IWQW

Institut für Wirtschaftspolitik und Quantitative
Wirtschaftsforschung

Diskussionspapier
Discussion Papers

No. 03/2013

Quasi-maximum likelihood estimation in generalized polynomial autoregressive conditional heteroscedasticity models

Fabian Tinkl
University of Erlangen-Nuremberg

ISSN 1867-6707

Friedrich-Alexander-Universität
IWQW
Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

Submitted to the Bernoulli
arXiv: math.PR/0000000

Quasi-maximum likelihood estimation in generalized polynomial autoregressive conditional heteroscedasticity models

FABIAN TINKL[1,*],

[1] *Department of Statistics and Econometrics, University of Erlangen-Nuremberg, Lange Gasse 20, D-90403 Nuremberg. E-mail:* [*]Fabian.Tinkl@fau.de

Abstract

In this article, consistency and asymptotic normality of the quasi-maximum likelihood estimator (QMLE) in the class of polynomial augmented generalized autoregressive conditional heteroscedasticity models (GARCH) is proven. The result extends the results of the standard GARCH model to the class of polynomial augmented GARCH models which contains many commonly employed GARCH models as special cases. The results are obtained under mild conditions.

Keywords: Asymptotic normality; consistency; polynomial augmented GARCH models; quasi-maximum likelihood estimation.

1. Introduction

Since the introduction of the autoregressive conditional heteroscedastic model (ARCH) and its successful application to the variance of the UK's inflation rate by Engle [11], there has been a growing interest in these models. Especially, the extension to the linear GARCH (LGARCH) model of Bollerslev [5] has made it possible to capture many characteristics of financial data with one single model. One of these characteristics is that the conditional standard deviation of stock returns, usually referred to as volatility, seems to vary over time. Second, the variation of volatility shows some clustering behavior, meaning the existence of periods in which volatility is high and periods in which it is low, cf. McNeil et al. [31]. Bollerslev et al. [6] considers not only stock returns but also interest rates or exchange rates as possible fields of application.

This article focuses on parameter estimation of general GARCH models. Typically, the Gaussian maximum-likelihood estimator, also called quasi-maximum likelihood estimator (QMLE) is employed to adapt a GARCH process to data. Starting with Weiss [44], consistency and asymptotic normality have been derived for the QMLE under different assumptions by various authors, cf. Lee and Hansen [29], Lumsdaine [30], Jeantheau [25], Berkes et al. [3] or Francq and Zakoïan [15] among others. Nevertheless, these results are only applicable to the QMLE in the LGARCH model. The first reference to derive general asymptotic results of the QMLE in a broader class of GARCH models,

imsart-bj ver. 2013/03/06 file: bernoulli_main.tex date: April 18, 2013

including the LGARCH, a general class of asymmetric GARCH (AGARCH) models as well as the exponential GARCH (EGARCH) model of Nelson [34] is given by Straumann and Mikosch [42]. However, they only show consistency and asymptotic normality of the QMLE for the LGARCH and AGARCH setting. The consistency of the QMLE for the EGARCH model is shown under strict assumptions and the necessary moment conditions on the process are hard to verify. The asymptotic normality of the QMLE remains an open question. Recently, Pan et al. [35] and Hamadeh and Zakoïan [17] have shown consistency and asymptotic normality of the QMLE within the framework of power transformed threshold GARCH models that generalize the threshold GARCH model (TGARCH) of Zakoïan [46]. It can be concluded from Pan et al. [35, Theorem 1] that the assumptions needed to ensure consistency and asymptotic normality are quite similar for the TGARCH and the LGARCH model. The aim of this article is to generalize these results to augmented GARCH models that contain all the models mentioned above as special cases. The conditions are easily verified as soon as a specific GARCH model is considered.

As an illustration of this approach, consistency and asymptotic normality of the QMLE in the LGARCH and TGARCH setting is shown. It can be seen that the assumptions coincide with the ones in Francq and Zakoïan [15] and Hamadeh and Zakoïan [17]. Moreover, consistency and asymptotic normality of the QMLE for all the members of the polynomial GARCH and of the power transformed GARCH family discussed in Hentschel [20] can be derived. Additionally, an estimator for the variance-covariance matrix is proposed.

This article is structured as follows. In Section 2 the probabilistic structure of the augmented GARCH(1,1) model is briefly discussed and some results that are needed in the subsequent discussions are provided. Consistency and asymptotic normality of the QMLE in augmented GARCH models is shown in Section 3. To illustrate the results, Section 4 shows asymptotic properties of the QMLE in the LGARCH and TGARCH setting. Finally, Section 5 contains proofs.

2. Probabilistic structure of augmented GARCH models

Throughout this article, the gradient of a function $f : \boldsymbol{\theta} \to \mathbb{R}$ is written as $\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta})$ which is a column-vector of dimension d when $\boldsymbol{\theta} \in \mathbb{R}^d$ and the Hessian matrix as

$$\nabla^2_{\boldsymbol{\theta}} f(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\partial^2 f(\boldsymbol{\theta})}{\partial \theta_1^2} & \cdots & \frac{\partial^2 f(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\boldsymbol{\theta})}{\partial \theta_d \partial \theta_1} & \cdots & \frac{\partial^2 f(\boldsymbol{\theta})}{\partial \theta_d^2} \end{pmatrix}$$

which is of dimension $d \times d$. The notation $\nabla^3_{\boldsymbol{\theta}} f(\boldsymbol{\theta})$ denotes the third partial derivative of f w.r.t $\boldsymbol{\theta}$. The third partial derivative can be calculated as in Rao and Rao [37, p. 225

ff.]. Define $H := \nabla^2_{\boldsymbol{\theta}} f(\boldsymbol{\theta})$ then, the derivative

$$\frac{\partial H}{\partial \boldsymbol{\theta}} = \nabla_{\boldsymbol{\theta}} H$$

is understood as the following matrix derivative:

$$\nabla_{\boldsymbol{\theta}} H = \frac{\partial \mathrm{vec}(H)}{\partial (\mathrm{vec}(\boldsymbol{\theta}))'},$$

where $\mathrm{vec}(H)$ is the vec-operator formed by writing the columns of the matrix H one below the other. Therefore, $\mathrm{vec}(H)$ is of dimension d^2 and $\nabla_{\boldsymbol{\theta}} H$ is a $d^2 \times d$ matrix. The notation $\|A\| < \infty$ means that $\max_{i,j} |a_{i,j}| < \infty$. The expected value of a matrix is taken componentwise and $\boldsymbol{A}_n \overset{a.s.}{\to} 0$ indicates that every component of the matrix $\boldsymbol{A}_n$ converges a.s. to 0 as $n \to \infty$. A sequence of random elements $(f_n)_{n \in \mathbb{Z}}$ with values in some normed vector $(B, \|\|)$ converges a.s. towards 0 if $\lim_{n\to\infty} P(\sup_{m \geq n} \|f_m\| \geq \epsilon) = 0$ for all $\epsilon > 0$. The abbreviation e.a.s. is employed for indicating that

$$\rho^n \|f_n\| \overset{a.s.}{\to} 0$$

for $n \to \infty$ and some $\rho > 1$, similar to Straumann and Mikosch [42]. Thus, if $\|f_n\| \leq Z\rho^{-n}$ a.s. for some positive random variable Z, it follows that $\|f_n\| \overset{e.a.s}{\to} 0$. Finally, let $\log^+ x := \log(\max(x, 1))$. In this section, the probabilistic structure of the augmented GARCH model introduced by Duan [10] is discussed. Based on the work of e.g. Duan [10], Carrasco and Chen [8], Aue et al. [1] and Hörmann [23] many useful results on stationarity, ergodicity, and the existence of moments are available. First, some results of Hörmann [23] on strict stationarity and the existence of moments of these models are reviewed. To keep notation simple, only the case $p = q = 1$ is considered. The generalization to arbitrary order models is straightforward with minor changes in the proofs, because all results on stationarity and ergodicity carry over to GARCH models of arbitrary order, cf. Lee and Shin [28]. Apart from that, Hansen and Lunde [18] demonstrated that a simple LGARCH(1,1) specification fits financial data reasonable well. The following definition of an augmented GARCH(1,1) model is similar to the definition given by Francq and Zakoïan [16] and Hörmann [23].

Definition 2.1. *Let $g(x)$ and $c(x)$ be real-valued and measurable functions and assume that ϵ_t is an i.i.d. sequence. Assume that the stochastic recurrence equation*

$$h_t := h(\sigma_t^2) = c(\epsilon_{t-1})h_{t-1} + g(\epsilon_{t-1}), \tag{2.1}$$

has a strictly stationary solution and assume $h_t : \mathbb{R}^+ \to \mathbb{R}^+$ is an invertible function. Then, an augmented GARCH(1,1) process $(X_t)_{t \in \mathbb{Z}}$ is defined by the equations

$$\begin{aligned} X_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= h_t^{-1}. \end{aligned} \tag{2.2}$$

For the functions $c(x)$, $g(x)$ and $h(x)$ some assumptions have to be made in order to get a solution to the stochastic equations in (2.1). A solution is called non-anticipative if h_t is independent from $\sigma(\epsilon_s : s \geq t)$ and irreducible if $h_t \neq x \in \mathbb{R}$, such that $g(\epsilon_0) = x(1 - c(\epsilon_0))$. In the latter case $g(\epsilon_0)$ and $c(\epsilon_0)$ are linear combinations of each other and h_t reduces to a constant value x. For instance, let $g(\epsilon_0) \equiv 1$, $c(\epsilon_0) \equiv \beta$ with $0 < \beta < 1$ then, it holds that

$$h_t = \sum_{i=1}^{\infty} \beta^{i-1} = 1/(1-\beta)$$

which is constant. Hence, $x = 1/(1-\beta) \in \mathbb{R}^+$ and instead of an augmented GARCH process a simple white noise process with $E[X_0] = 0$ and $E[X_0^2] = (1-\beta)^{-1}$ is considered. Using the notation of non-anticipativity and irreducibility, the following theorem can be derived from Hörmann [23, Theorem 1] and Hörmann [23, Theorem 2]

Theorem 2.1. *Assume that ϵ_t is i.i.d. and that*

$$E[\log |c(\epsilon_0)|] < 0, \tag{2.3}$$

$E[\log^+ |c(\epsilon_0)|]$ and $E[\log^+ |g(\epsilon_0)|]$ are finite. Then, for every $t \in \mathbb{Z}$ the series

$$h_t = \sum_{i=1}^{\infty} g(\epsilon_{t-i}) \prod_{1 \leq j < i} c(\epsilon_{t-j}), \tag{2.4}$$

is a.s. convergent and h_t is the unique and strictly stationary solution of the equation (2.1). Conversely, if ϵ_t is i.i.d. and h_t is irreducible, the equation (2.2) has a strictly stationary non-anticipative solution. The series in (2.4) converges a.s. and is the unique stationary solution of equation (2.2).

Let $h(\sigma_t^2) = \sigma_t^2$, $c(\epsilon_t) = \beta + \alpha\epsilon_t^2$ and $g(\epsilon_t) = \omega$, then, the LGARCH(1,1) model of Bollerslev [5] is obtained. To ensure that $h_t : \mathbb{R}^+ \to \mathbb{R}^+$ is bijective and increasing it is required to have $\omega > 0, \alpha, \beta \geq 0$. Since $c(x)$ and $g(x)$ are continuous functions of x, they are measurable. The conditions for strict stationarity becomes

$$-\infty < E[\log(\beta + \alpha\epsilon_0^2)] < 0 \tag{2.5}$$

. Hence, the unique and strictly stationary solution h_t is of the form

$$\sigma_t^2 = \omega \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} (\beta + \alpha\epsilon_{t-j}^2). \tag{2.6}$$

It is shown in Nelson [33, Theorem 2] that, if $\omega > 0, \alpha > 0$ and $\beta \geq 0$, the LGARCH(1,1) model admits a strictly stationary solution if and only if $-\infty < E[\log(\beta+\alpha\epsilon_0^2)] < 0$. Moreover, he proves the uniqueness of this solution. A generalization to arbitrary GARCH(p,q) process is given by Bougerol and Picard [7]. Conclude from Jensen's inequality that

$$E[\log(\beta + \alpha\epsilon_0^2)] \leq \log(\beta + \alpha E[\epsilon_0^2]) < 0.$$

Therefore, if $E[\epsilon_0^2] = 1$ is assumed, a sufficient criterion for the existence of a strictly stationary solution is given by $\alpha + \beta < 1$. Yet, this is also the condition for weak stationarity of the LGARCH(1,1) process, cf. Bollerslev [5, Theorem 1]. Thus, the existence of a weakly stationary solution implies the existence of a strictly stationary solution provided ϵ_t is an i.i.d. sequence and both solutions must coincide. In particular, if $\alpha + \beta = 1$ the LGARCH(1,1) is sometimes called an integrated GARCH(1,1) (IGARCH(1,1)) model. This process is strictly stationary but not weakly stationary as $Var(X_0) = \infty$. Note that ergodicity is obtained because h is a measurable function of the i.i.d. sequence ϵ_t. The next theorem gives conditions for the existence of $E[h(\sigma_0^2)^s]$ and $E[X_0^s]$ for some $s > 0$. This result can be found in Aue et al. [1, Theorem 2.2] and Hörmann [23, Theorem 3].

Theorem 2.2. *Let $(X_t, h_t)_{t\in\mathbb{Z}}$ be strictly stationary, ϵ_t an i.i.d. sequence and $E[|g(\epsilon_0)|^s] < \infty$ for some $s > 0$.*

1. *If $E[|c(\epsilon_0)|^s] < 1$, then $E[|h(\sigma_0^2)|^s] < \infty$ holds.*
2. *If $c(\epsilon_0) \geq 0$, $g(\epsilon_0) \geq 0$ a.s. and $E[|h(\sigma_0^2)|^s] < \infty$, then $E[|c(\epsilon_0)|^s] < 1$ follows.*

If $E[\epsilon_0^2] = 1$ is assumed, then in the LGARCH case it is well-known that $E[|c(\epsilon_0)|] = E[\alpha\epsilon_0^2 + \beta] = \alpha + \beta < 1$ is necessary and sufficient for $E[\sigma_0^2] < \infty$ and, thus, $E[X_0^2] < \infty$.

3. Main results

Suppose, the volatility function h_t depends on an unknown parameter-vector $\boldsymbol{\theta} \in \mathbb{R}^d$; i.e.

$$h_t(\boldsymbol{\theta}) := c_{\boldsymbol{\theta}}(\epsilon_{t-1})h_{t-1}(\boldsymbol{\theta}) + g_{\boldsymbol{\theta}}(\epsilon_{t-1}). \tag{3.1}$$

The parameter may appear in the function $c(x)$ or $g(x)$. For instance, consider the LGARCH(1,1) model. Then, the parameter of interest is given by $\boldsymbol{\theta} = (\omega, \alpha, \beta)$ and $c_{\boldsymbol{\theta}}(x) = \alpha x + \beta$ and $g_{\boldsymbol{\theta}} = \omega$. In order to guarantee that the process X_t is strictly stationary and ergodic it is assumed that the conditions of Theorem 2.1 are met. The conditional variance process has the following unique a.s. representation:

$$\sigma_t^2(\boldsymbol{\theta}) = h^{-1}\left(\sum_{i=1}^{\infty} g_{\boldsymbol{\theta}}(\epsilon_{t-1-i}) \prod_{1\leq j<i} c_{\boldsymbol{\theta}}(\epsilon_{t-i})\right). \tag{3.2}$$

In this section strong consistency and asymptotic normality of the QMLE in augmented GARCH(1,1) models is derived. The true, but unknown parameter is denoted by $\boldsymbol{\theta}_0$. When working with real data, only finitely many data points are observed. Therefore, it is common to consider the finite sample version to (3.2), that is,

$$\tilde{\sigma}_t^2 = h^{-1}\left(\sum_{i=1}^{t} g_{\boldsymbol{\theta}}(\epsilon_{t-1-i}) \prod_{1\leq j<i} c_{\boldsymbol{\theta}}(\epsilon_{t-i})\right), \tag{3.3}$$

cf. Berkes et al. [3], Francq and Zakoïan [15], Straumann and Mikosch [42] and Hörmann [23]. The logarithm of the Gaussian likelihood is given by

$$\tilde{L}_n(\boldsymbol{\theta}) = \sum_{t=1}^{n} \tilde{l}_t := -1/2 \sum_{t=1}^{n} \log \tilde{\sigma}_t^2(\boldsymbol{\theta}) + \frac{X_t^2}{\tilde{\sigma}_t^2(\boldsymbol{\theta})} \tag{3.4}$$

modulo a constant. A QMLE is obtained as the measurable solution $\hat{\boldsymbol{\theta}}_n$ of

$$\hat{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta}\in K} n^{-1}\tilde{L}_n(\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta}\in K} n^{-1} \sum_{t=1}^{n} \log \tilde{\sigma}_t^2 + \frac{X_t^2}{\tilde{\sigma}_t^2}, \tag{3.5}$$

where K is some compact set of $\mathbb{R}^d$. Note that $X_t = \sigma_t(\boldsymbol{\theta}_0)\epsilon_t$. However, since $\tilde{\sigma}_t^2$ is neither stationary nor ergodic, cf. Straumann and Mikosch [42], it is necessary to work with the ergodic and stationary approximation σ_t^2 from equation (3.2) instead of $\tilde{\sigma}_t$. Denote

$$L_n(\boldsymbol{\theta}) = \sum_{t=1}^{n} l_t := -1/2 \sum_{t=1}^{n} \log \sigma_t^2(\boldsymbol{\theta}) + \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})}$$

the Gaussian likelihood if σ_t^2 is employed instead of $\tilde{\sigma}_t^2$ and, similarly,

$$L(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}_0}[l_0(\boldsymbol{\theta})].$$

The results of this section are derived using the following two steps. First, assuming σ_t^2 is observable, consistency of an estimator $\hat{\boldsymbol{\theta}}_n^*$ is established that is obtained by replacing $\tilde{\sigma}_t^2$ in equation (3.5) with σ_t^2. Second, it is proven that the estimator $\hat{\boldsymbol{\theta}}_n$ of (3.5) converges a.s. to $\hat{\boldsymbol{\theta}}_n^*$. This procedure is standard in the literature, cf. Berkes et al. [3], Francq and Zakoïan [15] or Straumann and Mikosch [42]. To ensure that the conditions of Theorem 2.1 hold, it is necessary to restrict the feasible values of the parameter $\boldsymbol{\theta}_0$. Let $\boldsymbol{\theta}_0 \in K$, such that $(X_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic and let K be some compact subset of $\mathbb{R}^d$. Moreover, let $h(\sigma_t^2) = (\sigma_t^2)^\delta$, for some $\delta > 0$. In order to derive strong consistency the following restrictions on $c_{\boldsymbol{\theta}}(x)$ and $g_{\boldsymbol{\theta}}(x)$ are imposed.

(C1) It holds that a.s. $c_{\boldsymbol{\theta}_0}(\epsilon_0) \geq 0$, $g_{\boldsymbol{\theta}_0}(\epsilon_0) \geq 0$ and $\sigma_0^2(\boldsymbol{\theta}) \geq \bar{h} > 0$ for all $\boldsymbol{\theta} \in K$.
(C2) Additionally, it holds that $E_{\boldsymbol{\theta}_0}[|c_{\boldsymbol{\theta}_0}(\epsilon_0)|^s] < 1$ and $E_{\boldsymbol{\theta}_0}[|g_{\boldsymbol{\theta}_0}(\epsilon_0)|^s] < \infty$ for some $s > 0$.
(C3) The function $h(\sigma_t^2(\boldsymbol{\theta})) = \sigma_t^{2\delta}(\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta} \in K$ for all t.
(C4) For every $\boldsymbol{\theta} \in K$ the following identifier condition holds: $\sigma_0^2(\boldsymbol{\theta}) \overset{a.s.}{=} \sigma_0^2(\boldsymbol{\theta}_0)$ if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

The next theorem establishes consistency of the QMLE for augmented GARCH models with the choice $h(x) = x^\delta$. Observe that $E_{\boldsymbol{\theta}_0}[|g_{\boldsymbol{\theta}_0}(\epsilon_0)|^s] < \infty$ implies $E_{\boldsymbol{\theta}_0}[\log^+ |g_{\boldsymbol{\theta}_0}(\epsilon_0)|] < \infty$.

Theorem 3.1. *Let $0 < \delta$ and let $\boldsymbol{\theta}_0 \in K$ for some compact set $K \subset \mathbb{R}^d$. Assume that the process $(X_t)_{t\in\mathbb{Z}}$ is strictly stationary and (C1)-(C4). Let $E[\epsilon_0] = 0$ and $E[\epsilon_0^2] = 1$. Then, it follows that $\hat{\boldsymbol{\theta}}_n \overset{a.s.}{\to} \boldsymbol{\theta}_0$.*

***Remark* 3.1.** To the best of the author's knowledge, consistency of the QMLE for the class of augmented GARCH(1,1) models has not been studied before. Straumann and Mikosch [42] considered a general class of GARCH models including the LGARCH model and asymmetric GARCH model using a stochastic version of Banach's fixed point theorem, cf. Straumann and Mikosch [42, Theorem 2.8]. However, their conditions are hard to verify and detailed discussions are necessary for every specific model, cf. Straumann and Mikosch [42, Section 5]. In contrast to their assumptions, the conditions (C1)-(C4) are verified by straightforward calculations as soon as a specific augmented GARCH model is considered.

***Remark* 3.2.** The assumption $E[\epsilon_0] = 0$ is only made for convenience as it corresponds to the efficient market hypothesis $E[X_t|\mathfrak{F}_{t-1}] = 0$ a.s., cf. White [45]. In Francq and Zakoïan [15] the assumption $E[\epsilon_0] = 0$ is dropped, cf. their Remark 2.5. The assumption $E[\epsilon_t^2] = 1$ is not restrictive as long as $E[\epsilon_t^2] < \infty$, cf. Berkes and Horváth [2].

To establish asymptotic normality, it is convenient to make a first order Taylor series expansion of $\nabla_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}_n^*)$, cf. Straumann and Mikosch [42] to obtain for n large enough:

$$\nabla_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}_n^*) = \nabla_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}_0) + \nabla_{\boldsymbol{\theta}}^2 L_n(\boldsymbol{\xi})(\boldsymbol{\theta}_n^* - \boldsymbol{\theta}_0), \tag{3.6}$$

here $\boldsymbol{\xi}$ is between $\boldsymbol{\theta}_n^*$ and $\boldsymbol{\theta}_0$. Since $\boldsymbol{\theta}_n^*$ is the unique maximum of $L_n(\boldsymbol{\theta})$, it follows that $\nabla_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}_n^*) = \mathbf{0}$. If

$$n^{-1}\nabla_{\boldsymbol{\theta}}^2 L_n(\boldsymbol{\xi}) \overset{a.s.}{\to} \boldsymbol{B}_0^{-1} := E_{\boldsymbol{\theta}_0}[\nabla_{\boldsymbol{\theta}_0}^2 l_0(\boldsymbol{\theta}_0)] \tag{3.7}$$

rearranging equation (3.6), yields

$$\begin{aligned} n^{-1}\nabla_{\boldsymbol{\theta}}^2 L_n^{-1}(\boldsymbol{\xi})(\boldsymbol{\theta}_n^* - \boldsymbol{\theta}_0) &= -n^{-1}\nabla_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}_0) \\ \sqrt{n}(\boldsymbol{\theta}_n^* - \boldsymbol{\theta}_0) &= -\boldsymbol{B}_0^{-1}(1 + o_P(1))n^{-1/2}\nabla_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}_0). \end{aligned} \tag{3.8}$$

Applying a central limit theorem for martingale differences, cf. Billingsley [4, Theorem 18.3] or Heyde [22, Theorem 2] to the sum $n^{-1/2}\nabla_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}_0)$ yields

$$-\boldsymbol{B}_0^{-1}(1 + o_P(1))n^{-1/2}\nabla_{\boldsymbol{\theta}} L_n(\boldsymbol{\theta}_0) \overset{d}{\to} N(\mathbf{0}, \boldsymbol{B}_0^{-1}\boldsymbol{S}\boldsymbol{B}_0^{-1}),$$

where

$$\boldsymbol{S} = E_{\boldsymbol{\theta}_0}\left[\nabla_{\boldsymbol{\theta}} l_0(\boldsymbol{\theta}_0)\nabla_{\boldsymbol{\theta}}^{'} l_0(\boldsymbol{\theta}_0)\right].$$

To ensure that the variance-covariance matrix is well defined, it is assumed that

$$E_{\boldsymbol{\theta}_0}[\|\nabla_{\boldsymbol{\theta}} l_0(\boldsymbol{\theta}_0)\|] < \infty. \tag{3.9}$$

In addition, the condition

$$E_{\boldsymbol{\theta}_0}[\|\nabla_{\boldsymbol{\theta}}^2 l_0(\boldsymbol{\theta}_0)\|_{\tilde{K}}] < \infty \tag{3.10}$$

for some compact $\tilde{K}$ containing the true parameter $\boldsymbol{\theta}_0$ is employed to show that (3.7) holds. These assumptions are along the lines of Straumann and Mikosch [42] or Ferguson [14, Theorem 18]. However, Lumsdaine [30], Berkes et al. [3], Francq and Zakoïan [15] and Jensen and Rahbek [26] work with finite expectation of the third derivatives of l_0 in a neighborhood of the true parameter in order to bound the second derivatives. This may lead to some cumbersome calculations and is not necessary. In order to establish asymptotic normality, the derivatives of the likelihood function

$$l_t(\boldsymbol{\theta}) = -1/2\left(\log\sigma_t^2(\boldsymbol{\theta}) + \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})}\right) \tag{3.11}$$

are needed. Recall that for the augmented GARCH(1,1) model σ_t^2 is uniquely determined by the relationship

$$\sigma_t^2 = h^{-1}\left(\sum_{i=1}^{\infty} g_{\boldsymbol{\theta}}(\epsilon_{t-i})\prod_{j=1}^{i-1} c_{\boldsymbol{\theta}}(\epsilon_{t-j})\right) = \left(\sum_{i=1}^{\infty} g_{\boldsymbol{\theta}}(\epsilon_{t-i})\prod_{j=1}^{i-1} c_{\boldsymbol{\theta}}(\epsilon_{t-j})\right)^{1/\delta}.$$

Therefore, the partial derivative is obtained

$$\begin{aligned}
\nabla_{\boldsymbol{\theta}}\sigma_t^2 &= 1/\delta\left(\sum_{i=1}^{\infty} g_{\boldsymbol{\theta}}(\epsilon_{t-i})\prod_{j=1}^{i-1} c_{\boldsymbol{\theta}}(\epsilon_{t-j})\right)^{1/\delta-1}\nabla_{\boldsymbol{\theta}}\left(\sum_{i=1}^{\infty} g_{\boldsymbol{\theta}}(\epsilon_{t-i})\prod_{j=1}^{i-1} c_{\boldsymbol{\theta}}(\epsilon_{t-j})\right) \\
&= 1/\delta\left(\left(\sum_{i=1}^{\infty} g_{\boldsymbol{\theta}}(\epsilon_{t-i})\prod_{j=1}^{i-1} c_{\boldsymbol{\theta}}(\epsilon_{t-j})\right)^{1/\delta}\right)^{1-\delta}\nabla_{\boldsymbol{\theta}} h(\sigma_t^2) \\
&= 1/\delta(\sigma_t^2)^{1-\delta}\nabla_{\boldsymbol{\theta}}(\sigma_t^2)^{\delta}.
\end{aligned} \tag{3.12}$$

For instance, assuming $\delta = 1/2$, yields $\nabla_{\boldsymbol{\theta}}\sigma_t^2 = 2\sigma_t\nabla_{\boldsymbol{\theta}}\sigma_t$ and for $\delta = 1$ the above expression simplifies to $\nabla_{\boldsymbol{\theta}}\sigma_t^2 = \nabla_{\boldsymbol{\theta}}\sigma_t^2$. Thus, the first derivative of $l_t(\boldsymbol{\theta})$ is given by

$$\nabla_{\boldsymbol{\theta}} l_t(\boldsymbol{\theta}) = -1/2\left(1 - \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})}\right)\frac{\nabla_{\boldsymbol{\theta}}\sigma_t^2(\boldsymbol{\theta})}{\sigma_t^2(\boldsymbol{\theta})} \tag{3.13}$$

and its second derivative by

$$\begin{aligned}
\nabla_{\boldsymbol{\theta}}^2 l_t(\boldsymbol{\theta}) &= -1/2\left(1 - \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})}\right)\left(\frac{1}{\sigma_t^2(\boldsymbol{\theta})}\nabla_{\boldsymbol{\theta}}^2\sigma_t^2(\boldsymbol{\theta})\right) \\
&\quad - 1/2\left(2\frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})} - 1\right)\left(\frac{1}{\sigma_t^4(\boldsymbol{\theta})}\nabla_{\boldsymbol{\theta}}\sigma_t^2(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}^{'}\sigma_t^2(\boldsymbol{\theta})\right).
\end{aligned} \tag{3.14}$$

The existence of both derivatives is an immediate consequence of the differentiability of $\sigma_t^2(\boldsymbol{\theta})$. The following assumptions are needed to ensure the asymptotic normality of the QMLE.

(N1) The true parameter $\boldsymbol{\theta}_0$ lies in the interior of K, denoted with $\mathring{K}$.
(N2) The distribution of ϵ_t is such that $E[\epsilon_0^2] = 1$ and $E[\epsilon_0^4] < \infty$.
(N3) There exists a convex set $\tilde{K} \subset K$, containing $\boldsymbol{\theta}_0$, such that $\sigma_t^2(\boldsymbol{\theta})$ is three times continuously differentiable in $\boldsymbol{\theta}$ with measurable derivatives such that

(i) $E_{\boldsymbol{\theta}_0}[\log^+ \|\nabla_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta})\|_{\tilde{K}}] < \infty$,
(ii) $E_{\boldsymbol{\theta}_0}[\log^+ \|\nabla^2_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta})\|_{\tilde{K}}] < \infty$,
(iii) $E_{\boldsymbol{\theta}_0}[\log^+ \|\nabla^3_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta})\|_{\tilde{K}}] < \infty$.

In addition, the following moment conditions hold:

(iv) $E_{\boldsymbol{\theta}_0}[\|\nabla^2_{\boldsymbol{\theta}} l_0(\boldsymbol{\theta})\|_{\tilde{K}}] < \infty$,
(v) $E_{\boldsymbol{\theta}_0}\left[\left\|\frac{1}{\sigma_0^2(\boldsymbol{\theta}_0)} \nabla_{\boldsymbol{\theta}} \sigma_0^2(\boldsymbol{\theta}_0)\right\|\right] < \infty$,
(vi) $E_{\boldsymbol{\theta}_0}\left[\left\|\frac{1}{\sigma_0^2(\boldsymbol{\theta}_0)} \nabla^2_{\boldsymbol{\theta}} \sigma_0^2(\boldsymbol{\theta}_0)\right\|\right] < \infty$,
(vii) $E_{\boldsymbol{\theta}_0}\left[\left\|\frac{1}{\sigma_0^4(\boldsymbol{\theta}_0)} \nabla_{\boldsymbol{\theta}} \sigma_0^2(\boldsymbol{\theta}_0) \nabla'_{\boldsymbol{\theta}} \sigma_0^2(\boldsymbol{\theta}_0)\right\|\right] < \infty$.

(N4) The components of $\nabla_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta})$ are linearly independent random variables.

Assumption (N3) may easily fail to hold. For instance, in the TGARCH model of Zakoïan [46], $\sigma_t^2(\boldsymbol{\theta})$ is not continuous differentiable. Thus, the moment conditions (N3) have to be verified directly. For simplicity, it is assumed without loss of generality that $\tilde{K} = K$, similar to Straumann [41, p. 116]. The next theorem establishes the asymptotic normality of the QMLE.

Theorem 3.2. *Let $\delta > 0$ and let $\boldsymbol{\theta}_0 \in \mathring{K}$ for some compact set $K \subset \mathbb{R}^d$. Assume the conditions of Theorem 3.1 and (N1)-(N4). Then, it follows that*

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\Sigma} = E[(\epsilon_0^4 - 1)] E_{\boldsymbol{\theta}_0}\left[\frac{1}{\sigma_0^4(\boldsymbol{\theta})} \nabla_{\boldsymbol{\theta}} \sigma_0^2(\boldsymbol{\theta}) \nabla'_{\boldsymbol{\theta}} \sigma_0^2(\boldsymbol{\theta})\right]^{-1}.$$

***Remark* 3.3.** Assumption (N1) and (N2) are standard assumption, cf. Berkes et al. [3], Francq and Zakoïan [15] or Straumann and Mikosch [42] and are necessary to ensure that $\boldsymbol{\Sigma}$ is well defined and that the support of $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$ is $(-\infty, \infty)^d$. The assumptions (N3) and (N4) are similar to assumptions (N3) and (N4) in Straumann and Mikosch [42]. They are needed to demonstrate that $\boldsymbol{\Sigma}$ is well defined and that $\hat{\boldsymbol{\theta}}_n$ is asymptotic equivalent to $\hat{\boldsymbol{\theta}}_n^*$. The moment conditions (N3i)-(N3iv) are similar to conditions (D1)-(D3) in Straumann and Mikosch [42] and they may be easily verified as soon as a special GARCH model is considered. Therefore, it suffices to show that for some $s > 0$ $E[\|Y\|_K^s] < \infty$ because $E[\|Y\|_K^s] < \infty$ implies $E[\log^+ \|Y\|_K] < \infty$ and Y stands for one the derivatives of σ_t^2.

Model	Specification
LGARCH(1,1) of Bollerslev [5]	$\sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta\sigma_{t-1}^2$
AGARCH(1,1) of Ding et al. [9]	$\sigma_t^2 = \omega + \alpha(\lvert X_{t-1}\rvert - \gamma X_{t-1})^2 + \beta\sigma_{t-1}^2$
APGARCH(1,1) of Ding et al. [9]	$\sigma_t^d = \omega + \alpha((\lvert X_{t-1}\rvert - \gamma X_{t-1})^d + \beta\sigma_{t-1}^d$
NGARCH(1,1) of Engle and Ng [12]	$\sigma_t^2 = \omega + \alpha(\epsilon_{t-1} - c)^2\sigma_{t-1}^2 + \beta\sigma_{t-1}^2$
VGARCH(1,1) of Engle and Ng [12]	$\sigma_t = \omega + \alpha(\epsilon_{t-1} - c)^2 + \beta\sigma_{t-1}^2$
TSGARCH(1,1) of Schwert [40]	$\sigma_t = \omega + \alpha\lvert X_{t-1}\rvert + \beta\sigma_{t-1}$

Table 1. GARCH specification

***Remark* 3.4.** To the author's knowledge asymptotic normality of augmented GARCH models has not been considered before. Table 1 shows some examples of GARCH specification for which the result applies.

The TGARCH model of Zakoïan [46], however, does not fulfill assumption (N3) because it is not continuously differentiable. The same holds true for the more general class of power transformed TGARCH models as discussed among others in Hwang and Basawa [24], Pan et al. [35] and Hamadeh and Zakoïan [17]. However, the continuity of the derivatives is only needed to apply a multivariate mean value theorem in Lemma 5.5. When applied to the LGARCH(1,1) model, the assumptions actually coincide with the findings in Francq and Zakoïan [15] and applied to the AGARCH model the assumptions coincide with the ones given in Straumann and Mikosch [42]. In addition, asymptotic normality of the QMLE for the NGARCH, VGARCH, the TSGARCH model and other polynomial GARCH models may be derived which seemingly has not been established before.

From Theorem 3.2 standard errors may be calculated using the variance-covariance matrix $\mathbf{\Sigma}$. However, $\mathbf{\Sigma}$ can not be calculated explicitly as the common distribution of an augmented GARCH(1,1) model is not known. Thus, it seems reasonable to estimate $\mathbf{\Sigma}$ via a matrix $\tilde{\boldsymbol{V}}_n$ defined by:

$$\tilde{\boldsymbol{V}}_n(\hat{\boldsymbol{\theta}}_n) = \left(\frac{1}{n}\sum_{t=1}^{n}(\tilde{\epsilon_t}^4 - 1)\right)\left(\frac{1}{n}\sum_{t=1}^{n}\frac{1}{\tilde{\sigma}_t^4(\hat{\boldsymbol{\theta}}_n)}\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2(\hat{\boldsymbol{\theta}}_n)\nabla'_{\boldsymbol{\theta}}\tilde{\sigma}_t^2(\hat{\boldsymbol{\theta}}_n)\right)^{-1}, \tag{3.15}$$

where $\tilde{\epsilon_t} = X_t/\tilde{\sigma}_t(\hat{\boldsymbol{\theta}}_n)$. One can show, cf. Straumann and Mikosch [42, Remark 7.5] that $\tilde{\boldsymbol{V}}_n(\boldsymbol{\theta})$ is a strongly consistent estimator. Therefore, let $\boldsymbol{V}_n(\hat{\boldsymbol{\theta}}_n)$ be as $\tilde{\boldsymbol{V}}_n(\hat{\boldsymbol{\theta}}_n)$ with $\tilde{\sigma_t}$ replaced by σ_t. In a first step it can be shown that $\boldsymbol{V}_n(\hat{\boldsymbol{\theta}}_n)$ is strongly consistent for $\mathbf{\Sigma}$. In a second step, it is proven that

$$\|\tilde{\boldsymbol{V}}_n(\boldsymbol{\theta}) - \boldsymbol{V}_n(\boldsymbol{\theta})\|_K \overset{a.s.}{\to} 0.$$

The following result is mentioned in Straumann and Mikosch [42, Remark 7.5] and implicitly in Mukherjee [32, Proposition 3.11]. A proof is therefore omitted.

***Proposition* 3.3.** Assume the conditions of Theorem 3.2. Then, the estimator defined in (3.15) is strongly consistent for $\mathbf{\Sigma}$.

4. Applications

4.1. The QMLE in the LGARCH model

Let $K_c = [c, 1/c]^2 \times [0, 1-c]$ for some arbitrary small $c > 0$ such that $(X_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic for $\boldsymbol{\theta}_0 \in K_c$. Recall that the LGARCH(1,1) is obtained by setting in model (2.1) $g(\epsilon_t) = \omega$, $h(\sigma_t^2) = \sigma_t^2$ and $c(\epsilon_t) = \alpha + \beta\epsilon_t^2$. The next corollary shows consistency of the QMLE.

***Corollary* 4.1.** Let ϵ_t be an i.i.d. sequence, not concentrated at two points with $E[\epsilon_0] = 0$ and $E[\epsilon_0^2] = 1$. Suppose the true parameter $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0)$ is in K_c such that 2.5 holds. Then $\hat{\boldsymbol{\theta}}_n \overset{a.s.}{\to} \boldsymbol{\theta}_0$ if $\sigma_0(\boldsymbol{\theta}) \overset{a.s.}{=} \sigma_0(\boldsymbol{\theta}_0)$.

***Remark* 4.1.** Some cases are ruled out from consistency considerations. For example assume $\alpha = 0$, then $\sigma_t^2 = \omega + \beta\sigma_{t-1}^2$ and σ_t^2 is purely deterministic and independent of $(X_t)_{t\in\mathbb{Z}}$. Thus, ω and β are not identifiable. Nevertheless, $\beta = 0$ is allowed and essentially an ARCH(1) model is estimated. This result carries over to the GARCH(p,q) case as it is shown for instance in Francq and Zakoïan [15] under additional assumptions.

Proof of corollary 4.1. One has to show that (C1)-(C4) of Theorem 3.1 hold under the assumptions of Corollary 4.1. Since $g_{\boldsymbol{\theta}}(\epsilon_t) = \omega > 0$ and $c_{\boldsymbol{\theta}}(\epsilon_t) = \alpha\epsilon_t^2 + \beta > 0$, for $\alpha > 0, \beta \geq 0$ (C1) follows immediately. (C3) is obvious because $h(\sigma_t(\boldsymbol{\theta})) = \sigma_t^2(\boldsymbol{\theta}) = \omega + \alpha X_{t-1}^2 + \beta\sigma_{t-1}^2$ is a continuous function in $\boldsymbol{\theta} = (\omega, \alpha, \beta)$. Under the condition $\sigma_0(\boldsymbol{\theta}) \overset{a.s.}{=} \sigma_0(\boldsymbol{\theta}_0)$ it follows that $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ by the proof of Francq and Zakoïan [15, Theorem 2.1] and (C4) is fulfilled. It is left to show that (C2) holds. The second part of (C2) namely $E[|g_{\boldsymbol{\theta}}(\epsilon_t)|^s] = E[\omega^s] < \infty$ holds trivially for all $\omega > 0$ and $s > 0$. Additionally, deduce from Berkes et al. [3, Lemma 2.3] that if the LGARCH process is strictly stationary and ergodic there exists some $s > 0$ such that $E[\sigma_t^{2s}] < \infty$. Since (C1) holds an application of the second part of Theorem 2.2 yields that $E[|c_{\boldsymbol{\theta}}(\epsilon_t)|^s] < 1$. This completes the proof. □

The next corollary shows the asymptotic normality of the QMLE. The assumptions are the same as in Francq and Zakoïan [15, Theorem 2.2].

***Corollary* 4.2.** Assume the conditions of Corollary 4.1, $E[\epsilon_0^4] < \infty$ and, in addition, $\boldsymbol{\theta}_0 \in \mathring{K}_c$. Then, it follows that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \overset{d}{\to} N(\mathbf{0}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\Sigma} = E[(\epsilon_0^4 - 1)]E_{\boldsymbol{\theta}_0}\left[\frac{1}{\sigma_0^4(\boldsymbol{\theta})}\nabla_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}^{'}\sigma_0^2(\boldsymbol{\theta})\right]^{-1}.$$

Proof. The conditions (N1)-(N4) of Theorem 3.2 have to be verified. (N1) and (N2) are immediate. (N3iv)-(N3vii) are demonstrated in the proof of Francq and Zakoïan [15, Theorem 2.2] and (N4) is proven in Berkes et al. [3, Lemma 5.7] and Francq and Zakoïan [15, Theorem 2.2], respectively. It is left to verify conditions (N3i)-(N3iii). It is merely demonstrated that (N3i) holds since the remaining conditions are verified analogously. Since the process is assumed to be strictly stationary, it follows from Berkes et al. [3, Lemma 2.3] that there exists some $s > 0$ such that $E[\sigma_t^{2s}] < \infty$ and, thus, $E[X_t^{2s}] < \infty$. In order to establish (N3i), it may fruitful to discuss the almost sure representation of $\nabla_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta})$ which is obtained from the ARCH(∞) representation of σ_t^2 similar to Berkes et al. [3] and Francq and Zakoïan [15, equation (4.4)]. Hence, the first partial derivative of σ_t^2 is given by

$$\nabla_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta}) = \begin{pmatrix} \frac{1}{1-\beta} \\ \sum_{i=1}^{\infty} \beta^{i-1} X_{t-i}^2 \\ \frac{\omega}{(1-\beta)^2} + \alpha \sum_{i=1}^{\infty} (i-1)\beta^{i-2} X_{t-i}^2 \end{pmatrix}.$$

Observe that

$$E_{\boldsymbol{\theta}_0}[\log^+ \|(1-\beta)^{-1}\|_{K_c}] \le -\log^+ c < \infty.$$

Turning to $\frac{\partial \sigma_t^2}{\partial \alpha}$, it follows that $\frac{\partial \sigma_t^2}{\partial \alpha} \le \sigma_t^2/\alpha$. Using Berkes et al. [3, Lemma 2.3], there exists a $s > 0$ such that $E[\sigma_t^{2s}] < \infty$. Therefore, it holds with Jensen's inequality that

$$\begin{aligned} E_{\boldsymbol{\theta}_0}\left[\log^+ \left\| \frac{\partial \sigma_t^2}{\partial \alpha} \right\|_{K_c}\right] &\le E_{\boldsymbol{\theta}_0}\left[\log^+ \sigma_t^2 - \log^+ \|\alpha\|_{K_c}\right] \\ &\le 1/s \log^+ E_{\boldsymbol{\theta}_0}[\sigma_t^{2s}] - \log^+ c < \infty. \end{aligned}$$

From Hardy et al. [19, p. 24], it can be deduced that $(a+b)^s \le a^s + b^s$ for $0 < s < 1$ and, thus, together with Straumann and Mikosch [42, Lemma 2.2.] it follows that

$$\begin{aligned} E_{\boldsymbol{\theta}_0}\left[\log^+ \left\| \frac{\partial \sigma_t^2}{\partial \beta} \right\|_{K_c}\right] &\le s^{-1} \log^+ E_{\boldsymbol{\theta}_0}\left[\left(\left\| \frac{\omega}{(1-\beta)^2} \right\|_{K_c} + \alpha \sum_{i=1}^{\infty} (i-1) \|\beta^{i-2} X_{t-i}^2\|_{K_c}\right)^s\right] \\ &\le C_1 + \log^+ \sum_{i=1}^{\infty} (i-1)(1-c)^{(i-2)} E[X_{t-i}^{2s}] < \infty, \end{aligned}$$

where $C_1 := s^{-1} \log^+ \left(\frac{\omega}{(1-\beta)^2}\right)^s + \log^+ \alpha + 2\log 2 < \infty$ and the series converges by an application of Lemma 5.1. Thus, Theorem 3.2 applies and the assertion follows. $\square$

It is possible to derive consistency and asymptotic normality under weaker assumptions on the innovation process ϵ_t. Especially, the i.i.d. assumption may be dropped, cf. Lee and Hansen [29] or Escanciano [13]. An alternative is to employ a martingale difference structure for the stationary and ergodic sequence ϵ_t together with stronger moment assumptions as in Lee and Hansen [29] and Escanciano [13]. However, it seems questionable whether weak dependence assumptions on ϵ_t are indeed that important as asserted by Escanciano [13]. The main argument is that the i.i.d. assumption does not allow for time-varying skewness and kurtosis. Though Rockinger and Jondeau [38] and Jondeau and Rockinger [27] among others claim this empirical finding, it is not an established stylized fact and empirical studies are contradictory, cf. Herrmann [21] for an overview. Therefore, the classical framework is considered and the i.i.d. assumption for ϵ_t is upheld.

4.2. The QMLE in the TGARCH model

Zakoïan [46] proposed an asymmetric GARCH model allowing for modeling the leverage effect, i.e. that negative stock returns have a stronger influence on the volatility than positive ones. A phenomenon often observed for financial market data, cf. McNeil et al. [31]. Consistency and asymptotic normality of the QMLE are derived under similar conditions as in the LGARCH(1,1) case. However, some problems arise because $\sigma_t(\boldsymbol{\theta})$ is not continuous differentiable at least in the third component. The TGARCH(1,1) model is obtained by setting $g_{\boldsymbol{\theta}}(\epsilon_t) = \omega$, $c_{\boldsymbol{\theta}}(\epsilon_t) = \alpha^+|\epsilon_t| + \alpha^- \max(0, -\epsilon_t) + \beta$ and $\delta = 1/2$ with $\omega > 0, \alpha_1^+, \alpha_1^-, \beta \geq 0$ for positivity of σ_t. Consistency and asymptotic normality of the QMLE of this type of model was proven by Pan et al. [35] in a more general setting, namely the power transformed TGARCH model and more recently it was proven again using slightly different arguments by Hamadeh and Zakoïan [17]. Since the TGARCH model is not continuous differentiability of σ_t, the results of the previous section do not directly apply. Its first derivative w.r.t. $\boldsymbol{\theta}$ is given by:

$$\nabla_{\boldsymbol{\theta}} \sigma_t(\boldsymbol{\theta}) = (1, |X_{t-1}|, \max(0, -X_{t-1}), \sigma_{t-1})'$$

which is not continuous in its third component. Hence, assumption (N3) no longer holds. Defining $X_{t-1}^- := \max(0, -X_{t-1})$, σ_t may be written as:

$$\begin{aligned}\sigma_t &= \omega + \alpha^+|X_{t-1}| + \alpha^- X_{t-1}^- + \beta(\omega + \alpha^+ X_{t-2}^+ + \alpha^- X_{t-2}^- + \beta\sigma_{t-2}) \\ &= \omega(1 + \beta + \beta^2 + ...) + \alpha^+(|X_{t-1}| + \beta|X_{t-2}| + \beta^2|X_{t-3}| + ...) \\ &+ \alpha^-(X_{t-1}^- + \beta X_{t-2}^- + \beta^2 X_{t-3}^- + ...) \\ &= \frac{\omega}{1-\beta} + \alpha^+ \sum_{i=1}^{\infty} \beta^{i-1}|X_{t-i}| + \alpha^- \sum_{i=1}^{\infty} \beta^{i-1} X_{t-i}^-. \end{aligned} \tag{4.1}$$

Thus, a strictly stationary TGARCH process $(X_t)_{t\in\mathbb{Z}}$ can be represented as TARCH(∞) process. Again, as in the LGARCH setting it is worked with the approximation $\tilde{\sigma}_t$ for σ_t, based on a finite sample $X_1, ..., X_n$ of $(X_t)_{t\in\mathbb{Z}}$. Let

$$K_d = [d, 1/d]^3 \times [0, 1-d]$$

denote the compact parameter space for some $d > 0$. From Theorem 2.1 it can be deduced that the TGARCH(1,1) model is strictly stationary and ergodic with an a.s. convergent solution for σ_t if the condition

$$E_{\boldsymbol{\theta}_0}[\log(\beta_0 + \alpha_0^+|\epsilon_t| + \alpha_0^- \max(0, -\epsilon_t))] < 0 \tag{4.2}$$

holds. The next corollary states the consistency of the QMLE.

***Corollary* 4.3.** Let ϵ_t be an i.i.d. process, not concentrated at two points with $E[\epsilon_0] = 0$ and $E[\epsilon_0^2] = 1$. Suppose the true parameter $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0^+, \alpha_0^-, \beta_0)$ is in K_d such that (4.2) holds. Then, $\hat{\boldsymbol{\theta}}_n \overset{a.s.}{\to} \boldsymbol{\theta}_0$ if $\sigma_0(\boldsymbol{\theta}) \overset{a.s.}{=} \sigma_0(\boldsymbol{\theta}_0)$.

***Remark* 4.2.** Recently, Hamadeh and Zakoïan [17] proved this result for the power transformed TGARCH model of the form $X_t = \sigma_t^{1/\delta}\epsilon_t$. The present result is a very special case for $\delta = 1$. The case $\alpha^+ + \alpha^- = 0$ is excluded from the considerations because in that case σ_t is not irreducible.

Proof of corollary 4.3. The assumptions (C1) and (C3) are obviously fulfilled. (C4) is shown in Hamadeh and Zakoïan [17] because $\sigma_0(\boldsymbol{\theta}) \overset{a.s.}{=} \sigma_0(\boldsymbol{\theta}_0)$ implies $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, provided ϵ_t is an i.i.d. sequence, not concentrated at two points. Finally, (C2) holds because there exists a $s > 0$ such that $E[\omega^s] < \infty$ for all ω and $E[|\beta + \alpha^+|\epsilon_t| + \alpha^- \max(0, -\epsilon_t)|^s] < 1$, cf. Pan et al. [35, Theorem 6]. Thus, the conclusion follows from the second part of Theorem 2.2. □

The next theorem states the asymptotic normality of the QMLE for the parameters of the TGARCH(1,1) model. The same result was recently established by Hamadeh and Zakoïan [17] under the same conditions. However, some different arguments in the proof are used.

Theorem 4.4. *Assume the conditions of Corollary 4.3, $E[\epsilon_0^4] < \infty$ and, in addition, $\boldsymbol{\theta}_0 \in \mathring{K}_d$. Then, it follows that*

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \overset{d}{\to} N(\mathbf{0}, \boldsymbol{\Sigma}), \tag{4.3}$$

where

$$\boldsymbol{\Sigma} = E[(\epsilon_0^4 - 1)]E_{\boldsymbol{\theta}_0}\left[\frac{4}{\sigma_0^2(\boldsymbol{\theta}_0)}\nabla_{\boldsymbol{\theta}}\sigma_0(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_0(\boldsymbol{\theta}_0)\right]^{-1}.$$

5. Proofs

The following lemma from Straumann and Mikosch [42] will be applied abundantly in the following proofs.

***Lemma* 5.1.** Let $(f_t)_{t\in\mathbb{Z}}$ be a sequence of real random variables with $f_t \overset{e.a.s.}{\to} 0$ and let $(X_t)_{t\in\mathbb{Z}}$ be a sequence of identically distributed random variables in a separable Banach space $(B, \|\cdot\|)$. If $E[\log^+\|X_0\|] < \infty$, then $\sum_{t=0}^{\infty} f_t X_t$ converges a.s., and one has $f_n \sum_{t=0}^{n} X_t \overset{e.a.s.}{\to} 0$ and $f_n X_n \overset{e.a.s.}{\to} 0$ as $n \to \infty$.

5.1. Proof of Theorem 3.1

Before proving Theorem 3.1, two lemmas are needed. The next lemma shows consistency of the estimator $\hat{\boldsymbol{\theta}}_n^*$.

***Lemma* 5.2.** Under the conditions of Theorem 3.1 it follows for $n \to \infty$ that

$$\hat{\boldsymbol{\theta}}_n^* \overset{a.s.}{\to} \boldsymbol{\theta}_0.$$

Proof. The proof consists of the following steps. First, it is proven that

$$1/2 E_{\boldsymbol{\theta}_0}\left[\left|\log \sigma_0^2(\boldsymbol{\theta}_0) + \frac{X_0^2}{\sigma_0^2(\boldsymbol{\theta}_0)}\right|\right] = E_{\boldsymbol{\theta}_0}[|l_0(\boldsymbol{\theta}_0)|] < \infty. \tag{5.1}$$

Second, it is shown that for every $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ $L(\boldsymbol{\theta}) < L(\boldsymbol{\theta}_0)$ holds. Third, using an argument as in Pfanzagl [36] and Ferguson [14, Theorem 16b] the strong consistency follows. To demonstrate (5.1) observe that the conclusion follows if $E_{\boldsymbol{\theta}_0}[l_0^-(\boldsymbol{\theta}_0)] < \infty$ and $E_{\boldsymbol{\theta}_0}[l_0^+(\boldsymbol{\theta}_0)] < \infty$, where $l_0^+ = \max(0, l_0)$ and $l_0^- = \max(-l_0, 0)$. Under assumption (C1) it holds that

$$\begin{aligned} E_{\boldsymbol{\theta}_0}[l_0^+(\boldsymbol{\theta}_0)] &< E_{\boldsymbol{\theta}_0}[\log(\max(\sigma_0^2(\boldsymbol{\theta}_0), 1))] < \log E_{\boldsymbol{\theta}_0}[\max(\sigma_0^2(\boldsymbol{\theta}_0), 1)] \\ &\leq \max(0, -\log \bar{h}) < \infty. \end{aligned}$$

Additionally, under the conditions (C1) and (C2) there exists a $s > 0$ such that

$$E_{\boldsymbol{\theta}_0}[\log \sigma_0^2(\boldsymbol{\theta}_0)] = E_{\boldsymbol{\theta}_0}\left[1/s \log(\sigma_0^2(\boldsymbol{\theta}_0))^s\right] < 1/s \log E_{\boldsymbol{\theta}_0}[\left(\sigma_0^2(\boldsymbol{\theta})\right)^s].$$

It is left to show that

$$E_{\boldsymbol{\theta}_0}[\left(\sigma_0^2(\boldsymbol{\theta})\right)^s] = E_{\boldsymbol{\theta}_0}\left[\left(\sum_{i=1}^{\infty} g_{\boldsymbol{\theta}_0}(\epsilon_{-i}) \prod_{j=1}^{i-1} c_{\boldsymbol{\theta}_0}(\epsilon_{-j})\right)^{s/\delta}\right] < \infty. \tag{5.2}$$

Under assumption (C2) it follows from Hörmann [23, Theorem 3] that $E_{\boldsymbol{\theta}_0}[h(\sigma_0^2)^s] < \infty$ and for $\delta \geq 1$ an application of Jensen's inequality yields:

$$(E_{\boldsymbol{\theta}_0}[(\sigma_0^2)^s])^\delta \leq E_{\boldsymbol{\theta}_0}[h(\sigma_0^2)^s] < \infty$$

and equation (5.2) follows. For the case $0 < \delta < 1$ note that

$$E_{\boldsymbol{\theta}_0}[X_0^{2\delta}] = E_{\boldsymbol{\theta}_0}[\sigma_0^{2\delta}] E[\epsilon_0^{2\delta}] = E_{\boldsymbol{\theta}_0}[h(\sigma_0^2)] E[\epsilon_0^{2\delta}] < \infty$$

and (5.2) follows setting $s = \delta$ and an application of Hörmann [23, Theorem 3]. Hence, it follows that $E_{\boldsymbol{\theta}_0}[l_0^-(\boldsymbol{\theta}_0)] < \infty$ and (5.1) is proven. Next, it is shown that

$$-1/2 E_{\boldsymbol{\theta}_0}\left[\log \sigma_0^2(\boldsymbol{\theta}_0) + \frac{X_0^2}{\sigma_0^2(\boldsymbol{\theta}_0)}\right] \geq -1/2 E_{\boldsymbol{\theta}_0}\left[\log \sigma_0^2(\boldsymbol{\theta}) + \frac{X_0^2}{\sigma_0^2(\boldsymbol{\theta})}\right]$$

holds with equality if and only if $\sigma_0^2(\boldsymbol{\theta}) \overset{a.s.}{=} \sigma_0^2(\boldsymbol{\theta}_0)$. Rearranging the equation, leads to

$$-1/2E_{\boldsymbol{\theta}_0}\left[\log\sigma_0^2(\boldsymbol{\theta}_0) - \log\sigma_0^2(\boldsymbol{\theta}) + \frac{X_0^2}{\sigma_0^2(\boldsymbol{\theta}_0)} - \frac{X_0^2}{\sigma_0^2(\boldsymbol{\theta})}\right] \geq 0. \tag{5.3}$$

Recall that $X_0^2 = \sigma_0^2(\boldsymbol{\theta}_0)\epsilon_0^2$ and $E[\epsilon_0^2] = 1$. Therefore, the estimate in (5.3) is equivalent to

$$E_{\boldsymbol{\theta}_0}\left[\log\frac{\sigma_0^2(\boldsymbol{\theta}_0)}{\sigma_0^2(\boldsymbol{\theta})} - \frac{\sigma_0^2(\boldsymbol{\theta}_0)}{\sigma_0^2(\boldsymbol{\theta})}\right] \leq -1.$$

For $x > 0$ the estimate

$$\log(x) \leq x - 1$$

is strict, with equality if and only if $x = 1$. Set $x = \frac{\sigma_0^2(\boldsymbol{\theta}_0)}{\sigma_0^2(\boldsymbol{\theta})}$, then, it holds that

$$E_{\boldsymbol{\theta}_0}\left[\log\frac{\sigma_0^2(\boldsymbol{\theta}_0)}{\sigma_0^2(\boldsymbol{\theta})} - \frac{\sigma_0^2(\boldsymbol{\theta}_0)}{\sigma_0^2(\boldsymbol{\theta})}\right] = -1,$$

if and only if $\sigma_0^2(\boldsymbol{\theta}_0) = \sigma_0^2(\boldsymbol{\theta})$. Since condition (C4) ensures that $\sigma_0^2(\boldsymbol{\theta}_0) \overset{a.s.}{=} \sigma_0^2(\boldsymbol{\theta})$ if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_0$ is the maximizer of $L(\boldsymbol{\theta})$, $L(\boldsymbol{\theta}) < L(\boldsymbol{\theta}_0)$ follows for all $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. Following Ferguson [14, Theorem 16b] it is left to show that

$$P(\limsup_{n\to\infty}\sup_{\boldsymbol{\theta}\in C} n^{-1}L_n(\boldsymbol{\theta}) \leq \sup_{\boldsymbol{\theta}\in C} L(\boldsymbol{\theta})) = 1 \tag{5.4}$$

for any compact set $C \subset K$. The same argument is employed in Pfanzagl [36, Lemma 3.11] and Straumann [41, Theorem 5.3.1]. Recall that

$$E_{\boldsymbol{\theta}_0}\left[\left|\log\sigma_0^2(\boldsymbol{\theta}_0) + \frac{X_0^2}{\sigma_0^2(\boldsymbol{\theta}_0)}\right|\right] < \infty.$$

Hence, by the ergodic theorem it holds for every fixed $\boldsymbol{\theta}$ that

$$n^{-1}L_n(\boldsymbol{\theta}) \overset{a.s.}{\to} E_{\boldsymbol{\theta}_0}[l_0(\boldsymbol{\theta})].$$

Since (5.1) holds, 5.4 follows from Ferguson [14, Theorem 16b]. Consistency follows by compactness arguments analogue to Wald [43, Theorem 2]. □

In order to demonstrate $n^{-1}\|\tilde{L}_n(\boldsymbol{\theta}) - L_n(\boldsymbol{\theta})\|_K$, the following intermediate result is proven first.

***Lemma* 5.3.** Under the assumption of Theorem 3.1 it follows that

$$\|\log\sigma_t^2 - \log\tilde{\sigma}_t^2\|_K \overset{e.a.s.}{\to} 0$$

and

$$\|(\sigma_t^2)^{-1} - (\tilde{\sigma}_t^2)^{-1}\|_K \overset{e.a.s.}{\to} 0.$$

Proof. Suppress the dependence of σ_t^2 on $\boldsymbol{\theta}$ for easier notation. Since $h(x) = x^\delta$ is a continuous function for every $\delta > 0$, $x \in (0, \infty)$, and $y = h(x) > 0$ under assumption $c(\epsilon_0), g(\epsilon_0) \geq 0$ the inverse function $h^{-1}(y) = y^{1/\delta}$ is uniquely defined for every $y > 0$. Denote $\tilde{h}_t = h(\tilde{\sigma}_t^2)$. Then, applying Hörmann [23, Lemma 5] yields

$$P(|h_t - \tilde{h}_t| > \exp(-\alpha t)) \leq C_2 \exp(-\rho t) \to 0,$$

for $t \to \infty$ and some constants $C_2 > 0$ and $\rho > 0$. Denote $f_t := |h_t - \tilde{h}_t|$, then, a.s. $f_t \geq f_{t+1}$. In addition, f_t is continuous on K by assumption (C3) and $f_t \overset{e.a.s}{\to} 0$ for fixed $\boldsymbol{\theta} \in K$. Applying Rudin [39, Theorem 7.13], it may be concluded that $\|h_t - \tilde{h}_t\|_K \overset{e.a.s.}{\to} 0$. Recall that $h(x)$ is differentiable w.r.t. x and so is $h^{-1}(x)$. Thus, applying the mean value theorem, yields

$$\begin{aligned}
\|\sigma_t^2 - \tilde{\sigma}_t^2\|_K &= \|h^{-1}(h_t) - h^{-1}(\tilde{h}_t)\|_K \\
&\leq \left\| \frac{1}{\delta(\hat{\sigma}_t^2)^{\delta-1}} \right\|_K \|h_t - \tilde{h}_t\|_K \\
&\leq \frac{1}{\delta(\bar{h})^{\delta-1}} \|h_t - \tilde{h}_t\|_K \overset{e.a.s.}{\to} 0.
\end{aligned}$$

Here, $\hat{\sigma}_t^2 \geq \bar{h}$ is between σ_t^2 and $\tilde{\sigma}_t^2$. This shows that $\|\sigma_t^2 - \tilde{\sigma}_t^2\|_K \overset{e.a.s.}{\to} 0$ for $t \to \infty$. Applying the mean value theorem to the sequence $\log \sigma_t^2$, it follows that

$$\|\log \sigma_t^2 - \log \tilde{\sigma}_t^2\|_K \leq \left\| \frac{1}{\hat{\sigma_t}^2} \right\|_K \|\sigma_t^2 - \tilde{\sigma}_t^2\|_K \leq \frac{1}{\bar{h}} \|\sigma_t^2 - \tilde{\sigma}_t^2\|_K \overset{e.a.s.}{\to} 0$$

by assumption (C1). Proceeding essentially the same way, it follows that

$$\|(\sigma_t^2)^{-1} - (\tilde{\sigma}_t^2)^{-1}\|_K = \left\| \frac{\tilde{\sigma}_t^2 - \sigma_t^2}{\tilde{\sigma}_t^2 \sigma_t^2} \right\|_K \leq \left\| \frac{\tilde{\sigma}_t^2 - \sigma_t^2}{\bar{h}^2} \right\|_K \overset{e.a.s.}{\to} 0.$$

□

Proof of Theorem 3.1. First, recall that the augmented GARCH process is assumed to be strictly stationary and ergodic. An application of Lemma 5.2 shows that $\hat{\boldsymbol{\theta}}_n^* \overset{a.s.}{\to} \boldsymbol{\theta}_0$. It is left to show that

$$\frac{1}{n} \|L_n(\boldsymbol{\theta}) - \tilde{L}_n(\boldsymbol{\theta})\|_K \overset{a.s.}{\to} 0. \tag{5.5}$$

From Lemma 5.3 there exist constants $C_2, C_3 > 0$, such that

$$\|\log \sigma_t^2(\boldsymbol{\theta}) - \log \tilde{\sigma}_t^2(\boldsymbol{\theta})\|_K \leq C_2 \|\sigma_t^2(\boldsymbol{\theta}) - \tilde{\sigma}_t^2(\boldsymbol{\theta})\|_K$$

and

$$\left\|(\sigma_t^2(\boldsymbol{\theta}))^{-1} - (\tilde{\sigma}_t^2(\boldsymbol{\theta}))^{-1}\right\|_K \leq C_3 \|\sigma_t^2(\boldsymbol{\theta}) - \tilde{\sigma}_t^2(\boldsymbol{\theta})\|_K.$$

Hence, using the triangle inequality, yields

$$\begin{aligned}\|L_n(\boldsymbol{\theta}) - \tilde{L}_n(\boldsymbol{\theta})\|_K &= \left\|\sum_{t=1}^{n} \log \sigma_t^2(\boldsymbol{\theta}) + \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})} - \log \tilde{\sigma}_t^2(\boldsymbol{\theta}) - \frac{X_t^2}{\tilde{\sigma}_t^2(\boldsymbol{\theta})}\right\|_K \\ &\leq \sum_{t=1}^{n} \left\|\log \sigma_t^2(\boldsymbol{\theta}) - \log \tilde{\sigma}_t^2(\boldsymbol{\theta})\right\|_K + X_t^2 \left\|(\sigma_t^2(\boldsymbol{\theta}))^{-1} - (\tilde{\sigma}_t^2(\boldsymbol{\theta}))^{-1}\right\|_K \\ &\leq \sum_{t=1}^{\infty} (1 + C_4 X_t^2) C_2 \|\sigma_t^2(\boldsymbol{\theta}) - \tilde{\sigma}_t^2(\boldsymbol{\theta})\|_K,\end{aligned}$$

where $C_4 = C_3/C_2$. Since $\|\sigma_t^2(\boldsymbol{\theta}) - \tilde{\sigma}_t^2(\boldsymbol{\theta})\|_K \overset{e.a.s.}{\to} 0$, apply Lemma 5.1 to conclude that the series

$$\sum_{t=1}^{\infty} X_t^2 \|\sigma_t^2(\boldsymbol{\theta}) - \tilde{\sigma}_t^2(\boldsymbol{\theta})\|_K$$

converges a.s. if $E[\log^+ X_0^2] < \infty$. Since $(X_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic together with (C1) and (C2), it holds that $E[X_0^{2s}] < \infty$ for some $s > 0$, which in turn implies $E[\log^+ X_0^2] < \infty$ by Jensen's inequality. Hence, the conditions of Lemma 5.1 are fulfilled and equation (5.5) follows for $n \to \infty$ because $\|L_n(\boldsymbol{\theta}) - \tilde{L}_n(\boldsymbol{\theta})\|_K$ is a.s. bounded. This completes the proof of Theorem 3.1. □

5.2. Proof of Theorem 3.2

First, it is shown for the estimator $\hat{\boldsymbol{\theta}}_n^*$. Second, the asymptotic equivalence of $\hat{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n^*$ is proven. This is established by showing that $\sqrt{n}\|\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^*\| \overset{a.s.}{\to} 0$ for $n \to \infty$.

***Lemma* 5.4.** Assume the conditions of Theorem 3.2. Then, it follows that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^* - \boldsymbol{\theta}_0) \overset{d}{\to} N(\mathbf{0}, \boldsymbol{\Sigma}),$$

and $\boldsymbol{\Sigma}$ as in Theorem 3.2.

Proof. Since $\nabla_{\boldsymbol{\theta}} l_t(\boldsymbol{\theta})$ and $\nabla^2_{\boldsymbol{\theta}} l_t(\boldsymbol{\theta})$ are measurable functions of the strictly stationary and ergodic process $(X_t)_{t\in\mathbb{Z}}$, conclude that $n^{-1/2}\sum_{t=1}^n(\nabla_{\boldsymbol{\theta}} l_t)$ is a strictly stationary and ergodic zero-mean martingale difference. It is left to show that $\boldsymbol{\Sigma}$ is well defined. First, observe that

$$\begin{aligned}E_{\boldsymbol{\theta}_0}\left[\nabla^2_{\boldsymbol{\theta}} l_0(\boldsymbol{\theta}_0)\right] &= -1/2 E_{\boldsymbol{\theta}_0}\left[\left(1 - \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta}_0)}\right)\left(\frac{1}{\sigma_t^2(\boldsymbol{\theta}_0)} \nabla^2_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta}_0)\right)\right] \\ &\quad - 1/2 E_{\boldsymbol{\theta}_0}\left[\left(2\frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta}_0)} - 1\right)\left(\frac{1}{\sigma_t^4(\boldsymbol{\theta}_0)} \nabla_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta}_0) \nabla'_{\boldsymbol{\theta}} \sigma_t^2(\boldsymbol{\theta}_0)\right)\right].\end{aligned}$$

Recall, that $\frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta}_0)} = \epsilon_t^2$ and $E[\epsilon_t^2] = 1$. Together with the independence of ϵ_t from $\sigma_t^2(\boldsymbol{\theta})$ and $\nabla^2_{\boldsymbol{\theta}}\sigma_t^2(\boldsymbol{\theta}_0)$ and assumption (N3) it follows that

$$E_{\boldsymbol{\theta}_0}\left[-1/2\left(1 - \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta}_0)}\right)\left(\frac{1}{\sigma_t^2(\boldsymbol{\theta}_0)}\nabla^2_{\boldsymbol{\theta}}\sigma_t^2(\boldsymbol{\theta}_0)\right)\right] = 0.$$

It is left to show that the matrix

$$E_{\boldsymbol{\theta}_0}\left[\left(\frac{1}{\sigma_0^4(\boldsymbol{\theta}_0)}\nabla_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\right)\right]^{-1}$$

is positive definite. This follows directly by an application of Straumann and Mikosch [42, Lemma 7.2] together with (N4). Using $\epsilon_0 = X_0/\sigma_0(\boldsymbol{\theta}_0)$, $E[\epsilon_0^2] = 1$, the independence of ϵ_0 and $\nabla_{\boldsymbol{\theta}}\sigma_0^2$ and assumption (N3) again, it holds that

$$\begin{aligned}\boldsymbol{\Sigma} &= E_{\boldsymbol{\theta}_0}\left[\frac{1}{\sigma_0^4(\boldsymbol{\theta}_0)}\nabla_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\right]^{-1} E_{\boldsymbol{\theta}_0}\left[\left(1-\epsilon_0^2\right)^2\frac{1}{\sigma_0^4(\boldsymbol{\theta}_0)}\nabla_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\right] \\ &\quad\times E_{\boldsymbol{\theta}_0}\left[\frac{1}{\sigma_0^4(\boldsymbol{\theta}_0)}\nabla_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\right]^{-1} \\ &= E_{\boldsymbol{\theta}_0}[(\epsilon_0^4 - 1)]E_{\boldsymbol{\theta}_0}\left[\frac{1}{\sigma_0^4(\boldsymbol{\theta}_0)}\nabla_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_0^2(\boldsymbol{\theta}_0)\right]^{-1}\end{aligned}$$

□

The next lemma shows that $\hat{\boldsymbol{\theta}}_n^*$ is asymptotically equivalent to $\hat{\boldsymbol{\theta}}_n$ meaning that both have the same limiting distribution. This follows if $\sqrt{n}\|\hat{\boldsymbol{\theta}}_n^* - \hat{\boldsymbol{\theta}}_n\| \overset{a.s.}{\to} 0$ holds as $n \to \infty$. First, one has to show that

$$\left\|n^{-1}\nabla^2_{\boldsymbol{\theta}}\tilde{L}_n - \boldsymbol{B}_0\right\|_K \overset{a.s.}{\to} 0 \tag{5.6}$$

as $n \to \infty$. Since

$$\|n^{-1}\nabla^2_{\boldsymbol{\theta}}\tilde{L}_n - \boldsymbol{B}_0\|_K \le \|n^{-1}\nabla^2_{\boldsymbol{\theta}}\tilde{L}_n - n^{-1}\nabla^2_{\boldsymbol{\theta}}L_n\|_K + \|n^{-1}\nabla^2_{\boldsymbol{\theta}}L_n - \boldsymbol{B}_0\|_K$$

it suffices to prove that

$$\|n^{-1}\nabla^2_{\boldsymbol{\theta}}\tilde{L}_n - n^{-1}\nabla^2_{\boldsymbol{\theta}}L_n\|_K \overset{a.s.}{\to} 0$$

as $n \to \infty$ and (5.6) follows. Next, observe that

$$\begin{aligned}\sqrt{n}(\hat{\boldsymbol{\theta}}_n^* - \hat{\boldsymbol{\theta}}_n) &= \sqrt{n}(\hat{\boldsymbol{\theta}}_n^* - \boldsymbol{\theta}_0) - \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \\ &= -(\boldsymbol{B}_0^{-1} + o_P(1))\left(\frac{1}{\sqrt{n}}\sum_{t=1}^n \nabla_{\boldsymbol{\theta}} l_t - \frac{1}{\sqrt{n}}\sum_{t=1}^n \nabla_{\boldsymbol{\theta}}\tilde{l}_t\right).\end{aligned}$$

Hence, $\hat{\boldsymbol{\theta}}_n^*$ and $\hat{\boldsymbol{\theta}}_n$ are asymptotically equivalent if

$$\frac{1}{\sqrt{n}} \left\| \sum_{t=1}^{n} \nabla_{\boldsymbol{\theta}} l_t - \nabla_{\boldsymbol{\theta}} \tilde{l}_t \right\|_K \overset{a.s.}{\to} 0 \tag{5.7}$$

and

$$\frac{1}{n} \left\| \sum_{t=1}^{n} \nabla^2_{\boldsymbol{\theta}} l_t - \nabla^2_{\boldsymbol{\theta}} \tilde{l}_t \right\|_K \overset{a.s.}{\to} 0. \tag{5.8}$$

The next lemma entails all the ingredients being necessary to prove (5.7) and (5.8).

***Lemma* 5.5.** Under the condition of Theorem 3.2 it holds that

$$\|\sigma_t^{-2} \nabla_{\boldsymbol{\theta}} \sigma_t^2 - \tilde{\sigma}_t^{-2} \nabla_{\boldsymbol{\theta}} \tilde{\sigma}_t^2\|_K \overset{e.a.s.}{\to} 0, \tag{5.9}$$

$$\|\sigma_t^{-2} \nabla^2_{\boldsymbol{\theta}} \sigma_t^2 - \tilde{\sigma}_t^{-2} \nabla^2_{\boldsymbol{\theta}} \tilde{\sigma}_t^2\|_K \overset{e.a.s.}{\to} 0, \tag{5.10}$$

and

$$\|\sigma_t^{-4} \nabla_{\boldsymbol{\theta}} \sigma_t^2 \nabla'_{\boldsymbol{\theta}} \sigma_t^2 - \tilde{\sigma}_t^{-4} \nabla_{\boldsymbol{\theta}} \tilde{\sigma}_t^2 \nabla'_{\boldsymbol{\theta}} \tilde{\sigma}_t^2\|_K \overset{e.a.s.}{\to} 0 \tag{5.11}$$

as $n \to \infty$.

Proof. Write σ_t^2 instead of $\sigma_t^2(\boldsymbol{\theta})$ for easier notation. To prove equation (5.9) note that $\sigma_t^2(\boldsymbol{\theta}) \geq \bar{h} > 0$ for all $\boldsymbol{\theta} \in K$ by condition (C1). Since $\sigma_t^2(\boldsymbol{\theta})$ is three times differentiable w.r.t. $\boldsymbol{\theta}$, the mean value theorem for vector valued functions applies, cf. Rudin [39, Theorem 9.19]. Therefore, the additional assumption that K is convex is needed.

$$\begin{aligned}
\|\sigma_t^{-2} \nabla_{\boldsymbol{\theta}} \sigma_t^2 - \tilde{\sigma}_t^{-2} \nabla_{\boldsymbol{\theta}} \tilde{\sigma}_t^2\|_K &\leq \left\| \tilde{\sigma}_t^{-2} \frac{\partial}{\partial \boldsymbol{\theta}} \left(\sigma_t^2 - \tilde{\sigma}_t^2 \right) \right\|_K + \left\| \nabla_{\boldsymbol{\theta}} \sigma_t^2 \left(\frac{1}{\sigma_t^2} - \frac{1}{\tilde{\sigma}_t^2} \right) \right\|_K \\
&\leq \underbrace{\frac{1}{\bar{h}} \left\| \nabla^2_{\boldsymbol{\theta}} \sigma_t^2 \right\|_K \left\| (\sigma_t^2 - \tilde{\sigma}_t^2) \right\|_K}_{:=A_t} + \underbrace{\left\| \nabla_{\boldsymbol{\theta}} \sigma_t^2 \left(\frac{1}{\sigma_t^2} - \frac{1}{\tilde{\sigma}_t^2} \right) \right\|_K}_{:= B_t}.
\end{aligned}$$

Since $\|\sigma_t^2 - \tilde{\sigma}_t^2\|_K \overset{e.a.s.}{\to} 0$ by virtue of Lemma 5.3 and, additionally, $E_{\boldsymbol{\theta}_0}[\log^+ \left\| \nabla^2_{\boldsymbol{\theta}} \sigma_t^2 \right\|_K] < \infty$ by assumption (N3i)-(N3iii) for every $\boldsymbol{\theta} \in K$ it follows that $A_t \overset{e.a.s.}{\to} 0$ by virtue of Lemma 5.1. Similarly, $B_t \overset{e.a.s.}{\to} 0$ because by Lemma 5.3 $\|(\sigma_t^2)^{-1} - (\tilde{\sigma}_t^2)^{-1}\| \overset{e.a.s.}{\to} 0$ and assumption (N3i)-(N3iii). Hence equation (5.9) follows. Applying the mean value theorem again, yields

$$\begin{aligned}
\|\sigma_t^{-2} \nabla^2_{\boldsymbol{\theta}} \sigma_t^2 - \tilde{\sigma}_t^{-2} \nabla^2_{\boldsymbol{\theta}} \tilde{\sigma}_t^2(\boldsymbol{\theta})\|_K &\leq \|\tilde{\sigma}_t^{-2}\|_K \|\nabla^2_{\boldsymbol{\theta}} (\sigma_t^2 - \tilde{\sigma}_t)\|_K \\
&\quad + \|\nabla^2_{\boldsymbol{\theta}} \sigma_t^2\|_K \left\| \frac{1}{\sigma_t^2} - \frac{1}{\tilde{\sigma}_t^2} \right\|_K \\
&\leq \|\nabla^3_{\boldsymbol{\theta}} \sigma_t^2\|_K \left\| (\sigma_t^2 - \tilde{\sigma}_t^2) \right\|_K + \|\nabla^2_{\boldsymbol{\theta}} \sigma_t^2\|_K \left\| \frac{1}{\sigma_t^2} - \frac{1}{\tilde{\sigma}_t^2} \right\|_K.
\end{aligned}$$

An application of Lemma 5.3 together with the assumption (N3i)-(N3iv) and Lemma 5.1 finalizes the proof of equation (5.10). Finally, the validity of equation (5.11) is checked proceeding essentially the same way.

$$\begin{aligned}\|\sigma_t^{-4}\nabla_{\boldsymbol{\theta}}\sigma_t^2\nabla'_{\boldsymbol{\theta}}\sigma_t^2 - \tilde{\sigma}_t^{-4}\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2\nabla'_{\boldsymbol{\theta}}\tilde{\sigma}_t^2\|_K &\le \|\tilde{\sigma}_t^{-4}\|_K\|\nabla_{\boldsymbol{\theta}}\sigma_t^2\nabla'_{\boldsymbol{\theta}}\sigma_t^2 - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2\nabla'_{\boldsymbol{\theta}}\tilde{\sigma}_t^2\|_K \\ &\quad + \|\nabla_{\boldsymbol{\theta}}\sigma_t^2\nabla'_{\boldsymbol{\theta}}\sigma_t^2\|_K\|\sigma_t^{-4} - \tilde{\sigma}_t^{-4}\|_K \\ &\le \bar{h}^{-2}\|(\nabla_{\boldsymbol{\theta}}\sigma_t^2 - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2)(\nabla_{\boldsymbol{\theta}}\sigma_t^2 + \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2)'\|_K \\ &\quad + \|\nabla_{\boldsymbol{\theta}}\sigma_t^2\nabla'_{\boldsymbol{\theta}}\sigma_t^2\|_K\|\sigma_t^{-4} - \tilde{\sigma}_t^{-4}\|_K.\end{aligned}$$

Write $c_{i,j} = \frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}_i}\frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}_j}$. Since the sup-norm is sub-multiplicative it holds

$$E_{\boldsymbol{\theta}_0}[\log^+\|c_{i,j}\|_K] \le E_{\boldsymbol{\theta}_0}\left[\log^+\left\|\frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}_i}\right\|_K\right] + E_{\boldsymbol{\theta}_0}\left[\log^+\left\|\frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}_j}\right\|_K\right] < \infty$$

for all $i, j \in \{1, ..., d\}$. Thus, it holds that

$$E[\log^+\|\nabla_{\boldsymbol{\theta}}\sigma_t^2\nabla'_{\boldsymbol{\theta}}\sigma_t^2\|_K] < \infty. \tag{5.12}$$

In addition, applying the mean value theorem to $\|\sigma_t^{-4} - \tilde{\sigma}_t^{-4}\|_K$, it one can deduce that

$$\begin{aligned}\|\sigma_t^{-4} - \tilde{\sigma}_t^{-4}\|_K &\le 2\|\sigma_t^{-2}\|\|\sigma_t^{-2} - \tilde{\sigma}_t^{-2}\|_K \\ &\le 2\bar{h}^{-1}\|\sigma_t^{-2} - \tilde{\sigma}_t^{-2}\|_K \overset{e.a.s}{\rightarrow} 0.\end{aligned}$$

Therefore, an application of Lemma 5.1 shows that

$$\|\nabla_{\boldsymbol{\theta}}\sigma_t^2\nabla'_{\boldsymbol{\theta}}\sigma_t^2\|_K\|\sigma_t^{-4} - \tilde{\sigma}_t^{-4}\|_K \overset{e.a.s.}{\rightarrow} 0.$$

Finally, observe that

$$\|(\nabla_{\boldsymbol{\theta}}\sigma_t^2 - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2)(\nabla_{\boldsymbol{\theta}}\sigma_t^2 + \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2)'\|_K \le 2\|\nabla_{\boldsymbol{\theta}}\sigma_t^2\|_K\|\nabla_{\boldsymbol{\theta}}\sigma_t^2 - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2\|_K \overset{e.a.s.}{\rightarrow} 0 \tag{5.13}$$

by condition (N3ii) and Lemma 5.1. Putting the estimates (5.12) and (5.13) together equation (5.11) follows. □

***Lemma* 5.6.** Assume the conditions of Theorem 3.2. Then, (5.7) and (5.8) hold.

Proof. By virtue of (3.11) and the results of Lemma 5.5 deduce that

$$\begin{aligned}\left\|\sum_{t=1}^n \nabla_{\boldsymbol{\theta}} l_t - \nabla_{\boldsymbol{\theta}}\tilde{l}_t\right\|_K &\le \sum_{t=1}^n \|\nabla_{\boldsymbol{\theta}} l_t - \nabla_{\boldsymbol{\theta}}\tilde{l}_t\|_K \\ &\le \sum_{t=1}^n |1 + \bar{h}^{-1}X_t^2|\|\sigma_t^{-2}\nabla_{\boldsymbol{\theta}}\sigma_t^2 - \tilde{\sigma}_t^{-2}\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t^2\|_K \\ &< \infty, \text{ for } n \to \infty\end{aligned}$$

by accounting for $E[\log^+(1+X_0^2)]<\infty$ and the results of Lemma 5.5. An application of Lemma 5.1 proves (5.7). For the second hypothesis of Lemma 5.13, observe that

$$\begin{aligned}\left\|\sum_{t=1}^{n}\nabla^2_{\boldsymbol\theta} l_t-\nabla^2_{\boldsymbol\theta}\tilde l_t\right\|_K &\le \sum_{t=1}^{n}\left\|\left(1-\frac{X_t^2}{\sigma_t^2}\right)\left(\frac{1}{\sigma_t^2}\nabla^2_{\boldsymbol\theta}\sigma_t^2\right)\right.\\ &\quad+\left(2\frac{X_t^2}{\sigma_t^2}-1\right)\left(\frac{1}{\sigma_t^4}\nabla_{\boldsymbol\theta}\sigma_t^2\nabla'_{\boldsymbol\theta}\sigma_t^2\right)\\ &\quad-\left(1-\frac{X_t^2}{\tilde\sigma_t^2}\right)\left(\frac{1}{\tilde\sigma_t^2}\nabla^2_{\boldsymbol\theta}\tilde\sigma_t^2\right)\\ &\quad\left.-\left(2\frac{X_t^2}{\tilde\sigma_t^2}-1\right)\left(\frac{1}{\tilde\sigma_t^4}\nabla_{\boldsymbol\theta}\tilde\sigma_t^2\nabla'_{\boldsymbol\theta}\tilde\sigma_t^2\right)\right\|_K\\ &\le\sum_{t=1}^{n}|(1+\bar h^{-2}X_t^2)|\left(\left\|\frac{1}{\sigma_t^2}\nabla_{\boldsymbol\theta}\sigma_t^2-\frac{1}{\tilde\sigma_t^2}\nabla_{\boldsymbol\theta}\tilde\sigma_t^2\right\|_K\right.\\ &\quad\left.+\left\|\frac{1}{\sigma_t^4}\nabla_{\boldsymbol\theta}\sigma_t^2\nabla'_{\boldsymbol\theta}\sigma_t^2-\frac{1}{\tilde\sigma_t^4}\nabla_{\boldsymbol\theta}\tilde\sigma_t^2\nabla'_{\boldsymbol\theta}\tilde\sigma_t^2\right\|_K\right).\end{aligned}$$

Since $E[\log^+\|1+X_0^2\|_K]<\infty$,

$$\left\|\sum_{t=1}^{n}\nabla^2_{\boldsymbol\theta} l_t-\nabla^2_{\boldsymbol\theta}\tilde l_t\right\|_K<\infty$$

follows by an application of Lemma 5.6 and Lemma 5.1. This shows (5.8) and Lemma 5.6 is proven. □

The proof of Theorem 3.2 follows by subsequent application of the Lemmas 5.4-5.6.

5.3. Proof of Theorem 4.4

Before proving Theorem 4.4, the following intermediate results are shown first.

***Lemma* 5.7.** Assume the condition of Theorem 4.4. Then, it follows that

$$E_{\boldsymbol\theta_0}\left[\log^+\|\nabla_{\boldsymbol\theta}\sigma_t(\boldsymbol\theta)\|_{K_d}\right]<\infty, \tag{5.14}$$

$$E_{\boldsymbol\theta_0}\left[\log^+\left\|\nabla^2_{\boldsymbol\theta}\sigma_t(\boldsymbol\theta)\right\|_{K_d}\right]<\infty, \tag{5.15}$$

$$E_{\boldsymbol\theta_0}\left[\log^+\left\|\nabla^3_{\boldsymbol\theta}\sigma_t(\boldsymbol\theta)\right\|_{K_d}\right]<\infty. \tag{5.16}$$

Proof. Starting with $\nabla\sigma_t(\boldsymbol{\theta})$ yields

$$\nabla_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta}) = \begin{pmatrix} \frac{1}{1-\beta} \\ \sum_{i=1}^{\infty}\beta^{i-1}|X_{t-i}| \\ \sum_{i=1}^{\infty}\beta^{i-1}X_{t-i}^{-} \\ \frac{\omega}{(1-\beta)^2} + \alpha^{+}\sum_{i=1}^{\infty}(i-1)\beta^{i-2}|X_{t-i}| + \alpha^{-}\sum_{i=1}^{\infty}(i-1)\beta^{i-2}X_{t-i}^{-} \end{pmatrix}. \tag{5.17}$$

Since $\beta \le 1-d$ it holds that $-\log(1-\beta) < \infty$ for all $\beta \in K_d$. Observe that a.s.

$$\frac{\partial\sigma_t}{\partial\alpha^{+}} = \sum_{i=1}^{\infty}\beta^{i-1}|X_{t-i}| \le \sigma_t/\alpha^{+}.$$

Since $(X_t)_{t\in\mathbb{Z}}$ is strictly stationary and $\beta < 1$ it follows from Hamadeh and Zakoïan [17, Proposition A1] and Pan et al. [35, Theorem 6] that there exists a $s > 0$ such that $E[\sigma_t^s] < \infty$, which yields for every $\boldsymbol{\theta} \in K_d$

$$E_{\boldsymbol{\theta}_0}\left[\log^{+}\left\|\frac{\partial\sigma_t}{\partial\alpha^{+}}\right\|_{K_d}\right] \le s^{-1}\log^{+}E_{\boldsymbol{\theta}_0}\left[\left\|\frac{\sigma_t^s}{\alpha_0^{+s}}\right\|_{K_d}\right] < \infty.$$

A similar argument leads to $E_{\boldsymbol{\theta}_0}\left[\log^{+}\left\|\frac{\partial\sigma_t}{\partial\alpha^{-}}\right\|_{K_d}\right] < \infty$. For every $\boldsymbol{\theta} \in K_d$, the fourth component can be estimated by

$$\begin{aligned} &E_{\boldsymbol{\theta}_0}\left[\log^{+}\left\|\frac{\partial\sigma_t}{\partial\beta}\right\|_{K_d}\right] \\ &= E_{\boldsymbol{\theta}_0}\left[\log^{+}\left\|\frac{\omega}{(1-\beta)^2} + \sum_{i=1}^{\infty}(i-1)\beta^{i-2}(\alpha^{+}|X_{t-i}| + \alpha^{-}X_{t-i}^{-})\right\|_{K_d}\right] \\ &\le C_5 + 1/s\log^{+}E_{\boldsymbol{\theta}_0}\left[\left\|\sum_{i=1}^{\infty}(i-1)\beta^{i-2}(\alpha^{+}|X_{t-i}| + \alpha^{-}X_{t-i}^{-})\right\|_{K_d}^{s}\right], \end{aligned}$$

where $C_5 = \log 2 + \log^{+}\omega - \log(1-\beta) < \infty$. Using $(a+b)^s \le a^s + b^s$ for $0 < s < 1$ and $a, b \ge 0$, observe that $|X_t|^s = \sigma_t^s|\epsilon_t|^s$ and, thus, $E[|X_t|^s] = E[\sigma_t^s]E[|\epsilon_t|^s] < \infty$ by Pan et al. [35, Theorem 6]. Hence, it can be deduced that

$$\begin{aligned} &E_{\boldsymbol{\theta}_0}\left[\left\|\sum_{i=1}^{\infty}(i-1)\beta^{i-2}(\alpha^{+}|X_{t-i}| + \alpha^{-}X_{t-i}^{-})\right\|_{K_d}^{s}\right] \\ &\le \sum_{i=1}^{\infty}(i-1)^s\beta^{s(i-2)}E_{\boldsymbol{\theta}_0}[(\alpha^{+} + \alpha^{-})|X_{t-i}|)^s] \\ &< \infty \end{aligned}$$

and (5.14) follows by an application of Lemma 5.1. To show (5.15) one proceeds essentially the same way. Therefore, it suffices to discuss the following vector

$$\frac{\partial}{\partial\beta}(\nabla_{\boldsymbol{\theta}}\sigma_t) = \begin{pmatrix} \frac{1}{(1-\beta)^2} \\ \sum_{i=1}^{\infty}(i-1)\beta^{i-2}|X_{t-i}| \\ \sum_{i=1}^{\infty}(i-1)\beta^{i-2}X_{t-i}^- \\ \frac{2\omega}{(1-\beta)^3} + \sum_{i=3}^{\infty}(i-1)(i-2)\beta^{i-3}(\alpha^+|X_{t-i}| + \alpha^- X_{t-i}^-) \end{pmatrix}$$

Following the steps to prove (5.14), it can be shown that

$$E_{\boldsymbol{\theta}_0}\left[\log^+\left\|\frac{\partial}{\partial\beta}(\nabla_{\boldsymbol{\theta}}\sigma_t)\right\|_{K_d}\right] < \infty$$

which in turn implies (5.15). For the third derivative it is enough to regard $\frac{\partial}{\partial\beta}(\nabla^2_{\boldsymbol{\theta}}\sigma_t)$. Since similar arguments are used as before, details are omitted. □

It is left to verify the remaining moment conditions (N3iv)-(N3vii). The assertions of the next lemma are shown in the proof of Hamadeh and Zakoïan [17, Theorem 2.2]. A proof is therefore omitted.

***Lemma* 5.8.** Assume the condition of Theorem 4.4. Then, it follows that

$$E_{\boldsymbol{\theta}_0}\left[\left\|\frac{\nabla_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta}_0)}\right\|\right] < \infty, \tag{5.18}$$

$$E_{\boldsymbol{\theta}_0}\left[\left\|\frac{\nabla^2_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta}_0)}\right\|\right] < \infty, \tag{5.19}$$

$$E_{\boldsymbol{\theta}_0}\left[\left\|\frac{\nabla_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t^2(\boldsymbol{\theta}_0)}\right\|\right] < \infty, \tag{5.20}$$

$$E_{\boldsymbol{\theta}_0}\left[\left\|\nabla^2_{\boldsymbol{\theta}}l_t(\boldsymbol{\theta})\right\|_{\tilde{K}}\right] < \infty, \tag{5.21}$$

for some neighborhood $\tilde{K} \subset K_d$ containing $\boldsymbol{\theta}_0$.

Proof of Theorem 4.4. First, it is shown that $\sqrt{n}(\hat{\boldsymbol{\theta}}_n^* - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma})$. Using the equations (3.12), (3.13) and (3.14) with $\delta = 1/2$ it holds that

$$\nabla_{\boldsymbol{\theta}}l_t(\boldsymbol{\theta}) = -1/2\left(1 - \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})}\right)\sigma_t^{-1}\nabla_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta}).$$

$$\nabla^2_{\boldsymbol{\theta}}l_t(\boldsymbol{\theta}) = -1/2\left(\left(1 - \frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})}\right)\left(\frac{1}{\sigma_t}\nabla^2_{\boldsymbol{\theta}}\sigma_t\right) + \left(3\frac{X_t^2}{\sigma_t^2(\boldsymbol{\theta})} - 1\right)\frac{1}{\sigma_t^2(\boldsymbol{\theta})}\nabla_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta})\nabla'_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta})\right).$$

Using the results of Lemma 5.8, it follows that $E_{\boldsymbol{\theta}_0}[\|\nabla_{\boldsymbol{\theta}}l_t(\boldsymbol{\theta}_0)\|] < \infty$ and $E_{\boldsymbol{\theta}_0}[\|\nabla^2_{\boldsymbol{\theta}}l_t(\boldsymbol{\theta}_0)\|] < \infty$. Since ϵ_0 and σ_0 are independent, it may be shown that

$$\boldsymbol{\Sigma} = E[(\epsilon_0^4 - 1)]E_{\boldsymbol{\theta}_0}\left[4\sigma_0^{-2}(\boldsymbol{\theta}_0)\nabla_{\boldsymbol{\theta}}\sigma_0(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_0(\boldsymbol{\theta}_0)\right]^{-1}.$$

From Hamadeh and Zakoïan [17, Theorem 2.2] it can be deduced that the matrix $\mathbf{\Sigma}$ is well defined because $E_{\boldsymbol{\theta}_0}\left[4\nabla_{\boldsymbol{\theta}}\sigma_0(\boldsymbol{\theta}_0)\nabla'_{\boldsymbol{\theta}}\sigma_0(\boldsymbol{\theta}_0)\right]$ is invertible. It is left show that $\hat{\boldsymbol{\theta}}_n$ is asymptotically equivalent to $\hat{\boldsymbol{\theta}}_n^*$ as in Lemma 5.6. Since $\sigma_t^2(\boldsymbol{\theta})$ is not continuously differentiable w.r.t. $\boldsymbol{\theta}$ Lemma 5.5 is not directly applicable. Nevertheless, the following equations can be demonstrated analogously to Lemma 5.5.

$$\left\|\frac{1}{\sigma_t}\nabla_{\boldsymbol{\theta}}\sigma_t - \frac{1}{\tilde{\sigma}_t}\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t\right\|_{K_d} \overset{e.a.s.}{\to} 0, \tag{5.22}$$

$$\left\|\frac{1}{\sigma_t}\nabla^2_{\boldsymbol{\theta}}\sigma_t - \frac{1}{\tilde{\sigma}_t}\nabla^2_{\boldsymbol{\theta}}\tilde{\sigma}_t\right\|_{K_d} \overset{e.a.s.}{\to} 0, \tag{5.23}$$

and

$$\left\|\frac{1}{\sigma_t^2}\nabla_{\boldsymbol{\theta}}\sigma_t\nabla'_{\boldsymbol{\theta}}\sigma_t - \frac{1}{\tilde{\sigma}_t^2}\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t\nabla'_{\boldsymbol{\theta}}\tilde{\sigma}_t\right\|_{K_d} \overset{e.a.s.}{\to} 0. \tag{5.24}$$

First, conclude that

$$\left\|\frac{1}{\sigma_t}\nabla_{\boldsymbol{\theta}}\sigma_t - \frac{1}{\tilde{\sigma}_t}\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t\right\|_{K_d} \le \underbrace{\left\|\frac{1}{\sigma_t} - \frac{1}{\tilde{\sigma}_t}\right\|_{K_d}\|\nabla_{\boldsymbol{\theta}}\sigma_t\|_{K_d}}_{:=A_t} + \underbrace{\left\|\frac{1}{\tilde{\sigma}_t}\right\|_{K_d}\|\nabla_{\boldsymbol{\theta}}\sigma_t - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t\|_{K_d}}_{:=B_t}.$$

Since $\|\sigma_t - \tilde{\sigma}_t\|_{K_d} \overset{e.a.s.}{\to} 0$ and $E_{\boldsymbol{\theta}_0}[\log^+\|\nabla_{\boldsymbol{\theta}}\sigma_t\|_{K_d}] < \infty$ by virtue Lemma 5.7, an application of Lemma 5.1 yields $A_t \overset{e.a.s.}{\to} 0$. Turning to B_t, observe that

$$\left\|\frac{\partial\sigma_t}{\partial\alpha^+} - \frac{\partial\tilde{\sigma}_t}{\partial\alpha^+}\right\|_{K_d} \le \left\|\frac{1}{\alpha^+}(\sigma_t - \tilde{\sigma}_t)\right\|_{K_d} \overset{e.a.s.}{\to} 0.$$

The same conclusion holds for $\frac{\partial\sigma_t}{\partial\alpha^-}$. For $\frac{\partial\sigma_t}{\partial\beta}$ it holds that

$$\begin{aligned}\left\|\frac{\partial\sigma_t}{\partial\beta} - \frac{\partial\tilde{\sigma}_t}{\partial\beta}\right\|_{K_d} &\le \sum_{i=t+1}^{\infty}(i-1)(1-d)^{i-2}\left\|\alpha^+|X_{t-1}| + \alpha^- X_{t-1}^-\right\|_{K_d}\\ &\le (1-d)^t\sum_{j=1}^{\infty}(j+t-1)(1-d)^{j-2}\left\|\alpha^+|X_{t-1}| + \alpha^- X_{t-1}^-\right\|_{K_d}\\ &\le (1-d)^t C_6 \overset{e.a.s.}{\to} 0,\end{aligned}$$

by virtue of Lemma 5.7 and Lemma 5.1. Hence,

$$\|\nabla_{\boldsymbol{\theta}}\sigma_t - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t\|_{K_d} \overset{e.a.s.}{\to} 0$$

and $B_t \overset{e.a.s.}{\to} 0$ because under the stationarity condition $\sigma_t \geq \omega > 0$ for all $\boldsymbol{\theta} \in K_d$. This proves (5.22). Note that (5.23) is shown if

$$\left\|\nabla^2_{\boldsymbol{\theta}}\sigma_t - \nabla^2_{\boldsymbol{\theta}}\tilde{\sigma}_t\right\|_{K_d} \overset{e.a.s.}{\to} 0.$$

Since this follows from the same considerations as before, details are omitted. It is left to show (5.24). Observe that

$$\begin{aligned}\left\|\frac{1}{\sigma_t^2}\nabla_{\boldsymbol{\theta}}\sigma_t\nabla^{'}_{\boldsymbol{\theta}}\sigma_t - \frac{1}{\tilde{\sigma}_t^2}\nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t\nabla^{'}_{\boldsymbol{\theta}}\tilde{\sigma}_t\right\|_{K_d} &\leq \omega^{-2}\|\nabla_{\boldsymbol{\theta}}\sigma_t - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t\|_{K_d}\|\nabla^{'}_{\boldsymbol{\theta}}\sigma_t\|_{K_d} \\ &\quad + \|\nabla_{\boldsymbol{\theta}}\sigma_t\nabla^{'}_{\boldsymbol{\theta}}\sigma_t\|_{K_d}\|\sigma_t^{-2} - \tilde{\sigma}_t^{-2}\|_{K_d}.\end{aligned}$$

Recall that $E[\log^+\|\nabla_{\boldsymbol{\theta}}\sigma_t\|_{K_d}] < \infty$, $E[\log^+\|\nabla_{\boldsymbol{\theta}}\sigma_t\nabla^{'}_{\boldsymbol{\theta}}\sigma_t\|_{K_d}] < \infty$ and, additionally, $\|\sigma_t^{-2} - \tilde{\sigma}_t^{-2}\|_{K_d} \overset{e.a.s}{\to} 0$. This implies

$$\|\nabla_{\boldsymbol{\theta}}\sigma_t - \nabla_{\boldsymbol{\theta}}\tilde{\sigma}_t\|_{K_d} \overset{e.a.s}{\to} 0.$$

Now, (5.24) follows by the same arguments as in Lemma 5.6 by an application of Lemma 5.1. Thus, it may be deduced that

$$n^{-1/2}\sum_{t=1}^{n}\|\nabla_{\boldsymbol{\theta}}l_t - \nabla_{\boldsymbol{\theta}}\tilde{l}_t\|_{K_d} \overset{a.s.}{\to} 0$$

and

$$n^{-1}\sum_{t=1}^{n}\|\nabla^2_{\boldsymbol{\theta}}l_t - \nabla^2_{\boldsymbol{\theta}}\tilde{l}_t\|_{K_d} \overset{a.s.}{\to} 0$$

hold as $n \to \infty$. This in turn implies the asymptotic equivalence of $\hat{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}^*_n$. Hence, the assertion of Theorem 4.4 follows. □

References

[1] Aue, A., Berkes, I., and Horváth, L. (2006). Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli*, 12(4):583–608.

[2] Berkes, I. and Horváth, L. (2004). The efficiency of the estimators of the parameters in GARCH processes. *Ann. Statist.*, 32(2):633–655.

[3] Berkes, I., Horváth, L., and Kokoszka, P. (2003). GARCH processes: structure and estimation. *Bernoulli*, 9(2):201–227.

[4] Billingsley, P. (1968). *Convergence of probability measures.* New York: John Wiley & Sons.

[5] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 31(3):307–327.

[6] Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *J. Econometrics*, 52:5–59.

[7] Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *J. Econometrics*, 52(1-2):115–127.
[8] Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory*, 18(1):17–39.
[9] Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *J. Emp. Finance*, 1:83–106.
[10] Duan, J.-C. (1997). Augmented GARCH(p, q) process and its diffusion limit. *J. Econometrics*, 79(1):97–127.
[11] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007.
[12] Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *J. Finance*, 48:1749–1778.
[13] Escanciano, J. C. (2009). Quasi-maximum likelihood estimation of semi-strong GARCH models. *Econometric Theory*, 25(2):561–570.
[14] Ferguson, T. S. (1996). *A course in large sample theory.* Texts in Statistical Science Series. London: Chapman & Hall.
[15] Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605–637.
[16] Francq, C. and Zakoïan, J.-M. (2006). Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory*, 22(5):815–834.
[17] Hamadeh, T. and Zakoïan, J.-M. (2011). Asymptotic properties of LS and QML estimators for a class of nonlinear GARCH processes. *J. Statist. Plann. Inference*, 141(1):488–507.
[18] Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH$(1, 1)$? *J. Appl. Econometrics*, 20(7):873–889.
[19] Hardy, G. H., Littlewood, J. E., and Pólya, G. (1952). *Inequalities.* Cambridge: Cambridge University Press, 2nd edition.
[20] Hentschel, L. (1995). All in the family nesting symmetric and asymmetric GARCH models. *J. Financ. Econ.*, 39:71–104.
[21] Herrmann, K. (2011). *Maximum Entropy Models for Time-Varying Moments Applied to Daily Financial Returns.* Berichte aus der Statistik. Shaker, Aachen.
[22] Heyde, C. C. (1974). On the central limit theorem for stationary processes. *Z. Wahrsch. Verw. Gebiete*, 30:315–320.
[23] Hörmann, S. (2008). Augmented GARCH sequences: dependence structure and asymptotics. *Bernoulli*, 14(2):543–561.
[24] Hwang, S. Y. and Basawa, I. V. (2004). Stationarity and moment structure for Box-Cox transformed threshold GARCH$(1, 1)$ processes. *Statist. Probab. Lett.*, 68(3):209–220.
[25] Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models. *Econometric Theory*, 14(1):70–86.
[26] Jensen, S. T. and Rahbek, A. (2004). Asymptotic inference for nonstationary GARCH. *Econometric Theory*, 20(6):1203–1226.
[27] Jondeau, E. and Rockinger, M. (2003). Conditional volatility, skewness, and kurtosis:

existence, persistence, and comovements. *J. Econom. Dynam. Control*, 27(10):1699–1737.

[28] Lee, O. and Shin, D. W. (2005). On stationarity and β-mixing property of certain nonlinear GARCH(p, q) models. *Statist. Probab. Lett.*, 73(1):25–35.

[29] Lee, S.-W. and Hansen, B. E. (1994). Asymptotic theory for the GARCH$(1, 1)$ quasi-maximum likelihood estimator. *Econometric Theory*, 10(1):29–52.

[30] Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH$(1, 1)$ and covariance stationary GARCH$(1, 1)$ models. *Econometrica*, 64(3):575–596.

[31] McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative risk management.* Princton: Princeton University Press.

[32] Mukherjee, K. (2008). M-estimation in GARCH models. *Econometric Theory*, 24(6):1530–1553.

[33] Nelson, D. B. (1990). Stationarity and persistence in the GARCH$(1, 1)$ model. *Econometric Theory*, 6(3):318–334.

[34] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59(2):347–370.

[35] Pan, J., Wang, H., and Tong, H. (2008). Estimation and tests for power-transformed and threshold GARCH models. *J. Econometrics*, 142(1):352–378.

[36] Pfanzagl, J. (1969). On the measurability and consistency of minimum contrast estimates. *Metrika*, 14:249–272.

[37] Rao, C. R. and Rao, M. B. (1998). *Matrix algebra and its applications to statistics and econometrics.* River Edge: World Scientific Publishing.

[38] Rockinger, M. and Jondeau, E. (2002). Entropy densities with an application to autoregressive conditional skewness and kurtosis. *J. Econometrics*, 106(1):119–142.

[39] Rudin, W. (1976). *Principles of mathematical analysis.* New York: McGraw-Hill, 3rd edition.

[40] Schwert, G. W. (1989). Why does stock market volatility change over time? *J. Finance*, 44:1115–1153.

[41] Straumann, D. (2005). *Estimation in conditionally heteroscedastic time series models*, volume 181 of *Lecture Notes in Statistics.* Berlin: Springer.

[42] Straumann, D. and Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equations approach. *Ann. Statist.*, 34(5):2449–2495.

[43] Wald, A. (1949). Note on the consistency of the maximum likelihood estimate. *The Annals of Mathematical Statistics*, 20(4):595–601.

[44] Weiss, A. (1986). Asymptotic theory for ARCH models: Estimation and testing. *Econometric Theory*, 2:107–131.

[45] White, H. (2001). *Asymptotic theory for econometricians.* New York: Academic press.

[46] Zakoïan, J. (1994). Threshold heteroskedastic models. *J. Econom. Dynam. Control*, 18:931–955.

Diskussionspapiere 2013
Discussion Papers 2013

01/2013 **Wrede, Matthias**: Rational choice of itemized deductions

02/2013 **Wrede, Matthias**: Fair Inheritance Taxation in the Presence of Tax Planning

Diskussionspapiere 2012
Discussion Papers 2012

01/2012 **Wrede, Matthias**: Wages, Rents, Unemployment, and the Quality of Life

02/2012 **Schild, Christopher-Johannes:** Trust and Innovation Activity in European Regions - A Geographic Instrumental Variables Approach

03/2012 **Fischer, Matthias:** A skew and leptokurtic distribution with polynomial tails and characterizing functions in closed form

04/2012 **Wrede, Matthias**: Heterogeneous Skills and Homogeneous Land: Segmentation and Agglomeration

05/2012 **Ardelean, Vlad**: Detecting Outliers in Time Series

Diskussionspapiere 2011
Discussion Papers 2011

01/2011 **Klein, Ingo, Fischer, Matthias and Pleier, Thomas:** Weighted Power Mean Copulas: Theory and Application

02/2011 **Kiss, David:** The Impact of Peer Ability and Heterogeneity on Student Achievement: Evidence from a Natural Experiment

03/2011 **Zibrowius, Michael:** Convergence or divergence? Immigrant wage assimilation patterns in Germany

04/2011 **Klein, Ingo and Christa, Florian:** Families of Copulas closed under the Construction of Generalized Linear Means

05/2011 **Schnitzlein, Daniel**: How important is the family? Evidence from sibling correlations in permanent earnings in the US, Germany and Denmark

06/2011 **Schnitzlein, Daniel**: How important is cultural background for the level of intergenerational mobility?

07/2011 **Steffen Mueller**: Teacher Experience and the Class Size Effect - Experimental Evidence

08/2011 **Klein, Ingo:** Van Zwet Ordering for Fechner Asymmetry

09/2011 **Tinkl, Fabian and Reichert Katja:** Dynamic copula-based Markov chains at work: Theory, testing and performance in modeling daily stock returns

10/2011 **Hirsch, Boris and Schnabel, Claus:** Let's Take Bargaining Models Seriously: The Decline in Union Power in Germany, 1992 – 2009

11/2011 **Lechmann, Daniel S.J. and Schnabel**, **Claus** : Are the self-employed really jacks-of-all-trades? Testing the assumptions and implications of Lazear's theory of entrepreneurship with German data

12/2011 **Wrede, Matthias**: Unemployment, Commuting, and Search Intensity

13/2011 **Klein, Ingo**: Van Zwet Ordering and the Ferreira-Steel Family of Skewed Distributions

Diskussionspapiere 2010
Discussion Papers 2010

01/2010 **Mosthaf, Alexander, Schnabel, Claus and Stephani, Jens:** Low-wage careers: Are there dead-end firms and dead-end jobs?

02/2010 **Schlüter, Stephan and Matt Davison:** Pricing an European Gas Storage Facility using a Continuous-Time Spot Price Model with GARCH Diffusion

03/2010 **Fischer, Matthias, Gao, Yang and Herrmann, Klaus:** Volatility Models with Innovations from New Maximum Entropy Densities at Work

04/2010 **Schlüter, Stephan and Deuschle, Carola:** Using Wavelets for Time Series Forecasting – Does it Pay Off?

05/2010 **Feicht, Robert and Stummer, Wolfgang:** Complete closed-form solution to a stochastic growth model and corresponding speed of economic recovery.

06/2010 **Hirsch, Boris and Schnabel, Claus:** Women Move Differently: Job Separations and Gender.

07/2010 **Gartner, Hermann, Schank, Thorsten and Schnabel, Claus:** Wage cyclicality under different regimes of industrial relations.

08/2010 **Tinkl, Fabian**: A note on Hadamard differentiability and differentiability in quadratic mean.

Diskussionspapiere 2009
Discussion Papers 2009

01/2009 **Addison, John T. and Claus Schnabel:** Worker Directors: A German Product that Didn't Export?

02/2009 **Uhde, André and Ulrich Heimeshoff:** Consolidation in banking and financial stability in Europe: Empirical evidence

03/2009 **Gu, Yiquan and Tobias Wenzel:** Product Variety, Price Elasticity of Demand and Fixed Cost in Spatial Models

04/2009 **Schlüter, Stephan:** A Two-Factor Model for Electricity Prices with Dynamic Volatility

05/2009 **Schlüter, Stephan and Fischer, Matthias**: A Tail Quantile Approximation Formula for the Student t and the Symmetric Generalized Hyperbolic Distribution

06/2009 **Ardelean, Vlad:** The impacts of outliers on different estimators for GARCH processes: an empirical study

07/2009 **Herrmann, Klaus:** Non-Extensitivity versus Informative Moments for Financial Models - A Unifying Framework and Empirical Results

08/2009 **Herr, Annika:** Product differentiation and welfare in a mixed duopoly with regulated prices: The case of a public and a private hospital

09/2009 **Dewenter, Ralf, Haucap, Justus and Wenzel, Tobias**: Indirect Network Effects with Two Salop Circles: The Example of the Music Industry

10/2009 **Stuehmeier, Torben and Wenzel, Tobias**: Getting Beer During Commercials: Adverse Effects of Ad-Avoidance

11/2009 **Klein, Ingo, Köck, Christian and Tinkl, Fabian**: Spatial-serial dependency in multivariate GARCH models and dynamic copulas: A simulation study

12/2009 **Schlüter, Stephan**: Constructing a Quasilinear Moving Average Using the Scaling Function

13/2009 **Blien, Uwe, Dauth, Wolfgang, Schank, Thorsten and Schnabel, Claus:** The institutional context of an "empirical law": The wage curve under different regimes of collective bargaining

14/2009 **Mosthaf, Alexander, Schank, Thorsten and Schnabel, Claus:** Low-wage employment versus unemployment: Which one provides better prospects for women?

Diskussionspapiere 2008
Discussion Papers 2008

01/2008 **Grimm, Veronika and Gregor Zoettl**: Strategic Capacity Choice under Uncertainty: The Impact of Market Structure on Investment and Welfare

02/2008 **Grimm, Veronika and Gregor Zoettl**: Production under Uncertainty: A Characterization of Welfare Enhancing and Optimal Price Caps

03/2008 **Engelmann, Dirk and Veronika Grimm**: Mechanisms for Efficient Voting with Private Information about Preferences

04/2008 **Schnabel, Claus and Joachim Wagner**: The Aging of the Unions in West Germany, 1980-2006

05/2008 **Wenzel, Tobias:** On the Incentives to Form Strategic Coalitions in ATM Markets

06/2008 **Herrmann, Klaus**: Models for Time-varying Moments Using Maximum Entropy Applied to a Generalized Measure of Volatility

07/2008 **Klein, Ingo and Michael Grottke:** On J.M. Keynes' "The Principal Averages and the Laws of Error which Lead to Them" - Refinement and Generalisation