; \theta_0 \right) \right] = E \left[ \frac{Dh_{0t}^* (Dh_{0t}^*)_t h_{0t}^{*-2}}{h_{0t}^{*-2}} \right].
$$

Assume that $\Lambda (\theta_0)$ is not positive definite; then there exists a vector $\lambda \in R^d \setminus \{0\}$ such that $\lambda^T D H (x_{t-1}^*; \theta_0) = 0$ a.s. since $h_{0t}^* > 0$ a.s. Using the mean value theorem, this implies that for some $\theta \neq \theta_0$ $H (x_{t-1}^*; \theta) = H (x_{t-1}^*; \theta_0)$ a.s. which is ruled out by (C.5).

Lemma 6. Under (C.1)-(C.5) and (N.1)-(N.2), \( \sqrt{n} S_n (\theta_0) \xrightarrow{D} N \left( 0, \Omega (\theta_0) \right) \).

Proof. This follows from Theorem 1 yielding

$$
\sqrt{n} S_n (\theta_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} f (x_{t-1}, z_t) \xrightarrow{D} N \left( 0, \Omega (\theta_0) \right),
$$

with $\Omega (\theta_0) = E \left[ s (y_t^* | x_{t-1}^*; \theta_0) s (y_t^* | x_{t-1}^*; \theta_0)^T \right]$ since $E \left[ f (x_{t-1}^*; z_t) | x_{t-1}^* \right] = 0$ by (C.1).

One easily check that $\Omega (\theta_0) = \kappa_4 \Lambda (\theta_0)$.

Lemma 7. (N.2), (i)-(iii) hold if for some $\delta > 0$: $E[z_t^{4+\delta}] < \infty$, and

$$
EH (x; \theta_0) / H (x; \theta) \leq m_0 (x), \quad \| D_i H (x; \theta) \| / H (x; \theta) \leq m_i (x), \quad i = 1, 2,
$$

where

$$
E \left[ m_1^{2+\delta} (x_t^*) \right] < \infty, \quad E \left[ m_0 ^2 (x_t^*) \right] < \infty, \quad E \left[ m_0 (x_t^*) \left( m_3 (x_t^*) + 1 \right) \right] < \infty.
$$
Proof. Observe that
\[ s_i(y_t|x_{t-1}; \theta) = \frac{D_i h_t}{h_t} \left( 1 - \frac{y_t^2}{h_t^2} \right) \] (14)
and
\[ D_j s_i(y_t|x_{t-1}; \theta) = \frac{D_i h_t D_j h_t}{h_t^2} \left( 2 \frac{y_t^2}{h_t} - 1 \right) + \frac{D^2_{ij} h_t}{h_t} \left( 1 - \frac{y_t^2}{h_t} \right), \] (15)
\[ = \left( D_{ij}^2 h_t - D_i h_t D_j h_t \right) \left( 1 - \frac{y_t^2}{h_t^2} \right) + \frac{D_i h_t D_j h_t y_t^2}{h_t^2} \] (16)
where \( D h_t = D H(x_{t-1}; \theta) \) and \( D^2 h_t = D^2 H(x_{t-1}; \theta) \).

Using (14) and the assumed bounds,
\[ |s_i(y_t|x_{t-1}; \theta)| \leq \frac{|D_i H(x_{t-1}; \theta)|}{H(x_{t-1}; \theta)} \left[ 1 + \frac{H(x_{t-1}; \theta_0) z_t^2}{H(x_{t-1}; \theta) z_t^2} \right] \leq m_1(x_{t-1}) \left[ 1 + m_0(x_{t-1}) z_t^2 \right], \]
while
\[ |s_i(y_t|x_{t-1}; \theta)| \leq \frac{|D_i H(x_{t-1}; \theta)|}{H(x_{t-1}; \theta)} \left[ 1 + \frac{H(x_{t-1}; \theta_0) z_t^2}{H(x_{t-1}; \theta_0) z_t^2} \right] \leq m_1(x_{t-1}) \left[ 1 + z_t^2 \right]. \]

Applying the assumed bounds on \( D_j s_i \),
\[ |D_j s_i(y_t|x_{t-1}; \theta)| \leq \left| \frac{D_i H(x_{t-1}; \theta)}{H(x_{t-1}; \theta)} \right| \left| \frac{D_j H(x_{t-1}; \theta)}{H(x_{t-1}; \theta)} \right| \left( 2 \frac{H(x_{t-1}; \theta_0)}{H(x_{t-1}; \theta)} z_t^2 + 1 \right) \]
\[ + \left| \frac{D^2_{ij} H(x_{t-1}; \theta)}{H(x_{t-1}; \theta)} \right| \left( 1 + \frac{H(x_{t-1}; \theta_0)}{H(x_{t-1}; \theta)} z_t^2 \right) \]
\[ \leq C \left\{ m_0(x_{t-1}) + m_0(x_{t-1}) m_3(x_{t-1}) + m_3(x_{t-1}) \right\} (z_t^2 + 1). \]
Since \( \|D s(y_t|x_{t-1}; \theta)\|^2 = \sum_{i,j=1}^d |D_j s_i(y_t|x_{t-1}; \theta)|^2 \), the asserted bound holds. \( \square \)

B. PROOFS: EXAMPLES

Proof of Corollary 2. We have
\[ H(x; \theta) = \left( \omega + \sum_{i=1}^q \alpha_i (|x_i - b_i| + \gamma_i (x_i - b_i))^a \right)^{2/\mu}, \]
where \( |x - b_i| + \gamma(x - b_i) = (\gamma + 1) (x - b_i) + (\gamma - 1) (x - b_i) 1 \{ x - b_i < 0 \} \geq 0 \) for \(-1 \leq \gamma \leq 1.\)

Consistency: Define
\[ C_0 = \{ \theta \in \Theta | \omega \leq \omega_0 + \delta, \mu \leq \mu_0, 0 \leq \alpha_i \leq \alpha_0 + \delta, -1 \leq \gamma_i \leq 1, i = 1, ..., q \}. \]
We have that

$$2 \log (\omega) / \mu \leq \log (H (x; \theta)) \leq 2 \log (B (\theta)) + V (x)/\mu$$

where $V (x) = 1 + \log (1 + \sum_{i=1}^{q} |x_i - b_i|^\mu)$ and $B (\theta) = \omega + \sum_{i=1}^{q} \alpha_i (1 + |\gamma_i|)^\mu$.

By Lemma 8 with $\beta = (\mu, \gamma)', \tau (\beta) = 2/\mu$, $\tau = 2/\mu$, and

$$g_j (x) = 1 (|x - b_i| > 1/\varepsilon) |x - b_i|^{1/2} \varepsilon^{1/2} + 1 (|x - b_i| \leq 1/\varepsilon) |x - b_i|^{1/2} \varepsilon^{1/2},$$

we have that (C.4) holds, since $\left( g_1 (x^*_1, \ldots, g_q (x^*_q) \right)$ is non-degenerate.

Identification???

**Asymptotic Normality:** Define $\beta = (\omega, \gamma, \alpha)', g_i (x_i; \gamma_i) = |x_i - b_i| + \gamma_i (x_i - b_i)$, and $A (x; \theta) = \omega + \sum_{i=1}^{q} \alpha_i g_i (x_i, \gamma)^\mu$. Then, $\log (H (x; \theta)) = \frac{2}{\mu} \log (A (x; \theta))$ such that

$$D_\theta H (x; \theta) = H (x; \theta) D_\theta \log (H (x; \theta)),$$

where

$$D_\mu \log (H (x; \theta)) = -\frac{2}{\mu^2} \log (A (x; \theta)) + \frac{2}{\mu} \frac{D_\mu A (x; \theta)}{A (x; \theta)},$$

$$D_\beta \log (H (x; \theta)) = \frac{2}{\mu} \frac{D_\beta A (x; \theta)}{A (x; \theta)},$$

and

$$D_\mu A (x; \theta) = \sum_{i=1}^{q} \alpha_i g_i (x_i, \gamma)^\mu \log (g (x_i; \gamma_i)),

D_\omega A (x; \theta) = 1, \quad D_{\alpha_i} A (x; \theta) = g_i (x_i; \gamma_i)^\mu,

D_{\gamma_i} A (x; \theta) = \alpha_i \mu g_i (x_i; \gamma_i)^{\mu-1} (x_i - b_i).$$

Similarly,

$$D_{\theta \theta}^2 H (x; \theta) = H (x; \theta) \left\{ D_\theta \log (H (x; \theta)) D_\theta \log (H (x; \theta))' + D_{\theta \theta}^2 \log (H (x; \theta)) \right\},$$

where

$$D_{\mu \mu}^2 \log (H (x; \theta)) = \frac{4}{\mu^3} \log (A (x; \theta)) - \frac{2}{\mu^2} \frac{D_\mu A (x; \theta)}{A (x; \theta)} - \frac{2}{\mu} \left( \frac{D_\mu A (x; \theta)}{A (x; \theta)} \right)^2 + \frac{2}{\mu} \frac{D_{\mu \mu} A (x; \theta)}{A (x; \theta)},$$

$$D_{\mu \beta}^2 \log (H (x; \theta)) = \frac{2}{\mu^2} \frac{D_\mu A (x; \theta)}{A (x; \theta)} - \frac{2}{\mu} \frac{D_\mu A (x; \theta) D_\beta A (x; \theta)}{A (x; \theta)} + \frac{2}{\mu} \frac{D_{\mu \beta} A (x; \theta)}{A (x; \theta)},$$

$$D_{\beta \beta}^2 \log (H (x; \theta)) = \frac{2}{\mu} \frac{D_\beta A (x; \theta) D_\beta A (x; \theta)'}{A (x; \theta)} + \frac{2}{\mu} \frac{D_{\beta \beta} A (x; \theta)}{A (x; \theta)},
and
\[ D_{\mu}^2 A (x; \theta) = \sum_{i=1}^{q} \alpha_i g_i (x_i; \gamma_i)^\mu \log (g (x_i; \gamma_i))^2, \]
\[ D_{\alpha_i}^2 A (x; \theta) = g_i (x_i; \gamma_i)^\mu \log (g (x_i; \gamma_i)), \]
\[ D_{\alpha_i}^2 A (x; \theta) = \alpha_i \mu g_i (x_i; \gamma_i)^{\mu - 1} (x_i - b_i) \{ \log (g (x_i; \gamma_i)) + 1 \} \]
\[ D_{\alpha_i \gamma_i}^2 A (x; \theta) = \mu g_i (x_i; \gamma_i)^{\mu - 1} (x_i - b_i), \]
and all other entries equal to zero. Thus,
\[ \left( \frac{H_0 (x)}{H (x)} \right)^{\mu_0 / 2} \leq \frac{\omega_0 + \sum_{i=1}^{q} \alpha_{0,i} \left( |x_i - b_i| + \gamma_i (x_i - b_i) \right)^{\mu_0}}{\left( \omega + \sum_{i=1}^{q} \alpha_i \left( |x_i - b_i| + \gamma_i (x_i - b_i) \right)^{\mu_0 / \mu} \right)^{\mu_0 / \mu}} \leq \sum_{i=1}^{q} \frac{\omega_0 + \alpha_{0,i} \left( |x_i - b_i| + \gamma_i (x_i - b_i) \right)^{\mu_0}}{\left( \omega + \alpha_i \left( |x_i - b_i| + \gamma_i (x_i - b_i) \right)^{\mu_0 / \mu} \right)^{\mu_0 / \mu}} \leq C, \]

since
\[ \frac{\omega_0 + \alpha_{0,i} \left( |x_i - b_i| + \gamma_i (x_i - b_i) \right)^{\mu_0}}{\left( \omega + \alpha_i \left( |x_i - b_i| + \gamma_i (x_i - b_i) \right)^{\mu_0 / \mu} \right)^{\mu_0 / \mu}} \leq \frac{\omega_0 + \alpha_{0,i} \left( 1 + |\gamma_{0,i}| \right)^{\mu_0}}{\min \left\{ \alpha_{0,i}^{\mu_0 / \mu}, \alpha_i^{\mu_0 / \mu} \right\}}, \quad |x_i - b_i| \leq 1, \]

while
\[ \frac{\omega_0 + \alpha_{0,i} \left( |x_i - b_i| + \gamma_i (x_i - b_i) \right)^{\mu_0}}{\left( \omega + \alpha_i \left( |x_i - b_i| + \gamma_i (x_i - b_i) \right)^{\mu_0 / \mu} \right)^{\mu_0 / \mu}} \leq \max \left\{ \alpha_{0,i}^{\mu_0 / \mu}, \alpha_i^{\mu_0 / \mu} \right\} \max \left\{ \frac{1 + \gamma_{0,i}}{1 + \gamma_i}, \frac{1 - \gamma_{0,i}}{1 - \gamma_i} \right\}, \]

for \(|x_i| > 1\).

Next, we show that
\[ \left\| \frac{DH (x; \theta)}{H (x; \theta)} \right\| \leq C \left( 1 + \sum_{i=1}^{q} \log^+ (|x_i|) \right), \]

for \(\theta \in N_0\), where
\[ N_0 = \left\{ \theta \in \Theta | \omega < \omega < \bar{\omega}, \mu < \bar{\mu}, \alpha_i < \bar{\alpha}_i, \gamma_i < \bar{\gamma}_i, i = 1, \ldots, q \right\}, \]

and the bounds are chosen such that \(\theta_0 \in N_0 \subset \Theta\). In particular, \(\alpha_i > 0\), \(\gamma_i > -1\) and \(\bar{\gamma}_i < 1\). We have \(\|D_\theta H (x; \theta)\| / H (x; \theta) = \|D_\theta \log (H (x; \theta))\|\), where
\[ |D_\mu \log (H (x; \theta))| \leq \frac{2}{\mu^2} \left| \log (A (x; \theta)) \right| + \frac{2}{\mu} \frac{D_\mu A (x; \theta)}{A (x; \theta)}, \]
\[ \|D_\beta \log (H (x; \theta))\| \leq \frac{2 \left\| \frac{D_\beta A (x; \theta)}{A (x; \theta)} \right\|}{\mu}, \]

22
and
\[
\frac{|D_\mu A(x; \theta)|}{A(x; \theta)} \leq C + \sum_{i=1}^{q} \log^+ (|x_i|),
\]
\[
\frac{|D_\alpha A(x; \theta)|}{A(x; \theta)} \leq \frac{1}{\alpha}, \quad \frac{|D_\alpha, A(x; \theta)|}{A(x; \theta)} \leq \frac{1}{\alpha},
\]
\[
\frac{|D_\gamma A(x; \theta)|}{A(x; \theta)} \leq \frac{\mu}{\min \{\gamma_i + 1, \gamma_i - 1\}}.
\]

Finally, we claim that
\[
\frac{\|D^2 H(x; \theta)\|}{H(x; \theta)} \leq C \left(1 + \sum_{i=1}^{q} \log^+ (|x_i|^2)\right),
\]
for \(\theta \in \mathcal{N}_0\). We have
\[
\frac{\|D^2_{\theta \theta} H(x; \theta)\|}{H(x; \theta)} \leq \|D_{\theta} \log (H(x; \theta))\|^2 + \|D^2_{\theta \theta} \log (H(x; \theta))\|,
\]
where
\[
|D^2_{\mu \mu} \log (H(x; \theta))| \leq \frac{4}{\mu^3} |\log (A(x; \theta))| + \frac{4}{\mu^2} \frac{|D_\mu A(x; \theta)|}{A(x; \theta)} + \frac{2}{\mu} \frac{(D_\mu A(x; \theta))^2}{A^2(x; \theta)} + \frac{2}{\mu} \frac{|D^2_{\mu \mu} A(x; \theta)|}{A(x; \theta)},
\]
\[
\frac{\|D^2_{\mu \beta} \log (H(x; \theta))\|}{A(x; \theta)} \leq \frac{2}{\mu^2} \frac{\|D_\beta A(x; \theta)\|}{A(x; \theta)} + \frac{2}{\mu} \frac{|D_\beta A(x; \theta)| \|D_\beta A(x; \theta)\|}{A^2(x; \theta)} + \frac{2}{\mu} \frac{\|D^2_{\mu \beta} A(x; \theta)\|}{A(x; \theta)},
\]
\[
\frac{\|D^2_{\beta \beta} \log (H(x; \theta))\|}{A(x; \theta)} = \frac{2}{\mu} \frac{\|D_\beta A(x; \theta)\|^2}{A^2(x; \theta)} + \frac{2}{\mu} \frac{\|D_\beta A(x; \theta)\|}{A(x; \theta)},
\]
and
\[
\frac{|D^2_{\mu \mu} A(x; \theta)|}{A(x; \theta)} \leq C + \sum_{i=1}^{q} \log^+ (|x_i - b_i|)^2,
\]
\[
\frac{|D^2_{\mu \alpha} A(x; \theta)|}{A(x; \theta)} \leq C + \log^+ (|x_i - b_i|),
\]
\[
\frac{D^2_{\mu \gamma} A(x; \theta)}{A(x; \theta)} \leq \frac{\bar{\mu} (\log^+ (|x_i - b_i|) + 1)}{\min \{\gamma_i + 1, \gamma_i - 1\}}
\]
\[
\frac{D^2_{\alpha \gamma} A(x; \theta)}{A(x; \theta)} \leq \frac{\bar{\mu}}{\alpha_i \min \{\gamma_i + 1, \gamma_i - 1\}}.
\]

We conclude that \(V(x) \geq 1 + \sum_{i=1}^{q} \log^+ (g_i(x; \gamma_i))^2\) is sufficient for asymptotic normality.

**V-geometricity:** The sufficient conditions for geometric ergodicity follows from Lemma X \(\Box\)

23
**Proof of Corollary 3.** We have

\[
H(x; \theta) = \omega + \sum_{i=1}^{q} \left\{ \alpha_{1i}1_{(x_i < 0)} |x_i|^\mu + \alpha_{2i}1_{(x_i \geq 0)} |x_i|^\mu \right\}.
\]

**Consistency:** Define

\[
C_0 = \{ \theta \in \Theta | \omega \leq \omega_0 + \delta, \mu \leq \mu \leq \mu_0, 0 \leq \alpha_{ij} \leq \alpha_{0,ij} + \delta, i = 1, 2, j = 1, \ldots, q \}.
\]

We have \( \log(\omega) \leq \log(H(x; \theta)) \leq \log(B(\theta)) + V(x) \) where \( V(x) = 1 + \log(1 + \sum_{i=1}^{q} |x_i|^\mu) \) and \( B(\theta) = \omega + \sum_{i=1}^{q} \{ \alpha_{1i} + \alpha_{2i} \} \).

Using the same arguments as in the proof of Lemma 8, we have that (C.4) holds.

**Identification?**

**Asymptotic Normality:**

\[
D_\omega H(x; \theta) = 1,
\]

\[
D_\mu H(x; \theta) = \sum_{i=1}^{q} \log(|x_i|) \left\{ \alpha_{1i}1_{(x_i < 0)} |x_i|^\mu + \alpha_{2i}1_{(x_i \geq 0)} |x_i|^\mu \right\},
\]

\[
D_{\alpha_{1i}} H(x; \theta) = 1_{(y_{r-i} < 0)} |x_i|^\mu,
\]

\[
D_{\alpha_{2i}} H(x; \theta) = 1_{(y_{r-i} \geq 0)} |x_i|^\mu.
\]

\[
D_{\omega,\theta} H(x; \theta) = 0,
\]

\[
D_{\mu,\mu} H(x; \theta) = \sum_{i=1}^{q} \log(|x_i|)^2 \left\{ \alpha_{1i}1_{(x_i < 0)} |x_i|^\mu + \alpha_{2i}1_{(x_i \geq 0)} |x_i|^\mu \right\},
\]

\[
D_{\mu,\alpha_{1i}} H(x; \theta) = \log(|x_i|) 1_{(x_i < 0)} |x_i|^\mu,
\]

\[
D_{\mu,\alpha_{2i}} H(x; \theta) = \log(|x_i|) 1_{(x_i \geq 0)} |x_i|^\mu,
\]

\[
D_{\alpha_{1i},\alpha_{1j}} H(x; \theta) = D_{\alpha_{1i},\alpha_{2j}} H(x; \theta) = 0.
\]

We have

\[
\frac{\omega_0 + \sum_{i=1}^{q} \{ \alpha_{0,1i}1_{(x_i < 0)} |x_i|^\mu_0 + \alpha_{0,2i}1_{(x_i \geq 0)} |x_i|^\mu_0 \}}{\omega + \sum_{i=1}^{q} \{ \alpha_{1i}1_{(x_i < 0)} |x_i|^\mu + \alpha_{2i}1_{(x_i \geq 0)} |x_i|^\mu \}} \leq \sum_{i=1}^{q} \frac{\omega_0 + \alpha_{0,1i}1_{(x_i < 0)} |x_i|^\mu_0}{\omega + \alpha_{1i}1_{(x_i < 0)} |x_i|^\mu} + \sum_{i=1}^{q} \frac{\omega_0 + \alpha_{0,2i}1_{(x_i \geq 0)} |x_i|^\mu_0}{\omega + \alpha_{2i}1_{(x_i \geq 0)} |x_i|^\mu},
\]

24
Thus, where
\[
\frac{\omega_0 + \alpha_0 \lambda_1 |x_i|^{\mu_0}}{\omega + \alpha_1 \lambda_1 |x_i|^{\mu}} \leq \frac{\omega_0 + \alpha_0 \lambda_1}{\omega}, \quad |x_i| \leq 1,
\]
and
\[
\frac{\omega_0 + \alpha_0 \lambda_1 |x_i|^{\mu_0}}{\omega + \alpha_1 \lambda_1 |x_i|^{\mu}} \leq \frac{\omega_0 + \alpha_0 \lambda_1 |x_i|^q}{\omega}, \quad |x_i| > 1,
\]
Furthermore, since \( \log (|x_i|) |x_i|^\mu \to 0 \) as \( |x_i| \to 0 \) for any \( \beta > 0 \),
\[
\frac{\log (|x_i|) \lambda_1(1) |x_i|^\mu}{\omega + \alpha_1 \lambda_1 |x_i|^{\mu}} \leq c + \frac{1}{\alpha_1} \log (|x_i|) 1 \{ |x_i| > 1 \},
\]
and
\[
\frac{\log (|x_i|) \lambda_1(2) |x_i|^\mu}{\omega + \alpha_2 \lambda_1 |x_i|^{\mu}} \leq c + \frac{1}{\alpha_2} \log (|x_i|) 1 \{ |x_i| > 1 \}.
\]
Thus,
\[
\frac{H_0(x; \theta)}{H(x; \theta)} \leq c \left( 1 + \sum_{i=1}^q |x_i|^\delta \right).
\]
\[
\|DH(x; \theta)\| \leq c \left( 1 + \sum_{i=1}^q \log (|x_i|) 1 \{ |x_i| > 1 \} \right),
\]
\[
\|D^2H(x; \theta)\| \leq c \left( 1 + \sum_{i=1}^q \log (|x_i|)^2 1 \{ |x_i| > 1 \} \right).
\]
So we wish to verify the drift criterion with \( V(x) = 1 + \sum_{i=1}^q |x_i|^\delta \) for arbitrarily small \( \delta > 0 \).

**V-geometricity:** It follows from Carrasco and Chen (2002, Corollary 10) that for \( q = 1 \) and \( \mu = 2 \alpha_1 + E [\alpha_2 \max \{-z_i, 0\}^2] < 1 \) is sufficient condition for \( V(x) = 1 + x^2 \). For \( \mu < 2 \), \( q > 1 \)?

**Proof of Corollary 4.** We have
\[
H(x; \theta) = \omega + \sum_{i=1}^q \alpha_i |x_i|^{\mu}.
\]
Consistency: Define

\[ C = \{ \theta \in \Theta \mid \omega \leq \omega \leq \omega_0 + \delta, \mu \leq \bar{\mu}, 0 \leq \alpha_i \leq \alpha_{0,i} + \delta, i = 1, \ldots, q \} . \]

It holds that \( \log (\omega) \leq \log (H (x; \theta)) \leq \log (B (\theta)) + V (x) \) where \( V (x) = 1 + \log \left( 1 + \sum_{i=1}^{q} |x_i|^\bar{\mu} \right) \).

By Lemma 8, we have that (C.4) holds.

Identification?

Asymptotic Normality:

\[
\begin{align*}
D_\omega H (x; \theta) &= 1, \\
D_\mu H (x; \theta) &= \sum_{i=1}^{q} \alpha_i \log (|x_i|) |x_i|^\mu, \\
D_\alpha_i H (x; \theta) &= |x_i|^\mu, \\
D_{\omega, \theta} H (x; \theta) &= 0, \\
D_{\mu, \mu} H (x; \theta) &= \sum_{i=1}^{q} \alpha_i \log (|x_i|)^2 |x_i|^{2\mu}, \\
D_{\mu, \alpha_i} H (x; \theta) &= \log (|x_i|)^2 |x_i|^\mu, \\
D_{\alpha_i, \alpha_j} H (x; \theta) &= 0.
\end{align*}
\]

Using the same arguments as before,

\[
\begin{align*}
\frac{H_0 (x; \theta)}{H (x; \theta)} &\leq c \left( 1 + \sum_{i=1}^{q} |x_i|^{\delta} \right), \\
\frac{\|DH (x; \theta)\|}{H (x; \theta)} &\leq c \left( 1 + \sum_{i=1}^{q} \log^+ (|x_i|) \right), \\
\frac{\|D^2 H (x; \theta)\|}{H (x; \theta)} &\leq c \left( 1 + \sum_{i=1}^{q} \log^+ (|x_i|)^2 \right).
\end{align*}
\]

Again, we have to check that \( V (x) := 1 + \sum_{i=1}^{q} |x_i|^{\delta} \) for arbitrarily small \( \delta > 0 \) satisfies a drift criterion.

V-geometricity: Choosing \( V (x) := 1 + \sum_{i=1}^{q} |x_i|^2 \), for \( \mu_0 < 2 \), we have geometric ergodicity without any restrictions on \( \alpha \), while for \( \mu_0 = 2 \), \( \sum_{i=1}^{q} \alpha_{0,i} < 1 \) is a necessary and sufficient condition, c.f. Lu (1998, Theorem 4.1).
Proof of Corollary 5. We have

\[ H(x; \theta) = \omega + \sum_{i=1}^{q} \alpha_i \left[ 2F \left( \frac{x_i^2}{\mu} \right) - 1 \right]. \]

Consistency: Define

\[ C_0 = \{ \theta \in \Theta | \omega \leq \omega \leq \omega_0 + \delta, \mu \leq \mu \leq \bar{\mu}, 0 \leq \alpha_i \leq \alpha_{0,i} + \delta, i = 1, \ldots, q \}. \]

We have \( \log(\omega) \leq \log \left( H(x; \theta) \right) \leq \log \left( B(\theta) \right) + V(x) \) where \( V(x) = 1 \), and where \( \underline{\mu} \leq \mu < \infty \). By Lemma 8, we have that (C.4) holds.

Identification?

Asymptotic Normality:

\[ D_\omega H(x; \theta) = 1, \]
\[ D_\mu H(x; \theta) = -\frac{2}{\mu^2} \sum_{i=1}^{q} \alpha_i F'(y_{i-i}^2/\mu) y_{i-i}^2, \]
\[ D_{\alpha_i} H(x; \theta) = 2F \left( \frac{y_{i-i}^2}{\mu} \right) - 1, \]
\[ D_{\omega,\theta} H(x; \theta) = 0, \]
\[ D_{\mu,\mu} H(x; \theta) = \frac{4}{\mu^3} \sum_{i=1}^{q} \alpha_i F'(y_{i-i}^2/\mu) y_{i-i}^2 - \frac{2}{\mu^4} \sum_{i=1}^{q} \alpha_i F''(y_{i-i}^2/\mu) y_{i-i}^2, \]
\[ D_{\mu,\alpha_i} H(x; \theta) = -\frac{2}{\mu^2} F'(y_{i-i}^2/\mu) y_{i-i}^2 \]
\[ D_{\alpha_i,\alpha_j} H(x; \theta) = 0. \]

Thus,

\[ \frac{H(x; \theta_0)}{H(x; \theta)} = \frac{\omega_0 + \sum_{i=1}^{q} \alpha_{0,i} \left[ 2F \left( \frac{x_i^2}{\mu_0} \right) - 1 \right]}{\omega + \sum_{i=1}^{q} \alpha_i \left[ 2F \left( \frac{x_i^2}{\mu} \right) - 1 \right]} \leq \frac{\omega_0}{\omega} + \sum_{i=1}^{q} \frac{\alpha_{0,i}}{\omega}, \]

\[ \frac{\|DH(x; \theta)\|}{H(x; \theta)} \leq c \left( 1 + \sum_{i=1}^{q} x_i^4 \right), \]

\[ \frac{\|D^2H(x; \theta)\|}{H(x; \theta)} \leq c \left( 1 + \sum_{i=1}^{q} x_i^8 \right). \]

So we have to check that \( V(x) := 1 + \sum_{i=1}^{q} x_i^8 \) satisfies a drift criterion.
**V-geometricity:** Since $F$ is bounded, we can prove geometric ergodicity with $V (x) := 1 + \sum_{i=1}^{q} x_i^{2q}$, for any $q > 0$ without restricting the parameters, c.f. Lu (1998, Remark 4.3).

**Proof of Corollary 6.** We have

$$H (x; \theta) = \left( \omega + \sum_{i=1}^{q} \alpha_i x_i^{2\mu} \right)^{1/\mu}.$$  

**Consistency:** Define

$$\mathcal{C} = \{ \theta \in \Theta | \omega \leq \omega, \mu \leq \mu, 0 \leq \alpha_i \leq \alpha_i, i = 1, \ldots, q \}.$$  

Then $\frac{1}{\mu} \log (\omega) \leq \log (H (x; \theta)) \leq \log (B (\theta)) + V (x)$ where $V (x) = 1 + \log \left( 1 + \sum_{i=1}^{q} y_i^{2\mu} \right)$ when $\frac{1}{\mu} \leq 1$.

**Identification?**

**Asymptotic Normality:** Define $\beta = (\omega, \alpha_1, \ldots, \alpha_q)'$. Then, $DH (x; \theta) = (D_{\beta}H (x; \theta)', D_{\mu}H (x; \theta))'$ where $D_{\beta}H (x; \theta) = \frac{1}{\mu} H (x; \theta)^{1-\mu} A_{\beta} (x; \theta)$, and

$$D_{\mu}H (x; \theta) = \frac{1}{\mu} H (x; \theta)^{1-\mu} A_{\mu} (x; \theta) + H (x; \theta) B_{\mu} (x; \theta),$$

where $A_{\beta} (x; \theta) = (A_{\omega} (x; \theta), A_{\alpha_1} (x; \theta), \ldots, A_{\alpha_q} (x; \theta))'$ with

$$A_{\omega} (x; \theta) = 1, \quad A_{\alpha_i} (x; \theta) = x_i^{2\mu},$$

while

$$A_{\mu} (x; \theta) = \sum_{i=1}^{q} \alpha_i \log \left( x_i^{2\mu} \right) x_i^{2\mu},$$

$$B_{\mu} (x; \theta) = -\frac{1}{\mu^2} \log \left( \omega + \sum_{i=1}^{q} \alpha_i x_i^{2\mu} \right).$$

Similarly,

$$D_{\beta\beta}^{2}H (x; \theta) = \frac{1-\mu}{\mu^2} H (x; \theta)^{1-2\mu} A_{\beta} (x; \theta) A_{\beta} (x; \theta)'',$$

$$D_{\mu\beta}^{2}H (x; \theta) = \frac{1-\mu}{\mu^2} H (x; \theta)^{1-2\mu} A_{\mu} (x; \theta) A_{\beta} (x; \theta)' + \frac{2}{\mu} H (x; \theta)^{1-\mu} D_{\mu}A_{\beta} (x; \theta),$$

$$D_{\mu\mu}^{2}H (x; \theta) = -\frac{2}{\mu^2} H (x; \theta)^{1-\mu} \{ A_{\mu} (x; \theta) + D_{\mu}A_{\mu} (x; \theta) + A_{\mu} (x; \theta) B_{\mu} (x; \theta) \}$$

$$+ H (x; \theta) \{ B_{\mu}^{2} (x; \theta) + D_{\mu}B_{\mu} (x; \theta) \}.$$
where $D_\mu A_\alpha (x; \theta) = \log (x_i^2) x_i^2 \mu$ and $D_\mu A_\omega (x; \theta) = 0$, while

$$D_\mu A_\mu (x; \theta) = 2 \sum_{i=1}^q \alpha_i \log (x_i^2)^2 x_i^2 \mu,$$

$$D_\mu B_\mu (x; \theta) = \frac{4}{\mu^3} \log \left( \omega + \sum_{i=1}^q \alpha_i x_i^2 \mu \right) - \frac{2}{\mu^2} \sum_{i=1}^q \alpha_i \log (x_i^2) x_i^2 \mu.$$

Thus,

$$\left( \frac{H_0 (x)}{H (x)} \right)^{\mu_0} \leq \frac{\omega_0 + \sum_{i=1}^q \alpha_0,i x_i^2 \mu_0}{\left( \omega + \sum_{i=1}^q \alpha_i x_i^2 \mu \right)^{\mu_0/\mu}} \leq \sum_{i=1}^q \frac{\omega_0 + \alpha_0,i x_i^2 \mu_0}{\left( \omega + \alpha_i x_i^2 \right)^{\mu_0/\mu}} \leq \text{const.},$$

since

$$\frac{\omega_0 + \alpha_0,i x_i^2 \mu_0}{\left( \omega + \alpha_i x_i^2 \right)^{\mu_0/\mu}} \leq \frac{\omega_0 + \alpha_0,i}{\omega} \frac{1}{|x_i|} \leq 1,$$

$$\frac{\omega_0 + \alpha_0,i x_i^2 \mu_0}{\left( \omega + \alpha_i x_i^2 \right)^{\mu_0/\mu}} \leq \frac{\omega_0}{\max \{ \omega^{\mu_0/\mu}, \bar{\omega}^{\mu_0/\mu} \}} + \frac{\alpha_0,i}{\max \{ \alpha_i^{\mu_0/\mu}, \bar{\alpha}_i^{\mu_0/\mu} \}}, \quad |x_i| > 1.$$

Also,

$$\frac{\|D_\beta H (x; \theta)\|}{H (x; \theta)} = \frac{1}{\mu} \frac{H (x; \theta)^{\mu} \|A_\beta (x; \theta)\|}{H (x; \theta)},$$

$$\frac{\|D_\mu H (x; \theta)\|}{H (x; \theta)} = \frac{2}{\mu} \frac{H (x; \theta)^{-\mu} \|A_\mu (x; \theta)\|}{H (x; \theta)} + \|B_\mu (x; \theta)\|,$$

where

$$\frac{|A_\omega (x; \theta)|}{H (x; \theta)^{\mu}} = \frac{1}{\omega + \sum_{i=1}^q \alpha_i x_i^2 \mu} \leq \frac{1}{\omega},$$

$$\frac{|A_\alpha (x; \theta)|}{H (x; \theta)^{\mu}} = \frac{x_i^2 \mu}{\omega + \sum_{j=1}^q \alpha_j x_j^2 \mu} \leq \frac{1}{\alpha_i},$$

$$\frac{\|A_\mu (x; \theta)\|}{H (x; \theta)^{\mu/2}} \leq \frac{2 \sum_{j=1}^q \alpha_i |\log (x_i^2)| x_i^2 \mu}{\omega + \sum_{j=1}^q \alpha_j x_j^2 \mu} \leq 2 \sum_{j=1}^q \log^+ (x_j^2).$$

Similarly,

$$\frac{D_{\beta_\alpha}^2 H (x; \theta)}{H (x; \theta)} = 2 - \frac{\mu}{\mu} H (x; \theta)^{-2\mu} A_\beta (x; \theta) A_\beta (x; \theta)',$$

$$\frac{D_{\mu_\alpha}^2 H (x; \theta)}{H (x; \theta)} = 2 - \frac{\mu}{\mu} H (x; \theta)^{-2\mu} A_\mu (x; \theta) A_\beta (x; \theta)' + \frac{2}{\mu} H (x; \theta)^{-\mu} D_\mu A_\beta (x; \theta),$$

$$\frac{D_{\mu_\mu}^2 H (x; \theta)}{H (x; \theta)} = -\frac{2}{\mu^2} H (x; \theta)^{-\mu} \{ A_\mu (x; \theta) + D_\mu A_\mu (x; \theta) + A_\mu (x; \theta) B_\mu (x; \theta) \}$$

$$+ B_\mu^2 (x; \theta) + D_\mu B_\mu (x; \theta).$$
We conclude that
\[
\frac{\|DH(x; \theta)\|}{H(x; \theta)} \leq C \left( 1 + \sum_{i=1}^{q} |\log^+ (x_i^2)| \right),
\]
\[
\frac{\|D^2H(x; \theta)\|}{H(x; \theta)} \leq C \left( 1 + \sum_{i=1}^{q} |\log^+ (x_i^2)|^2 \right),
\]
in a neighbourhood of \( \theta_0 \). So we require that \( V(x) \geq 1 + \sum_{i=1}^{q} |\log^+ (x_i^2)|^2 \).

**V-geometricity:** Choosing \( V(x) := 1 + \sum_{i=1}^{q} x_i^{2\alpha_i} \), \( E \left[ z_i^{2\alpha_i} \right] \sum_{i=1}^{q} \alpha_{0,i} < 1 \) is a necessary and sufficient condition for V-geometric ergodicity, c.f. Lu (1998, Theorem 4.2).

**Proof of Corollary 10.** Here, \( H(x; \theta) = \exp [\omega \prod_{i=1}^{q} x_i^{2\alpha_i}] \).

**Consistency:** We have with \( V(x) = 1 + \sum_{i=1}^{q} |\log (x_i^2)| \) that \( |\log H(x; \theta)| \leq B(\theta) V(x) \), for \( B(\theta) = |\omega| + |\alpha_1| + \ldots + |\alpha_q| \).

Identification: With \( \delta = \theta - \theta_0 \),
\[
\log \left( \frac{H(x; \theta)}{H(x; \theta_0)} \right) = \log H(x; \theta) - \log H(x; \theta_0) = \delta_0 + \sum_{i=1}^{q} \delta_i \log (x_i^2) \neq 0 \text{ a.s.}
\]
so we require that \( P \left( \sum_{i=1}^{q} \delta_i \log (y_i^2) = c \right) = 0 \) for all \( \delta \in \mathbb{R}^q \) and \( c \in \mathbb{R} \).

**Asymptotic Normality:** We have
\[
\frac{H(x; \theta_0)}{H(x; \theta)} = \frac{\tilde{\omega}_0 \prod_{i=1}^{q} x_i^{\tilde{\alpha}_i}}{\tilde{\omega} \prod_{i=1}^{q} x_i^{\bar{\alpha}_i}} = \frac{\tilde{\omega}_0}{\tilde{\omega}} \prod_{i=1}^{q} x_i^{\tilde{\alpha}_{0,i} - \bar{\alpha}_i},
\]
while the first order partial derivatives satisfy
\[
\frac{DH(x; \theta)}{H(x; \theta)} = D \log H(x; \theta) = (1, \log (x_1^2), \ldots, \log (x_q^2))',
\]
and the second order ones
\[
D^2 \log H(x; \theta) = 0.
\]

Thus,
\[
s_i (y_t|x_{t-1}; \theta) = D \log h_t \left( 1 - \frac{y_t^2}{h_t} \right) \quad (17)
\]
and
\[
Ds (y_t|x_{t-1}; \theta) = D \log h_t (D \log h_t)' \frac{y_t^2}{h_t}, \quad (18)
\]
such that
\[
|s_i (y_t|x_{t-1}; \theta)| \leq c \left( 1 + \sum_{i=1}^{q} |\log (y_i^2)_{t-1}| \right) \left( 1 + z_t^2 \prod_{i=1}^{q} \max \{ y_{t-i}, y_{t-i}^\delta \} \right),
\]

30
Lemma 8. Assume that:

1. \( H (x; \theta) = (\omega + \sum_{i=1}^{q} \alpha_i g (x_i; \beta))^{\tau (\beta)} \), where \( g (x_i; \beta) \geq g (x_i) \geq 0 \) and \( \tau (\beta) \geq \tau \).

2. \( \theta = (\omega, \alpha, \beta) \in \Theta \), where \( \beta \in \mathcal{B} \),

\[ \Theta \equiv \{ (\omega, \alpha) \mid |\omega| \leq \omega, 0 \leq \alpha_i, i = 1, \ldots, q \} \times \mathcal{B} \]

and \( \mathcal{B} \) compact.

3. \( P (g (x_{1,t-1}^m) > \epsilon, \ldots, g (x_{q,t-1}^m) > \epsilon) > 0 \) for some \( \epsilon > 0 \).

Then there exists \( \delta > 0 \) such that with

\[ \mathcal{C}_0 \equiv \{ (\omega, \alpha) \mid |\omega| \leq \omega_0 + \delta, 0 \leq \alpha_i \leq \alpha_{0,i} + \delta, i = 1, \ldots, q \} \times \mathcal{B}, \quad (19) \]

\[ \inf_{\theta \in \Theta \setminus \mathcal{C}_0} l (y|x; \theta) \geq l (y|x) \text{ with } E \left[ l (y^*_t | x^*_t - 1) \right] > L (\theta_0). \]

Proof. Let \( \theta \in \Theta \setminus \mathcal{C}_0 \) be given; then either (i) \( \omega_0 + \delta_0 < \omega \) or (ii) \( \alpha_{0,i} + \delta_i < \alpha_i \), for at least one \( i \in \{ 1, \ldots, m \} \). In the first case,

\[ l (y|x; \theta) \geq \log (H (x; \theta)) \geq \tau (\beta) \log (\omega) > \tau \log (\omega + \delta) \geq \tau \log (\delta) \geq \tau \log (\delta) \geq \tau \log (\delta) 1 \{ x_1 > 1, \ldots, x_q > 1 \} \]
while in the second case, choosing $\omega < 1$ and $\delta > 1$ and $\epsilon > 0$,
\[
    l(y|x; \theta) \geq \log (H(x; \theta)) \\
    \geq \tau \log (\omega + \alpha g(x_i; \beta)) \\
    \geq \tau \log (\omega + \alpha_0 g(x_i; \beta) + \delta g(x_i; \beta)) \\
    \geq \tau \log (\omega + \delta g(x_i)) \\
    = \tau \log (\omega + \delta g(x_i)) 1 \{ g(x_i) \leq \epsilon \} + \tau \log (\omega + \delta g(x_i)) 1 \{ g(x_i) > \epsilon \} \\
    \geq \tau \log (\omega) 1 \{ g(x_i) \leq \epsilon \} + \mu \log (\delta \epsilon) 1 \{ g(x_i) > \epsilon \} \\
    \geq \tau \log (\omega) + \tau \log (\delta \epsilon) 1 \{ g(x_i) > \epsilon \} \\
    \geq \tau \log (\omega) + \tau \log (\delta \epsilon) 1 \{ g(x_1) > \epsilon, \ldots, g(x_q) > \epsilon \}.
\]

Defining
\[
    l(y|x) := \tau \log (\omega) + \tau \log (\delta \epsilon) 1 \{ g(x_1) > \epsilon, \ldots, g(x_q) > \epsilon \},
\]
we then have $l(y_t|x_{t-1}; \theta) \geq l(y_t|x_{t-1})$ and
\[
    E l(y^*_t|x^*_{t-1}) = \tau \log (\omega) + \tau \log (\delta \epsilon) P (g(x^*_1|t-1) > \epsilon, \ldots, g(x^*_q|t-1) > \epsilon).
\]

Choosing
\[
    \delta = \frac{1}{\epsilon} \exp \left( \left( L(\theta_0) + \epsilon - \tau \log (\omega) \right) / \left( \tau P (g(x^*_1|t-1) > \epsilon, \ldots, g(x^*_q|t-1) > \epsilon) \right) \right) \in (1, \infty)
\]
for some $\epsilon > 0$, we obtain $E \left[ l(y^*_t|x^*_{t-1}) \right] = L(\theta_0) + \epsilon$. \qed

**Lemma 9.** For the APARCH-model, under (10)-(11), $\{x_t\}$ is geometrically ergodic.

**Proof.** For $q = 1$, we have
\[
    h^\mu_t = \omega + \alpha \left\{ (1 + \gamma) \mu |z_{t-1} \geq 0 \right\} + (1 - \gamma) \mu |z_{t-1} < 0 \right\} h^\mu_{t-1}.
\]
The result now follows from Chen and Carrasco (2004, Proposition 5). For $q > 1$, we apply Lu (1998, Theorem 3.1) with $s(x) = |x|^\mu/2$: Defining $\bar{\gamma}_i = \max \{ 1 + \gamma_i, 1 - \gamma_i \}$, we have for $\sum_i |y_{t-i}| \to \infty$, that
\[
    s(h_t) = \omega + \sum_{i=1}^q \alpha_i (|y_{t-i} - b_i| + \gamma_i (y_{t-i} - b_i))^\mu \\
    \leq \omega + \sum_{i=1}^q \alpha_i \bar{\gamma}_i^\mu |y_{t-i} - b_i|^\mu \\
    \leq C + \sum_{i=1}^q \alpha_i \bar{\gamma}_i^\mu s(y^2_t).
\]
\]
C. Auxiliary Results

Here, we state some results for geometrically ergodic Markov chains that are used in the paper. We first define the concept of $V$-geometric ergodicity. Let $\{Z_t\}$ denote a Markov chain on $\mathbb{R}^q$ characterised by its transition probabilities $P^t_Z(\cdot, \cdot)$, $P_Z(Z_t \in A | Z_0 = x) = P^t_Z(x, A)$.

**Proposition 1.** Assume that $\{Z_t\}$ is $V$-geometrically ergodic and that the function $f : \mathbb{R}^q \mapsto \mathbb{R}$ satisfies either:

1. $f^2 \leq V$, or
2. $\pi_Z(f^{2+\delta}) < \infty$ for some $\delta > 0$,

If $\pi (f) > 0$, then $\bar{f} (z) = f(z) - \pi_Z(f)$ satisfies

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \bar{f} (Z_t) \to^d N (0, \sigma^2),$$

where $\sigma^2 = E_\pi [\bar{f}^2 (Z_0)] + 2 \sum_{t=1}^{\infty} E_\pi [\bar{f} (Z_0) \bar{f} (Z_t)] < \infty$.

**Proof.** The result follows from Meyn and Tweedie (1993, Theorem 17.0.1) and Håggström (2005, Theorem 1.2) respectively. 

\[\square\]


