ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF ECONOMICS

WORKING PAPER SERIES **12-2012**

First Order Asymptotic Theory for the QMLE of the GQARCH(1,1) model

Stelios Arvanitis and Alexandros Louka

76 Patission Str., Athens 104 34, Greece
Tel. (++30) 210-8203911 - Fax: (++30) 210-8203301
www.econ.aueb.gr

First Order Asymptotic Theory for the QMLE of the GQARCH(1,1) model

Stelios Arvanitis and Alexandros Louka
Athens University of Economics and Business

October 11, 2012

Abstract

We examine the asymptotic properties of the QMLE for the $\mathrm{GQARCH}\,(1,1)$ model. Under suitable conditions, we establish that the asymptotic distribution of $n^{\frac{a-1}{a}}\,(\mathrm{QMLE}-\theta_0)$ is characterized by the one of the unique minimizer of a quadratic form over a closed and convex subset of $\mathbb{R}^4$. This respesents the squared distance (w.r.t. a p.d. matrix) from a random vector that follows a normal distribution when $a=2$ or is a linear transformation of an α-stable random vector when $a \in (1,2)$. When the parameter is an interior point this implies that we have distributional convergence to this random vector. Hence we have examined cases in which non normal asymptotic distributions are obtained either due to convergence of the estimator on the boundary of the parameter space, and/or due to the non existence of high order moments for random elements involved in this framework. Possible extensions concern the establishment of analogous results for indirect estimators based on the QMLE which could have desirable first order asymptotic properties.

KEYWORDS: Conditional heteroskedasticity, quadratic ARCH models, stochastic recurrence equation, stationarity, ergodicity, quasi likelihood, normal integrand, epi-convergence, martingale CLT, $\mathrm{CLT}'\,s$ to a-stable distributions, inner limit of set sequences, weak convergence to minimizers of quadratic forms over convex sets.

1 Introduction

We examine the asymptotic properties of the QMLE for the $\mathrm{GQARCH}\,(1,1)$ model. The latter falls into the general class of quadratic ARCH models which was introduced in 1995 by Sentana [11]. This provided the most general formulation of any element of the conditional variance process as a (non anticipative) quadratic function of elements of the QARCH process itself. This formulation allows for the representation of negative dynamic asymmetry (partly attributed to the so called "leverage effect") that is a frequent empirical exhibited in financial time series.

Our motivation stems from the fact that even though the analogous properties of the QMLE were studied for a wide class of conditionally heteroskedastic models (for a detailed catalogue see for example Straumann [9]), this model was not included in any of these results. Our methodology allows for cases where the true parameter vector lies on the boundary of the parameter space and/or the random variables of the innovation process upon which the conditional variance process is constructed do not possess finite fourth moments. The asymptotic theory of M-estimators when the true parameter lies on the boundary has already been studied in the context of the $\mathrm{GARCH}\,(p,q)$ model in the more general work of Andrews [1]. Our considerations involve a slightly different approach for the approximation of the sequence of shifted and normalized parameter spaces that seems more general. Finally, to our knowledge, there are no analogous results for the asymptotic properties of the QMLE even for the already studied heteroskedastic models when the fourth moment of the innovations is not finite.

Our methodology and the remaining structure of the paper is as follows. We first present the model and discuss how the chosen parameterization enables the subsequent approach. Secondly, using the stochastic recurrence equation (SRE) theory as in Straumann [9], we provide sufficient conditions for the existence of a unique stationary ergodic solution to the associated recurrence relation and prove that there exist (among others weakly non stationary) $\mathrm{GQARCH}\,(1,1)$ processes for which these are satisfied. The presence of the autoregressive parameter in the conditional variance SRE complicates the issue of the derivation of tighter conditions, yet the consideration of the QARCH case provides us with some directions for future research.

Third, given the previous we define the usual in applications version of the quasi likelihood function emerging from filtering the volatility by arbitrary initial conditions and its stationary and ergodic approximation. We show that these functions are P a.s. twice differentiable on the largest subset of $\mathbb{R}^4$ for which the model is well defined, even in cases where the associated partial derivatives are obtained by one sided differentiation as in Andrews [1]. Given results concerning approximations of the non stationary by the stationary and ergodic likelihoods and their derivatives, and via the use of an LLN concerning stationary and ergodic random functions which may attain the value $+\infty$, and then the use of a CLT for stationary and ergodic martingale differences when finite fourth moments exist, or of a CLT for convergence in distribution to an α-stable random vector (see Theorem B.1 of Surgailis [10]) when finite absolute moments of some order in $(3,4)$ exist, we show first consistency and then that the rate of convergence is $n^{\frac{a-1}{a}}$ for $a=2$ and $a\in(1,2)$ in the first and second case respectively.

Finally, when the parameter space can be suitably approximated by some closed and convex subset of $\mathbb{R}^4$, we obtain via the use of Lemma 7.13 of van der Vaart [13] that the asymptotic distribution of $n^{\frac{a-1}{a}}\,(\mathrm{QMLE}-\theta_0)$ is characterized as the one of the unique minimizer over that set of a quadratic form w.r.t. a p.d. matrix and a random vector that follows a normal distribution when $a=2$ or is a linear transformation of an α-stable random vector when $a\in(1,2)$. We conclude by posing some questions for future research emerging from our results. The proofs of the main propositions are presented in the main body of the paper. Several auxiliary technical lemmas are presented in the appendix.

2 Model Specification

In the following let $(\Omega,\mathcal{F},P)$ denote a complete probability space and Θ a non empty subset of the Euclidean space $\mathbb{R}^4$. Any concept of measurability is in any case handled w.r.t. to $\mathcal{F}$, the Borel σ-fields of $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{\mathbb{Z}}$ (the latter w.r.t. the product topology) and Θ or where appropriate w.r.t. analogous product σ-fields. Notice that separability and completeness imply measurability of $\inf_A \|\cdot\|$ where A is either $\mathbb{R}^{\mathbb{Z}}$ or any sequentially complete subset of Θ. Let $z : \Omega \to \mathbb{R}^{\mathbb{Z}}$ be an iid sequence of random variables, with $\mathbb{E}z_0 = 0$, and $\mathbb{E}z_0^2 = 1$. Define the volatility process $(\sigma_t^2)_{t\in\mathbb{Z}}$ as any ($\mathbb{Z}$ indexed) sequence of non negative random variables satisfying the following first order stochastic recurrence equation (see Sentana [11])

$$\sigma_t^2(\theta) = \omega + \alpha\left(z_{t-1}\sigma_{t-1}(\theta) + \frac{\gamma}{2\alpha}\right)^2 + \beta\sigma_{t-1}^2(\theta) \tag{1}$$

where $\theta = (\omega, a, \gamma, \beta)' \in \Theta$. *Given* the existence of $(\sigma_t^2)_{t\in\mathbb{Z}}$ construct the $\mathrm{GQARCH}(1,1)$ process $(y_t)_{t\in\mathbb{Z}}$ by

$$y_t = z_t\sigma_t$$

In the following paragraph we provide some justification for the choice of parameterization in (1), and simultaneously specify Θ as the largest subset of $\mathbb{R}^4$ so that any process satisfying (1) is well defined (i.e. σ_t^2 is finite and non-negative P a.s. for any t) *and* the methodology used for the derivation of asymptotic properties is valid.

Parameter Restrictions-Positivity Constraints. A necessary condition for the existence of a solution to (1) is that σ_t^2 is finite and non-negative P a.s. for any t. We easily obtain from the analogous $\mathrm{GARCH}(1,1)$ case and the choice of parameterization in (1) that this holds strictly iff $\omega > 0$, $a > 0$, $\beta \geq 0$. Notice that this encompasses *neither* the case of non random, yet time varying volatility *nor* the one of homoskedasticity since the choice $a = 0$ is obviously not allowed. A reformulation of (1) so as to contain the case of $a = 0$ could be obtained if we specified $\sigma_t^2 = \omega + \beta\sigma_{t-1}^2$ when $a = 0$, hence impose the convention that when $a = 0$ then $\gamma = 0$ and furthermore that $\frac{0}{0} = 0$. This however would invalidate the methods used for the determination of the rates of convergence and the asymptotic distribution of the QMLE below.[1]

Notice here that we could have considered two alternative parameterizations of (1) so as to disentangle this restriction. The first one would be $\sigma_t^2 = \omega^* + \alpha^* z_{t-1}^2\sigma_{t-1}^2(\theta) + \gamma^* z_{t-1}\sigma_{t-1} + \beta\sigma_{t-1}^2$. In this case it is already known from Sentana [11] that strict positivity P a.s. for any t holds iff $\omega^* > 0$, $a^* \geq 0$, $\beta^* \geq 0$ and $\omega^* a^* > \frac{(\gamma^*)^2}{4}$ when $a^* > 0$ and $\gamma^* = 0$ when $a^* = 0$. This case obviously invalidates lemma 2.1 below. The second one would be $\sigma_t^2 = \omega_* + \alpha_*\left(z_{t-1}\sigma_{t-1} + \gamma_*\right)^2 + \beta_*\sigma_{t-1}^2$. For this strict positivity P a.s. for any t holds iff $\omega_* > 0$, $a_* \geq 0$, $\beta_* \geq 0$. In this case when $a_* = 0$ if γ_* assumes more than one values then it remains non identified. This problem is fixed if γ_* assumes one value (say 0). Notice however that the derivatives of the volatility process w.r.t. γ_* (or γ^* in the previous case) evaluated at any point of the form $(\omega_*, 0, 0, 0)'$ would be identically zero, rendering the methodology used given consistency invalid. Hence we stick with the parametrization in (1) and restrict $a > 0$. Notice that this restriction implies that the asymptotic properties of the QMLE in any of the last two parameterizations can be directly recovered by the analogous properties of the QMLE in the case considered via the continuous

[1] If we had adopted this reformulation then given the framework (to be defined later) of the auxiliary volatility process and the likelihood function we have that the latter would remain a P a.s. lower semicontinuous function and the arguments in the proof of proposition 3.5 would remain intact. Hence the QMLE would be consistent even when $\theta_0 = (\omega_0, 0, 0, \beta_0)$ except for the case where $\beta_0 \neq 0$ since this would obviously invalidate lemma 3.3. Then it is easy to see that for $\theta_0 = (\omega_0, 0, 0, 0)$ lemma 3.6 would not hold (the likelihood function would not posses Frechet derivatives at θ_0) hence the subsequent methodology would break down.

mapping theorem and the delta method (which generally goes through by Theorem 20.8 of Van der Vaart [13]).

Furthermore, we allow γ to assume only non positive values, something that is in accordance to the empirical stylized fact of the leverage effect in financial time series. Hence this constraint has *econometric significance*. Finally in order to facilitate the following result (first among others) we further constrain β to be strictly less than 1. Given the previous we define the eligible parameter space as

$$\Theta = \mathbb{R}^{++} \times \mathbb{R}^{++} \times \mathbb{R}^{-} \times [0, 1)$$

In what follows the inference procedures to be examined will be defined so as to employ compact subsets of Θ. The following result is quite useful. It states that for any $\theta \in \Theta$ the volatility process is bounded away from zero.

Lemma 2.1 *If for some $\theta \in \Theta$ there exists a volatility process satisfying (1) then $\inf_{z\in\mathbb{R}^{\mathbb{Z}}} \sigma_t^2 \geq \frac{\omega}{1-\beta}$ $\forall t \in \mathbb{Z}$.*

Proof. Minimizing $\omega + \alpha \left(z_{t-1}\sigma_{t-1}(\theta) + \frac{\gamma}{2\alpha}\right)^2 + \beta\sigma_{t-1}^2(\theta)$ given σ_{t-1}^2 we obtain

$$\inf_{z_{t-1}\in\mathbb{R}} \sigma_t^2 = \omega + \beta\sigma_{t-1}^2$$

Proceeding recursively the result follows. ■

Existence-Stationarity-Ergodicity. Given the construction of Θ, we use the SRE approach in Straumann [9] in order to show that there exist elements of Θ for which (1) assumes a unique stationary ergodic solution.[2] From (1) and lemma 2.1 it is obvious that

$$\sigma_t^2 = \phi_t\left(\sigma_{t-1}^2\right), \quad t \in \mathbb{Z}$$

where $\phi_t : \left[\frac{\omega}{1-\beta}, \infty\right) \to \left[\frac{\omega}{1-\beta}, \infty\right)$ with

$$\phi_t(s) = \omega + \alpha\left(z_{t-1}\sqrt{s} + \frac{\gamma}{2\alpha}\right)^2 + \beta s.$$

Due to the properties of z and Proposition 2.1.1 of Straumann [9] this specifies a stationary ergodic sequence of random map and the

$$\Lambda\left(\phi_t\right) = \sup_{s\in\left[\frac{\omega}{1-\beta},\infty\right)} \left|\alpha z_{t-1}^2 + \beta + \frac{\gamma z_{t-1}}{2\sqrt{s}}\right| < \infty.$$

specifies the analogous stationary ergodic sequence of their random Lipschitz coefficients.

Lemma 2.2 $\mathbb{E}\left[\log^{+}\Lambda\left(\phi_t\right)\right] < \infty$.

[2] Uniqueness holds in the sense that any other solution belongs to the same equivalence class w.r.t. the exponentialy almost sure (w.r.t. P) convergence as $t \to \infty$. See Straumann [9], Theorem 2.6.1.

Proof. Observe that $\Lambda(\phi_t) \leq az_{t-1}^2 + \beta + \sup_{x\in\left[\frac{\omega}{1-\beta},\infty\right)} \frac{-\gamma|z_{t-1}|}{2\sqrt{x}}$ which, due to Lemma 2.1, equals $az_{t-1}^2 + \beta + \frac{-\gamma|z_{t-1}|}{2\sqrt{\frac{\omega}{1-\beta}}} < \infty$. Hence, $\mathbb{E}\Lambda(\phi_t) < \infty$ which completes the proof. ■

Hence Theorem 2.6.1 of Straumann [9], and Lemma 2.2, the SRE (1) admits a unique stationary ergodic solution $(\sigma_t^2)_{t\in\mathbb{Z}}$ when $\theta \in \Theta \cap \{\theta : \mathbb{E}[\log \Lambda(\phi_t)] < 0\}$ which has the P a.s. representation

$$\sigma_t^2(\theta) = \lim_{m\to\infty} \phi_{t-1} \circ \ldots \circ \phi_{t-m}(y), \quad t \in \mathbb{Z} \tag{2}$$

Furthermore by Proposition 2.1.1 of Straumann [9] $(y_t)_{t\in\mathbb{Z}}$ is also stationary ergodic.

We term the elements of Θ that also satisfy the aforementioned restriction as *ergodic* by an obvious abuse of terminology. The following remark implies that such ergodic θ exist and furthermore that there exist ergodic θ that imply second order *non* stationarity for the $\mathrm{GQARCH}(1,1)$ process.

Remark R.1 *From the analogous results for the* $\mathrm{GARCH}(1,1)$ *we know that there exist* $\alpha > 0$, $\beta > 0$ *so that the corresponding volatility process is stationary ergodic (see Nelson [6]). Let such a pair of* (a, β). *Also we have that so that* $E(\sigma_0^2)$ *does not exist. Now* $\sup_{s\in\left[\frac{\omega}{1-\beta},\infty\right)} \left|\alpha z^2 + \beta + \frac{\gamma z}{2\sqrt{s}}\right| \leq \alpha z^2 + \beta - \frac{\gamma|z|}{2\sqrt{\frac{\omega}{1-\beta}}} \leq \alpha z^2 + \beta - \frac{\gamma(1+z^2)}{2\sqrt{\frac{\omega}{1-\beta}}} = \left(\alpha - \frac{\gamma}{2\sqrt{\frac{\omega}{1-\beta}}}\right) z^2 + \left(\beta - \frac{\gamma}{2\sqrt{\frac{\omega}{1-\beta}}}\right)$. *Then due to the continuity of* $E\log(\alpha z_0^2 + \beta)$ *in* (α, β) *and Jensen's inequality there exist* ω *and* γ *(that can be chosen strictly negative) so that* $E\log(\alpha z_0^2 + \beta) < 0$ *implies that*

$$\mathbb{E}\log\left(\left(\alpha - \frac{\gamma}{2\sqrt{\frac{\omega}{1-\beta}}}\right) z^2 + \left(\beta - \frac{\gamma}{2\sqrt{\frac{\omega}{1-\beta}}}\right)\right) < 0.$$

Hence there exist ergodic θ. *Finally notice that from the results for the* $\mathrm{GARCH}(1,1)$ *case* a *and* β *can be chosen so that* $\alpha + \beta \geq 1$. *Hence there exist ergodic* θ *that imply first order non stationarity for the volatility process (hence covariance non stationarity for the* $\mathrm{GQARCH}(1,1)$ *process).*

The next remark implies that the condition described above can be restrictive. It concerns the $\mathrm{QARCH}(1)$ case.

Remark R.2 *Suppose that* $\beta = 0$, *then the condition* $\mathbb{E}\log(\alpha z_0^2) < 0$ *(which is also necessary and sufficient condition for the the existence and uniqueness of stationary and ergodic solution in the* $\mathrm{ARCH}(1)$ *case) is sufficient. To see this, consider the* SRE *describing the squared root of the conditional variance process in the* $\mathrm{QARCH}(1)$ *case, i.e.*

$$\sigma_t = \phi_t(\sigma_{t-1}), \quad t \in \mathbb{Z}$$

with $\phi_t(s) = \left[\omega + \alpha\left(z_{t-1}s + \frac{\gamma}{2\alpha}\right)^2\right]^{1/2}$. *Then* $|\phi_t'(s)| = \left|\frac{\sqrt{\alpha}\left(z_{t-1}s+\frac{\gamma}{2\alpha}\right)}{\phi_t(s)}\right| \sqrt{\alpha}\,|z_{t-1}| \leq \sqrt{\alpha}\,|z_{t-1}|$ *since* $\left|\frac{\sqrt{\alpha}\left(z_{t-1}s+\frac{\gamma}{2\alpha}\right)}{\phi_t(s)}\right| = \sqrt{\frac{\alpha\left(z_{t-1}s+\frac{\gamma}{2\alpha}\right)^2}{\omega+\alpha\left(z_{t-1}s+\frac{\gamma}{2\alpha}\right)^2}} < 1$ *as* $\omega > 0$. *This along with partial evidence from* *simulations lead us to* **conjecture** *that the necessary and sufficient condition for existence and uniqueness of stationary and ergodic solution for the* $\mathrm{GARCH}(1,1)$ *model, i.e.* $\mathbb{E}\log(\alpha z_0^2+\beta)<0$*, is also sufficient for the* $\mathrm{GQARCH}(1,1)$ *model.*

3 First Order Theory for Quasi Maximum Likelihood Estimation

We are interested in the behavior of the QMLE for θ_0 when the latter is an ergodic point. We restate known ergodic and non ergodic versions of the Quasi Likelihood function based on volatility filters (auxiliary volatility processes) constructed from the GQARCH process and the form of the SRE in (1). In the first case the filter is defined on $\mathbb{Z}$ whereas in the second one it is assumed to stem from an arbitrary initial condition posed on the same SRE.

In any of the two cases it is easily deduced that the necessary condition of finiteness and non-negativity P a.s. for any t, for the existence of processes that satisfy the analogous SRE (or the initial condition problem) are satisfied due to the definition of Θ (see the discussion in paragraph 2). Obviously only the QMLE associated with the second case is practically feasible.

Auxiliary Ergodic Volatility Process and the Ergodic Quasi Likelihood Function. We first consider the ergodic case.

Definition D.1 *For any* $\theta\in\Theta$ *and given* $(y_i(\theta_0))_{t\in\mathbb{Z}}$ *for* θ_0 *an ergodic point, define the random element* $(h_t)_{t\in\mathbb{Z}}$ *by the following* SRE

$$h_t=\omega+\alpha\left(y_{t-1}(\theta_0)+\frac{\gamma}{2\alpha}\right)^2+\beta h_{t-1}.$$

Lemma 3.1 *The previous* P *a.s. admits a unique stationary and ergodic solution* $(h_t)_{t\in\mathbb{Z}}$ *of the form*

$$h_t=\frac{\omega}{1-\beta}+\alpha\sum\nolimits_{i=0}^{\infty}\beta^i\left(y_{t-1-i}(\theta_0)+\frac{\gamma}{2\alpha}\right)^2.$$

Moreover $\inf_K\inf_{y\in\mathbb{R}^{\mathbb{Z}}}h_t(\theta)>0$ *and independent of* t *for* K *any compact subset of* Θ*.*

Proof. It follows from the definition of Θ, Proposition 5.2.12 of Straumann [9] and lemma 2.1. ■

The following defines the (infeasible w.r.t. applications) Quasi Likelihood function. The term is used in an abusive manner since the original function would be constructed as $-\frac{1}{2}*c_n(\theta)+\mathrm{const}$. This form enables the characterization of the QMLE as a minimizer.

Definition D.2 *For any* $\theta\in\Theta$ *and given* $(y_i(\theta_0))_{t\in\mathbb{Z}}$ *for* θ_0 *an ergodic point, consider*

$$c_n(\theta)=\frac{1}{n}\sum_{i=1}^{n}\ell_i(\theta)$$

where

$$\ell_i(\theta)=\log h_i(\theta)+\frac{y_i^2(\theta_0)}{h_i(\theta)}$$

Term c_n *the ergodic quasi likelihood function of the* $\mathrm{GQARCH}(1,1)$ *process.*

Remark R.3 *c_n is continuous on Θ due to the previous lemma, and jointly measurable, hence when restricted to any compact subset of Θ (say K) it is a normal integrand in the sense of definition 3.5 of [5]. Also due to the ergodicity of σ_t^2 and h_t, c_n defines an ergodic process for any θ.*

Non Ergodic Volatility Process and the non Ergodic Quasi Likelihood Function. We study versions of the processes in the previous definitions emerging from arbitrary initial conditions. In this stationarity and ergodicity are lost but feasibility w.r.t. applications is obtained.

Definition D.3 *For any $\theta \in \Theta$ and given $(y_i(\theta_0))_{t\in\mathbb{Z}}$ define the random element $(h_t)_{t\geq 0}$ by the following*

$$h_t^*(\theta) = \begin{cases} \varsigma_0 \text{ if } t = 0 \\ \omega + \alpha\left(y_{t-1} + \frac{\gamma}{2\alpha}\right)^2 + \beta h_{t-1}^*(\theta) \text{ if } t \geq 1 \end{cases}$$

for some positive random variable ς_0 where again $\theta = (\omega, a, \gamma, \beta)' \in \Theta$.

Remark R.4 *Obviously $(h_t^*)_{t\geq 0}$ is well defined but generally non stationary. Also $\inf_K \inf_{y\in\mathbb{R}^{\mathbb{Z}++}} h_t^*(\theta) > 0$ and independent of t for K any compact subset of Θ.*

Definition D.4 *Consider*

$$c_n^*(\omega, \theta) = \frac{1}{n}\sum_{i=1i}^{n} \ell_i^*(\theta)$$

where

$$\ell_i^*(\theta) = \ln h_i^*(\theta) + \frac{y_i^2(\theta_0)}{h_i^*(\theta)}$$

We term c_n^* the *non ergodic quasi likelihood function* of the $\mathrm{GARCH}(1,1)$ process.

Remark R.5 *c_n^* is also continuous on Θ as well as jointly measurable, hence it is a normal integrand in the sense of definition 3.5 of [5] when restricted to K an arbitrary compact subset of Θ.*

3.1 Existence and Consistency of the QMLE Estimator

Definition and Existence

The following assumption defines the parameter space that will be subsequently used.

Assumption A.1 *Given an ergodic θ_0, K is a compact subset of Θ for which $\theta_0 \in K$.*

Notice that given the definition of Θ, K could be chosen from some further available information for θ_0. The following propositions define and provide the existence for the QMLE w.r.t. the two versions of the likelihood functions presented before. We allow for the case that the estimators are approximate maximizers and thereby there exist optimization errors.

Proposition 3.2 *For an arbitrary P a.s. non negative random variable ε_n there exists a random element θ_n with values in K defined by*

$$c_n^* (\theta_n) \leq \inf_K c_n^* (\theta) + \varepsilon_n$$

Proof. The result follows by remark R.3 which renders applicable Proposition 3.12.iii and the fundamental selection theorem (Theorem 2.13) of [5]. ■

Notice that the continuity arguments for the existence and measurability of the estimator are essential only in the case that the optimization error is P a.s. zero.

Consistency

The following lemmas provide with an identification condition, existence of $\log$ moments and enable the approximation between the two versions of the likelihood function. They are used for the derivation of consistency. We enable them by the use of the following assumptions.

Assumption A.2 *The distribution of z_0 is not concentrated in two points.*

Assumption A.3 $\varepsilon_n \to 0$ P *a.s.*

Assumption A.4 $\mathbb{E} \log^+ (\varsigma_0) < \infty$.

The first one implies the validity of asymptotic identification. The second that the optimization error is asymptotically negligible. The third is not very restrictive and permits the derivation of consistency for θ_n from the approximation of c_n^* by c_n.

Lemma 3.3 *Under assumption A.2 the relation $h_t(\theta) = h_t(\theta_0)$ P a.s. $\forall \theta \in K$, $\forall t$ implies $\theta = \theta_0$.*

Proof. Towards a contradiction, suppose that there exist $\theta \neq \theta_0$, so that $h_t(\theta) = h_t(\theta_0)$ P a.s. $\forall t$. Then

$$(\omega^* - \omega_0^*) + (\alpha - \alpha_0) y_{t-1}^2 + (\gamma - \gamma_0) y_{t-1} + (\beta - \beta_0) h_{t-1} = 0 \quad \forall t$$

where we define $\omega^* = \omega + \frac{\gamma^2}{4\alpha}$ and $\omega_0^* = \omega_0 + \frac{\gamma_0^2}{4\alpha_0}$. However, this implies that h_{t-1} is at the same time a measurable function of z_{t-1} and independent of z_{t-1}. By Lemma 5.4.2 of Straumann [9] the only way for this to be possible is if h_{t-1} is constant $\forall t$ P a.s. Suppose $h_t = h$ constant This necessarily implies that $\alpha + \beta < 1$. Taking expectations on the volatility process we must have $h = \frac{\omega^*}{1-\alpha-\beta}$. But

$$h = \omega^* + (\alpha z_0^2 + \beta) h + \gamma z_0 \sqrt{h}$$

which is equivalent to

$$z_0^2 + \left(\frac{\gamma}{\alpha} \sqrt{\frac{1-\alpha-\beta}{\omega^*}} \right) z_0 - 1 = 0$$

and this is a second order equation in z_{t-1} with positive discriminant. It has two roots, one positive and one negative. This would imply that the support of the distribution of z_0 contains exactly two points. This violates assumption A.2 thus necessarily $\beta = \beta_0$. Now the following equation

$$(\alpha - \alpha_0)y^2 + (\gamma - \gamma_0)y + (\omega^* - \omega_0^*) = 0$$

is a second order equation in y. Since $(\omega^* - \omega_0^*, \alpha - \alpha_0, \gamma - \gamma_0) \neq (0,0,0)$, there are three distinct cases concerning its roots: **i)** a single root λ of multiplicity 2, which leads to contradiction due to the facts $\mathbb{E}(y) = 0$, $\mathbb{E}(y^2) \neq 0$, **ii)** two roots of the same sign, which leads to contradiction since either $\mathbb{E}(y)$ does not exist or equals 0, and **iii)** two roots of alternating sign, in which case the contradiction is a consequence of the fact that z_t is independent of h_t and h_t is not a constant. Then necessarily $(\omega^* - \omega_0^*, \alpha - \alpha_0, \gamma - \gamma_0) = (0,0,0)$. The definition of ω^* and the fact that γ^* is non positive completes the proof. ■

Lemma 3.4 $\mathbb{E}\log^+ \sigma_t^2 < \infty$, $\mathbb{E}\log^+ y_t^2 < \infty$ *and* $\mathbb{E}\log^+ h_t(\theta) < \infty$ *for any* $\theta \in K$.

Proof. First notice that due to lemma 2.1

$$\begin{aligned}
\sigma_t^2 &\leq \omega_0 + \frac{\gamma_0^2}{4a_0} + \left(a_0 z_{t-1}^2 + \beta_0 + \frac{\gamma_0 z_{t-1}}{\sigma_{t-1}} \mathbf{1}\{z_{t-1} < 0\} \mathbf{1}\left\{\sigma_{t-1} \geq 2\sqrt{\frac{\omega_0}{1-\beta_0}}\right\}\right) \sigma_{t-1}^2 \\
&\quad + \gamma_0 z_{t-1} \mathbf{1}\{z_{t-1} < 0\} \mathbf{1}\left\{\sigma_{t-1} < 2\sqrt{\frac{\omega_0}{1-\beta_0}}\right\} \sigma_{t-1} \\
&\leq \omega_0 + \frac{\gamma_0^2}{4a_0} + \left(a_0 z_{t-1}^2 + \beta_0 + \frac{\gamma_0 z_{t-1}}{2\sqrt{\frac{\omega_0}{1-\beta_0}}} \mathbf{1}\{z_{t-1} < 0\} \mathbf{1}\left\{\sigma_{t-1} \geq 2\sqrt{\frac{\omega_0}{1-\beta_0}}\right\}\right) \sigma_{t-1}^2 \\
&\quad + 2\sqrt{\frac{\omega_0}{1-\beta_0}} \gamma_0 z_{t-1} \mathbf{1}\{z_{t-1} < 0\} \\
&\leq A_t \sigma_{t-1}^2 + C_{t-1}
\end{aligned}$$

where $A_t \doteqdot a_0 z_{t-1}^2 + \beta_0 + \frac{\gamma_0 z_{t-1}}{2\sqrt{\frac{\omega_0}{1-\beta_0}}} \mathbf{1}\{z_{t-1} < 0\} < \Lambda(\phi_t)$ and $C_{t-1} \doteqdot \omega_0 + \frac{\gamma_0^2}{4a_0} + 2\sqrt{\frac{\omega_0}{1-\beta_0}} \gamma_0 z_{t-1} \mathbf{1}\{z_{t-1} < 0\}$. Then from the definition of θ_0 we have that $\mathbb{E}\log A_t \leq \mathbb{E}\log \Lambda(\phi_t) < 0$. If we define the auxiliary SRE

$$s_t^2 = A_{t-1} s_{t-1}^2 + C_{t-1}$$

we can use similar arguments as in Example 5.2.5 of Straumann [9] to show that $\exists \eta : 0 < \eta \leq 1$ so that $\mathbb{E}[s_t^2]^\eta < \infty$. Then

$$\mathbb{E}\log^+ \sigma_t^2 \leq \mathbb{E}\log^+ s_t^2(\theta) \leq \frac{1}{\eta}\mathbb{E}\log^+ \left[s_t^{2\eta}\right] \leq \frac{1}{\eta}\log^+ \left[\mathbb{E}s_t^{2\eta}\right] < \infty$$

The above implies $\mathbb{E}\log^+ y_t^2 < \infty$ as $\mathbb{E}(z_t^2) = 1$. Furthermore, $\mathbb{E}\log^+ h_t(\theta) < \infty$ can be shown by an application of the Minkowski inequality to the P a.s. representation

$$h_t(\theta) = \frac{\omega + \frac{\gamma^2}{4a}}{1-\beta} + \sum_{i=0}^{\infty} \beta^i \left(a y_{t-i-1}^2 + \gamma y_{t-i-1}\right)$$

which exists due to the fact that θ_0 is ergodic and $\beta < 1$, to obtain $\mathbb{E}[h_t(\theta)]^\eta < \infty$. ■

Given identification and existence of log moments strong consistency follows for the estimator under examination.

Proposition 3.5 *Under assumptions A.1, A.2, A.3 and A.4 θ_n is P-strongly consistent.*

Proof. If $E\log^+(\varsigma_0)<\infty$ then Proposition 5.2.12 of [9] holds which implies that

$$|c_n(\theta)-c_n^*(\theta)|_K\to 0\ P\text{ a.s.}$$

due to Part 1.(i) of the proof of Theorem 5.3.1 of Straumann [9]. Then notice that

$$\begin{aligned}\mathbb{E}\inf_{\theta\in K}\left|\ln h_0(\theta)+\frac{z_0^2\sigma_0^2(\theta_0)}{h_0(\theta)}\right| &\leq \inf_{\theta\in K}\left(\mathbb{E}\left|\ln h_0(\theta)\right|+\mathbb{E}\frac{\sigma_0^2}{h_0(\theta)}\right)\\ &= \mathbb{E}\left|\ln h_0(\theta_0)\right|+1<\infty\end{aligned}$$

where the last equality follows from Part 1.(iii) of the proof of Theorem 5.3.1 of [9] and $\mathbb{E}\left|\ln h_0(\theta_0)\right|$ exists due to lemma 3.4 and the fact that $h_0(\theta_0)=\sigma_0^2\ P$ a.s. from lemma 3.3. Second, notice that from the pointwise ergodic theorem, $c_n(\theta)$ converges almost surely to its expectation, a function with values on the extended real line, which does not assume the value $-\infty$, since by lemma 3.1 $\inf_K\inf_{y\in\mathbb{R}^{\mathbb{Z}}}h_t(\theta)>0$ independent of t, and is proper due to the argument of the previous sentence. Hence, from Theorem 2.1 of [4] c_n epiconverges almost surely to its expectation, which due to stationarity is $\mathbb{E}\ell_0$. Hence, due to the fact that the topology of uniform convergence is finer from the topology of epiconvergence, and due to the previous we have that c_n^* epiconverges almost surely to $\mathbb{E}\ell_0$. Therefore, due to the almost sure convergence of ε_n to zero, there exists a measurable $\Omega^*\subseteq\Omega$ with $P(\Omega^*)=1$ such that for any $\omega\in\Omega^*$, Theorem 7.31 of [8] implies that $\limsup_{n\to\infty}(\varepsilon_n-\arg\min_K)(c_n^*)\subseteq\arg\min_K(\mathbb{E}\ell_0)$.[3] Due to lemma 3.3 we have that for all $\omega\in\Omega$ $\arg\min_K(\mathbb{E}\ell_0)=\{\theta_0\}$ which implies $\lim_{n\to\infty}(\varepsilon_n-\arg\min_K)(c_n^*)=\{\theta_0\}$, P a.s. since by proposition 3.2 θ_n exists and belongs to $(\varepsilon_n-\arg\min_K)(c_n)$. ■

3.2 Rates of Convergence

Given consistency, we establish the remaining first order asymptotic properties of the estimator via the use of local quadratic approximations of the likelihood functions. These are enabled by the existence of first and second order (possibly one sided) partial derivatives of c_n and c_n^* at θ_0 and/or neighboring points, which are well defined even in cases where θ_0 is a boundary point of Θ, due to the form of Θ and the definitions of the likelihood functions and the volatility filters. Then the quadratic approximations emerge from second order Taylor expansions of the likelihoods in neighborhoods of θ_0 and remain valid when restricted to random elements assuming their values in K, even though the remainders may depend on random elements assuming values in compact subsets of Θ that contain θ_0 but are not subsets of K. In the following for $\delta>0$, $C(\theta,\delta)$, and $S(\theta,\delta)$ denote the open cube and the sphere in $\mathbb{R}^4$ centered at θ respectively.

In what follows we partially differentiate w.r.t. θ when perceived as a member of Θ. The vector of first order partial (possibly one sided) derivatives is denoted by c_n' (resp. $c_n'^*$) and the matrix of second order partial derivatives by c_n'' (resp. $c_n''^*$). The vector of first order partial derivatives is denoted by c_n' (resp. $c_n'^*$) and the matrix of second order partial derivatives by c_n'' (resp. $c_n''^*$). Their forms are well known in the literature concerning the asymptotic theory for the

[3] Notice that the P a.s. continuity of c_n and the mode of convergence implies the lower semicontinuity of $\mathbb{E}\ell_0$ (see proposition 7.4.a of [8]). This along with the compactness of K imply that $\inf_K\mathbb{E}\ell_0$ is well separated.

QMLE in GARCH type models. We denote with $\ell_t(\theta)$, and $h'_t(\theta)$ (resp. $\ell'^*_t(\theta)$ and h'^*_t) and $\ell''_t(\theta)$ and $h''_t(\theta)$ (resp. $\ell''^*_t(\theta)$ h''^*_t) the analogous structures concerning the derivatives of the random variables $\ell_t(\theta)$ and h_t (resp. $\ell_t(\theta)$ and h^*_t) w.r.t. θ on K^+. These are explicitly derived and used in the appendix for the establishment of several intermediate results. The arguments of the previous paragraph are based on the following lemma.

Lemma 3.6 *Assumption 2^{2^*}.(a) of Andrews [1] holds for Θ.*

Proof. When θ_0 belongs to the interior of Θ, choose $\delta < \inf \|\theta_0 - y\|$ where the infimum is taken w.r.t. the boundary points of Θ and it is strictly positive since the boundary is a closed set. Let (in the notation of Assumption 2^{2^*}) $K^+ = C(\theta_0, \delta)$ and notice that $K^+ - \theta_0 = C(0, \delta)$ and then for any $\delta^* < \delta$ $S(\theta_0, \delta) \subset C(\theta_0, \delta)$. When θ_0 belongs to the interior of Θ, choose $K^+ = C(\theta_0, \delta) \cap \Theta$ for arbitrary δ and notice that $K^+ - \theta_0$ is the intersection of $\mathbb{R}^4$ with $C(0, \delta)$ and then for any $\delta^* < \delta$, $S(\theta_0, \delta^*) \subset K^+$. ■

We establish the rate of convergence for θ_n. This given the previous lemma essentially depends on the differentiability (possibly in the sense of Theorem 6 of Andrews [1]) of the likelihoods w.r.t. $\theta \in \Theta$, their proximity and particular properties of the stationary distribution of the innovation process. The following lemma enables the main proposition of the present section. This result will obviously be also used for the establishment of the asymptotic distribution. We denote convergence in distribution with $\rightsquigarrow$.

Lemma 3.7 *Under assumption A.2* **i***) If $\mathbb{E}z_0^4 < +\infty$ then $\sqrt{n}c'_n(\theta_0) \xrightarrow{d} V_2 \sim N(\mathbf{0}, \mathbf{G}_0)$ as $n \to \infty$ where $\mathbf{G}_0 = (\mathbb{E}z_t^4 - 1)\mathbb{J}$ and $\mathbb{J}$ a positive definite matrix defined in lemma 4.6.* **ii***) If $P\left(z_0 \in \left[-\sqrt{x+1}, \sqrt{x+1}\right]\right) = 1 - \frac{c_2+o(1)}{x^\alpha} h(x)$ as $x \to +\infty$, for $c_2 > 0$, $\alpha \in (1,2)$ and h a slowly varying function at infinity (in the Karamata sense),[4] then $n^{\frac{\alpha-1}{a}} c'_n(\theta_0) \rightsquigarrow -V_\alpha$ where V_α follows an α-stable distribution on $\mathbb{R}^4$ characterized as follows: for any non zero $\lambda \in \mathbb{R}^4$ $\lambda^T V_\alpha$ follows an α-stable distribution on $\mathbb{R}$ with characteristic function $f_{\lambda^T c_\alpha}(\theta) = \exp\left(-c(\lambda)|\theta|^\alpha \left(1 - i\beta(\lambda)\operatorname{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right)\right)$ where $c(\lambda) = -\frac{\Gamma(2-\alpha)}{a-1} c_2 \cos\left(\frac{\pi\alpha}{2}\right) E\left(\left|\lambda^T \frac{h'_t(\theta_0)}{\sigma_t^2}\right|^\alpha\right)$ and $\beta(\lambda) = -\frac{E\left(\left|\lambda^T \frac{h'_t(\theta_0)}{\sigma_t^2}\right|^\alpha \operatorname{sgn}\left(\lambda^T \frac{h'_t(\theta_0)}{\sigma_t^2}\right)\right)}{E\left(\left|\lambda^T \frac{h'_t(\theta_0)}{\sigma_t^2}\right|^\alpha\right)}$.* **iii***) If assumption A.4 holds then for any K^* compact subset of Θ, $n^{\frac{\alpha-1}{a}} \|c'_n(\theta) - c^{*\prime}_n(\theta)\|_{K^*} \to 0$ P a.s., for $a = 2$ when $\mathbb{E}z_0^4 < +\infty$ or $\alpha \in (1,2)$ when the analogous condition in* **ii***) holds and therefore, if additionally assumption A.2 holds then $n^{\frac{\alpha-1}{a}} c^{*\prime}_n(\theta_0) \rightsquigarrow V_2$ in the first and $n^{\frac{\alpha-1}{a}} c^{*\prime}_n(\theta_0) \rightsquigarrow -V_a$ in the second case.*

Proof. First notice that due to the chain rule (that holds for one sided partial derivatives) we have that

$$-c'_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{y_t^2}{h_t(\theta)} - 1\right)\frac{h'_t(\theta)}{h_t(\theta)}$$

where h'_t is given in lemma 4.1 in the Appendix. Hence by exploiting the fact that $h_t(\theta_0) = \sigma_t^2$ P a.s. we obtain that

$$-c'_n(\theta_0) = \frac{1}{n}\sum_{t=1}^{n}(z_t^2 - 1)\frac{h'_t(\theta_0)}{\sigma_t^2}.$$

[4]Remember that $h : \mathbb{R}^+ \to \mathbb{R}^+$ is slowly varying at infinity (in the Karamata's sense) if and only if for any $c > 0$, $\lim_{x\to\infty} \frac{h(ax)}{h(x)} = 1$.

Then notice that the random element $\frac{h_t'(\theta_0)}{\sigma_t^2}$ is measurable with respect to $\mathcal{F}_{t-1} = \{z_{t-k}, k \geq 1\}$ and $\mathcal{F}_{t-1}$ is independent of z_t and $\mathbb{E}z_t^2 = 1$. **i**) Since $\mathbb{E}z_0^4 < +\infty$ and due to Lemma 4.1 the sequence $(\ell_t'(\theta_0))_{t\in\mathbb{N}}$ is a finite variance stationary ergodic zero-mean martingale difference sequence with respect to the filtration $(\mathcal{F}_{t-1})_{t\in\mathbb{N}}$. Consequently we can apply the central limit theorem for finite variance stationary ergodic martingale difference sequences to obtain the result. Furthermore

$$\begin{aligned}
\mathbf{G}_0 &= \mathbb{E}\left[\ell_t'(\theta_0)(\ell_t'(\theta_0))^T\right] = \mathbb{E}\left[\frac{h_t'(\theta_0)(h_t'(\theta_0))^T}{\sigma_0^4}\right]\mathbb{E}\left[(z_t^2-1)^2\right] \\
&= \mathbb{E}\left[\frac{h_t'(\theta_0)(h_t'(\theta_0))^T}{\sigma_0^4}\right]\left(\mathbb{E}z_t^4-1\right) = \left(\mathbb{E}z_t^4-1\right)\mathbb{J}
\end{aligned}$$

By Lemma 4.6 and since $\mathbb{E}z_t^4 = \mathbb{E}\left(z_t^2\right)^2 > \left(\mathbb{E}z_t^2\right)^2$, it follows that $\mathbf{G}_0$ is a positive definite matrix. **ii**) The assumption on the asymptotic behavior of the distribution of z_0 is equivalent to

$$P\left(z_0^2 - 1 \leq x\right) = 1 - \frac{c_2 + o(1)}{x^\alpha} h(x) \text{ as } x \to +\infty$$

Then due to Theorem 2.6.1 of Ibragimov and Linnik [3] it follows that the distribution of $z_0^2 - 1$ ($c_1 = 0$) lies in the domain of attraction of an α-stable distribution. Notice that this and Theorem 2.6.3 of Ibragimov and Linnik [3] imply that $E\left(z_0^2\right)^\gamma < +\infty$ for any $1 \leq \gamma < \alpha$. Then Theorem 2.6.5 of Ibragimov and Linnik [3] implies that in a neighborhood of the origin the characteristic function of $z_0^2 - 1$ is of the form $\exp\left(-c|\theta|^\alpha\left(1 - i\beta\operatorname{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right)\right)$ where $c = -\frac{\Gamma(2-\alpha)}{a-1}c_2\cos\left(\frac{\pi\alpha}{2}\right)$ and $\beta = -1$. Then Theorem B.1 of Surgailis [10] along with lemma 4.1 imply that $-n^{\frac{\alpha-1}{a}}\lambda^T c_n'(\theta_0)$ converges in distribution to an α-stable random variable for any non zero λ. The result follows by the Cramer-Wald device and the Continuous Mapping Theorem. **iii**) We have that P a.s.

$$\begin{aligned}
& n^{\frac{\alpha-1}{a}}\left\|c_n'(\theta_0) - c_n^{*\prime}(\theta_0)\right\|_{K^*} \\
\leq\; & \frac{1}{n^{\frac{1}{a}}}\sum_{t=1}^n\left\|\left(\frac{y_t^2}{h_t(\theta)} - 1\right)\frac{h_t'(\theta)}{h_t(\theta)} - \left(\frac{y_t^2}{h_t^*(\theta)} - 1\right)\frac{h_t^{*\prime}(\theta)}{h_t^*(\theta)}\right\|_{K^*} \\
\leq\; & \frac{1}{n^{\frac{1}{a}}}\sum_{t=1}^n y_t^2\left|h_t^{-2}(\theta) - h_t^{-2*}(\theta)\right|_{K^*}\left\|h_t'(\theta) - h_t^{*\prime}(\theta)\right\|_{K^*} \\
& + \frac{1}{n^{\frac{1}{a}}}\sum_{t=1}^n\left|h_t^{-1}(\theta) - h_t^{-1*}(\theta)\right|_{K^*}\left\|h_t'(\theta) - h_t^{*\prime}(\theta)\right\|_{K^*}
\end{aligned}$$

Exploiting the fact that h_t and h_t^* are P a.s. uniformly over t and any compact subset of Θ bounded away from zero we obtain from the mean value theorem that there exist positive constants c_1, c_2 for which the majorant side in the previous display is P a.s. less than or equal to

$$\begin{aligned}
& \frac{c_1}{n^{\frac{1}{a}}}\sum_{t=1}^n y_t^2\left|h_t(\theta) - h_t^*(\theta)\right|_{K^*}\left\|h_t'(\theta) - h_t^{*\prime}(\theta)\right\|_{K^*} \\
& + \frac{c_2}{n^{\frac{1}{a}}}\sum_{t=1}^n\left|h_t(\theta) - h_t^*(\theta)\right|_{K^*}\left\|h_t'(\theta) - h_t^{*\prime}(\theta)\right\|_{K^*}
\end{aligned}$$

The first result follows from lemma 4.10, Proposition 5.2.12 of Straumann [9] which holds due to the definition of θ_0, assumption A.4, lemma 3.4 which enables Proposition 2.5.1 of Straumann [9] and the fact that in any case a is positive. The second result is a trivial consequence of the previous. ■

Hence the following result obtained by restricting the rate of convergence of the optimization error.

Proposition 3.8 *Under assumptions A.1, A.2, A.4 and if moreover* $\varepsilon_n = o_p\left(n^{-k}\right)$ *for* $k \geq 2 - \frac{2}{a}$ *then*

$$n^{\frac{\alpha-1}{\alpha}}\left(\theta_n - \theta_0\right) = O_p\left(1\right)$$

with $\alpha = 2$ *when* $\mathbb{E}z_0^4 < \infty$ *and* $\alpha \in (1,2)$ *when* $P\left(z_0 \in \left[-\sqrt{x+1}, \sqrt{x+1}\right]\right) = 1 - \frac{c_2 + o(1)}{x^{\alpha}} h\left(x\right)$ *as* $x \to +\infty$, *for* $c_2 > 0$ *and* h *as in proposition 3.7.*

Proof. First notice that by the definition of θ_n we have

$$\sum\nolimits_{t=1}^{n} \ell_t^*\left(\theta_n\right) - \sum\nolimits_{t=1}^{n} \ell_t^*\left(\theta_0\right) \leq o_p\left(n^{1-k}\right)$$

From lemma 3.6 and the definition of ℓ_t^* and h_t^* we have that since in any case $1 < a \leq 2$ the previous implies

$$\nu_n^T \frac{1}{n^{\frac{1}{a}}} \sum\nolimits_{t=1}^{n} \ell_t'^*\left(\theta_0\right) + \frac{1}{2}\nu_n^T c_t''^*\left(\theta_n^*\right)\nu_n \leq o_p\left(n^{-\varepsilon}\right) \leq o_p\left(1\right)$$

where $\varepsilon = k - 2 + \frac{2}{a} \geq 0$ by hypothesis and $\nu_n = n^{\frac{a-1}{a}}\left(\theta_n - \theta_0\right)$ and θ_n^* is a random element with values in the line segment between θ_n and θ_0 P a.s. which can stay outside K with positive probability. Obviously due to lemma 3.4 $\bar{\theta}_n$ converges to θ_0 P a.s. Using this we can choose $\varepsilon > 0$ so that lemma 4.5 holds and by additionally employing lemma 4.11 in the Appendix we have that the previous can be expressed as

$$\nu_n^T \frac{1}{n^{\frac{1}{a}}} \sum\nolimits_{t=1}^{n} \ell_t'^*\left(\theta_0\right) + \frac{1}{2}\nu_n^T \left(E\ell_0''\left(\theta_0\right) + o_p\left(1\right)\right)\nu_n \leq o_p\left(1\right)$$

From lemma 4.6 we have that $E\ell_0''\left(\theta_0\right)$ is a positive definite matrix and using this along with lemma 4.5. **iii**) we obtain that there exists some positive $c > 0$ such that

$$O_p\left(\left\|\nu_n\right\|\right) - c\left\|\nu_n\right\|^2 + \left\|\nu_n\right\|^2 o_p\left(1\right) \geq o_p\left(1\right)$$

which implies that

$$O_p\left(1\right) \geq \left\|\nu_n\right\|^2\left(1 + o_p\left(1\right)\right) - 2\left\|\nu_n\right\|\left(1 + o_p\left(1\right)\right)O_p\left(1\right) + O_p\left(1\right)$$

Hence

$$\left\|\nu_n\right\|\left(1 + o_p\left(1\right)\right) \leq O_p\left(1\right)$$

and the result follows. ■

3.3 Asymptotic Distribution

In order to characterize the asymptotic distribution of $n^{\frac{\alpha-1}{\alpha}}(\theta_n-\theta_0)$ we first, following van der Vaart [13] (see paragraph 7.4), characterize the asymptotic parameter space as a convenient limit of the sequence of centered and rescaled parameter spaces and then impose some further structure in the assumption that follows. For a as in the previous section, we denote with $H_n(a)$ the $n^{\frac{a-1}{a}}(K-\theta_0)=\left\{n^{\frac{a-1}{a}}(x-\theta_0),x\in K\right\}$ and notice that given the assumption A.1 $H_n(a)$ is compact and contains 0.

Definition D.5 $H(a)=\limsup_{n\to\infty}H_n(a)$ *i.e. it is the set containing any* $x\in\mathbb{R}^4$ *such that* x *is a cluster point of some* $(x_n)_{n\in\mathbb{N}}$ *with* $x_n\in H_n(a)$.

Notice that $H(a)$ always exists and it is a closed subset of $\mathbb{R}^4$ (see Proposition 4.4 of Rockafellar and Wets [8]). In our case it is always different from $\varnothing$ since it contains 0. From exercise 4.2 of Rockafellar and Wets [8] we obtain that $x\in H(a)$ iff $\liminf_{n\to\infty}n^a\inf_{y\in K-\theta_0}\left\|\frac{x}{n^a}-y\right\|=0$. But the limit inferior equals to $\liminf_{n\to\infty}n^{a^*}\inf_{y\in\frac{1}{n^{a^*-a}}(K-\theta_0)}\left\|\frac{x}{n^{a^*}}-y\right\|$. Hence a sufficient condition for $H(a)=H(a^*)$ is that $K-\theta_0$ is a cone. The next assumption enables the uniqueness of the main result.

Assumption A.5 $H(a)$ *is convex.*

We first present the main result and then discuss some examples.

Proposition 3.9 *Under assumptions A.1, A.2, A.4, A.5 and if moreover* $\varepsilon_n=o_p\left(n^{-k}\right)$ *for* $k\geq 2-\frac{2}{a}$, *then*

$$n^{\frac{\alpha-1}{\alpha}}(\theta_n-\theta_0)\rightsquigarrow\tilde{h}$$

with $\tilde{h}$ *defined uniquely by* $q\left(\tilde{h}\right)=\inf_{h\in H(a)}q(h)$ *and* $q(h):=(h-\mathbb{J}^{-1}Z)'\,\mathbb{J}\,(h-\mathbb{J}^{-1}Z)$ *for* $\mathbb{J}=E\ell_0''(\theta_0)$ *positive definite, and using the notation and definitions and of lemma 3.7 i)* $Z=V_2$ *when* $\alpha=2$ *when* $Ez_0^4<\infty$ *i.e.* $Z\sim N\left(0,(\mathbb{E}z_t^4-1)\,\mathbb{J}\right)$ *or ii)* $Z=-V_\alpha$ *for* $\alpha\in(1,2)$ *when* $P\left(z_0\in\left[-\sqrt{x+1},\sqrt{x+1}\right]\right)=1-\frac{c_2+o(1)}{x^\alpha}h(x)$ *as* $x\to\infty$ *and* h *slowly varying at infinity and* $c_2>0$ *where* V_α *follows the* α*-stable distribution characterized in lemma 3.7.*

Proof. From lemma 3.6 and the definitions of the likelihoods in each of the cases described in the proposition, we can define $\varpi_n:\Theta\to\mathbb{R}$ as

$$\begin{aligned}\varpi_n(h) &\equiv n^{2\frac{a-1}{a}}\left(c_n^*\left(\theta_0+n^{-\frac{\alpha-1}{\alpha}}h\right)-c_n^*(\theta_0)\right)\\ &= h'\frac{1}{n^{\frac{1}{a}}}\sum_{t=1}^{n}\ell_t'^*(\theta_0)+\frac{1}{2}h'c_t''^*(\theta_0)\,h+\frac{1}{2}h'\left(c_n''^*\left(\bar{\theta}_n\right)-c_n''^*(\theta_0)\right)h\end{aligned}$$

where $\bar{\theta}_n$ is a random element with values in the line segment between $n^{-\frac{\alpha-1}{\alpha}}h+\theta_0$ and θ_0 P a.s. that lies inside Θ. Obviously due to lemma 3.4 $\bar{\theta}_n$ converges to θ_0 P a.s. Using this we can choose $\varepsilon > 0$ so that lemma 4.5 holds and by additionally employing lemma 4.11 in the Appendix and lemma 3.7, we have that for U an arbitrary compact subset of Θ

$$\left|\varpi_n(h) - h'Z - \frac{1}{2}h'\mathbb{J}h\right|_U = o_p\left(1\right). \tag{3}$$

Due to proposition 3.8 $h_n \doteqdot n^{\frac{\alpha-1}{\alpha}}\left(\theta_n - \theta_0\right) \in H_n\left(a\right) \cap B\left(0, n^{\frac{\alpha-1}{\alpha}}\varepsilon\right) \doteqdot H_n^*\left(a\right)$ with P-probability tending to 1 for some $\varepsilon > 0$. If F is a closed non empty subset of $\mathbb{R}^4$, and $h_n \in F$, then for large enough n, either $H_n^*\left(a\right) \subset F$, or $H_n^*\left(a\right) \nsubseteq F$ but $H_n^*\left(a\right) \cap F \neq \varnothing$. In either case due to the definitions of θ_n, ϖ_n and the fact that $\varepsilon_n = o_p\left(n^{-k}\right)$ for $k \geq 2 - \frac{2}{a}$

$$\inf_{h \in H_n^*(a) \cap F} \varpi_n\left(h\right) \leq \inf_{h \in H_n^*(a)} \varpi_n\left(h\right) + o_p\left(1\right)$$

and therefore due to Slutsky's lemma

$$\begin{aligned} P\left(h_n \in F\right) &\leq P\left(\inf_{h \in H_n^*(a) \cap F} \varpi_n\left(h\right) \leq \inf_{h \in H_n^*(a)} \varpi_n\left(h\right) + o_p\left(1\right)\right) \\ &\leq P\left(\inf_{h \in H_n^*(a) \cap F} \varpi_n\left(h\right) \leq \inf_{h \in H_n^*(a)} \varpi_n\left(h\right)\right) + o\left(1\right) \end{aligned}$$

Now notice that $H_n^* = H_n^* \cap \mathbb{R}^q$ and $\mathbb{R}^q$ is open, $\limsup_{n\to\infty} H_n^*\left(a\right) = H\left(a\right)$, since $\limsup_{n\to\infty} H_n\left(a\right) = H\left(a\right)$ and $n^{\frac{\alpha-1}{\alpha}} \to \infty$. Furthermore equation 3 and the continuous mapping theorem imply that Lemma 7.13.2-3 of van der Vaart [13] is applicable, so that the last probability is less than or equal to

$$P\left(\inf_{h \in H(a) \cap F} \varpi_n\left(h\right) \leq \inf_{h \in H(a)} \varpi_n\left(h\right) + o_p\left(1\right)\right) \leq P\left(\inf_{h \in H(a) \cap F} \varpi_n\left(h\right) \leq \inf_{h \in H(a)} \varpi_n\left(h\right)\right) + o\left(1\right)$$

due to Slutsky's Lemma. Again from equation 3, the continuous mapping theorem and Slutsky's Lemma we obtain that the last probability is less than or equal to

$$P\left(\inf_{h \in H(a) \cap F} h'Z - \frac{1}{2}h'\mathbb{J}h \leq \inf_{h \in H(a)} h'Z - \frac{1}{2}h'\mathbb{J}h\right) + o\left(1\right)$$

Now

$$\begin{aligned} &P\left(\inf_{h \in H(a) \cap F} h'Z + \frac{1}{2}h'\mathbb{J}h \leq \inf_{h \in H(a)} h'Z + \frac{1}{2}h'\mathbb{J}h\right) \\ = \;& P\left(\inf_{h \in H(a) \cap F} h'Z + \frac{1}{2}h'\mathbb{J}h \pm \frac{1}{2}Z'\mathbb{J}^{-1}Z \leq \inf_{h \in H(a)} h'Z + \frac{1}{2}h'\mathbb{J}h \pm \frac{1}{2}Z'\mathbb{J}^{-1}Z\right) \\ = \;& P\left(\inf_{h \in H(a) \cap F} \left(h - \mathbb{J}^{-1}Z\right)'\mathbb{J}\left(h - \mathbb{J}^{-1}Z\right) \leq \inf_{h \in H(a)} \left(h - \mathbb{J}^{-1}Z\right)'\mathbb{J}\left(h - \mathbb{J}^{-1}Z\right)\right) \end{aligned}$$

Assumption A.5 implies that $H\left(a\right)$ is closed and convex, and lemma 4.6 that $\mathbb{J}$ is positive definite. Hence due to uniqueness when

$$\inf_{h \in H(a) \cap F} \left(h - \mathbb{J}^{-1}Z\right)'\mathbb{J}\left(h - \mathbb{J}^{-1}Z\right) \leq \inf_{h \in H(a)} \left(h - \mathbb{J}^{-1}Z\right)'\mathbb{J}\left(h - \mathbb{J}^{-1}Z\right)$$

holds then

$$\tilde{h} \in H(a) \cap F$$

and therefore the last probability is less than or equal to

$$P\left(\tilde{h} \in H(a) \cap F\right) \leq P\left(\tilde{h} \in F\right)$$

hence we have proven that

$$\limsup_{n\to\infty} P(h_n \in F) \leq P\left(\tilde{h} \in F\right)$$

and the result follows from the Portmanteau theorem due to the fact that F is chosen arbitrarily. Notice that when $Ez_0^4 < \infty$ from lemma 3.7 we obtain that $a = 2$ and $Z \sim N\left(0, \left(\mathbb{E}z_t^4 - 1\right)\mathbb{J}\right)$ while from the same result when $a \in (1,2)$, $P\left(z_0 \in \left[-\sqrt{x+1}, \sqrt{x+1}\right]\right) = 1 - \frac{c_2+o(1)}{x^\alpha} h(x)$ as $x \to \infty$, $c_2 > 0$ and h slowly varying at infinity $Z = -V_a$ which follows the a-stable distribution in the second case of lemma 3.7. ■

We close this paragraph by providing some examples concerning the form of K and the subsequent form of $H(a)$. Some of these have obvious econometric significance. In the last one we examine a simple case where assumption A.5 does not hold.

Example: interior point. Suppose that $\theta_0 \in \operatorname{Int} K$. This implies that there exists an open ball centered at θ_0 that lies entirely inside K. Then definition D.5 implies that $H(a) = \mathbb{R}^4$. Hence in the first case of the previous proposition $n^{\frac{\alpha-1}{\alpha}}(\theta_n - \theta_0) \rightsquigarrow N\left(0, \left(\mathbb{E}z_t^4 - 1\right)\mathbb{J}^{-1}\right)$ and in the second one $n^{\frac{\alpha-1}{\alpha}}(\theta_n - \theta_0) \rightsquigarrow -\mathbb{J}^{-1}V_a$.

Example: $\mathrm{GARCH}(1,1)$**.** Suppose that $\theta_0 = (\omega_0, \alpha_0, 0, \beta_0)$ and $K = [\omega_l, \omega_u] \times [\alpha_l, \alpha_u] \times [\gamma_l, 0] \times [0, \beta_u]$ where $0 < \omega_l < \omega_0 < \omega_u$, $0 < \alpha_l < \alpha_0 < \alpha_u$, $0 < \beta_0 < \beta_u$ and $\gamma_l < 0$. Then the centered parameter space is $K - \theta_0 = [\omega_l - \omega_0, \omega_u - \omega_0] \times [\alpha_l - \alpha_0, \alpha_u - \alpha_0] \times [\gamma_l, 0] \times [-\beta_0, \beta_u - \beta_0]$. Hence $H(a) = \mathbb{R}^2 \times (-\infty, 0] \times \mathbb{R}$ and assumption A.5 holds. Given the econometric significance of the negative dynamic asymmetry, the result in the following proposition implies that the restriction of γ as above is asymptotically informative when $\gamma_0 = 0$.

Example: $\mathrm{QARCH}(1,1)$**.** Let $\theta_0 = (\omega_0, \alpha_0, \gamma_0, 0)$ and $K = [\omega_l, \omega_u] \times [\alpha_l, \alpha_u] \times [\gamma_l, 0] \times [0, \beta_u]$ where $0 < \omega_l < \omega_0 < \omega_u$, $0 < \alpha_0 < \alpha_u$, $\gamma_l < \gamma_0 < 0$ and $\beta_u > 0$. Then the centered parameter space is $K - \theta_0 = [\omega_l - \omega_0, \omega_u - \omega_0] \times [-\alpha_0, \alpha_u - \alpha_0] \times [\gamma_l - \gamma_0, -\gamma_0] \times [0, \beta_u]$ and $H(a) = \mathbb{R}^3 \times \mathbb{R}^+$, and again assumption A.5 holds.

Example: discrete K**.** Suppose that K is discrete hence due to assumption A.1 it is finite. Then $K - \theta_0$ is also finite and $H(a) = \{0\}$. Assumption A.5 holds and the previous proposition implies that in any of the two cases the asymptotic distribution is degenerate at 0.

Example: θ_0 **lies on a sphere.** Suppose that $\theta^* \in \operatorname{Int}\Theta$, $K = \bar{B}(\theta^*, \varepsilon)$ for $\varepsilon \leq \min(|\gamma^*|, \beta^*, 1 - \beta^*)$, and $\|\theta_0 - \theta^*\| = \varepsilon$. Then $K - \theta_0 = \bar{B}(\theta^* - \theta_0, \varepsilon)$. Since $K - \theta_0$ is closed and convex there exists a supporting hyperplane and thereby $H(a)$ is the closed half space containing $K - \theta_0$ and the hyperplane. If $K = S(\theta^*, \varepsilon)$ then $H(a)$ equals the supporting hyperplane itself. In any case assumption A.5 holds.

Example: θ_0 **lies on the boundary of an annulus.** Suppose that θ^* and ε are as before, and $K = \bar{B}(\theta^*, \varepsilon) / B(\theta^*, \varepsilon_0)$ for $\varepsilon_0 < \varepsilon$, and $\|\theta_0 - \theta^*\| = \varepsilon_0$ or $\|\theta_0 - \theta^*\| = \varepsilon$. Let $\mathcal{H}_\varepsilon$ denote the supporting hyperplane of $\bar{B}(\theta^* - \theta_0, \varepsilon)$ and $\mathcal{CH}_\varepsilon$ the closed half space containing $\bar{B}(\theta^* - \theta_0, \varepsilon)$ and $\mathcal{H}_\varepsilon$. In the first case $H(a) = \mathcal{CH}_\varepsilon^c \cup \mathcal{H}_\varepsilon$ and in the second $H(a) = \mathcal{H}_\varepsilon$. In both cases assumption A.5 holds.

Example: K **is a union of line segments.** For some ergodic $\theta_0 \in \Theta$ suppose that K is the union of a countable collection of line segments, each one of which contains θ_0 and lies inside $B(\theta_0, \delta)$ for $\delta < \inf_{y \in \mathrm{Bd}\,\Theta} \|\theta_0 - y\|$ where $\mathrm{Bd}\,\Theta$ denotes the boundary of Θ. Hence $K - \theta_0$ is the union of a countable collection of line segments which contain 0 and have length less than δ. $H(a)$ then equals the countable union of the one dimensional subspaces that are uniquely defined by each of the line segments in $K - \theta_0$. Obviously assumption A.5 holds iff all the initial line segments are colinear.

4 Further Research

We have examined the asymptotic properties of the QMLE of the $\mathrm{GQARCH}(1,1)$ model. Under our assumption framework, we established that the asymptotic distribution of $n^{\frac{a-1}{a}}(\mathrm{QMLE} - \theta_0)$ is characterized by the one of the unique minimizer of a quadratic form over a closed and convex subset of $\mathbb{R}^4$.[5] This respesents the squared distance (w.r.t. a p.d. matrix) from a random vector that follows a normal distribution when $a = 2$ or is a linear transformation of an α-stable random vector when $a \in (1, 2)$. When the parameter is an interior point this implies that we have distributional convergence to this random vector. Hence we have examined cases in which non normal asymptotic distributions are obtained either due to convergence of the estimator on the boundary of the parameter space, and/or due to the non existence of high order moments for random elements involved in this framework.

The previous results raise two possible directions for further research. The first concerns the investigation of the conjecture that the necessary and sufficient conditions for stationarity and ergodicity of the $\mathrm{GARCH}(1,1)$ are sufficient for the $\mathrm{GQARCH}(1,1)$. The second concerns the derivation of the asymptotic distribution for particular indirect estimators when the QMLE is used as an auxiliary in the same context. We expect that when θ_0 is a boundary point some of these estimators have "smaller" first order asymptotic bias and mse than the QMLE. We suspect that the derivation of these results would be facilitated by some locally uniform (w.r.t. θ_0) extension of the results in Andrews [1] that moreover enables local quadratic approximations of the likelihood function, without the use of derivatives. Hence it could also enable the incorporation in the asymptotic theory of θ_0 that correspond to conditional homoskedasticity or non random time varying volatility processes.

[5] Notice that we could also have examined the asymptotic properties of the *infeasible* QMLE defined as an approximate minimizer of the ergodic likelihood function. It is easy to see that our assumption framework establishes (without the need of assumption A.4 for this case) that it is asymptoticaly equivalent to the case examined.

References

[1] Andrews, D. W. K., "*Estimation when a parameter is on boundary*", *Econometrica, 67:1341-1383, 1999.*

[2] Berkes, I. , L. Horvath, and P. Kokoszka, "*GARCH processes: structure and estimation*", *Bernoulli,* 9:201-228, 2003.

[3] Ibragimov, I. A., and Yu V. Linnik, "*Independent and Stationary Sequences of Random Variables*", Wolters-Noordhoff, 1971.

[4] Korf, A.L., and Wets J-B, "*An Ergodic Theorem for Stochastic Programming Problems*", unpublished document.

[5] Molchanov, Ilya, "*Theory of Random Sets*" ,Probability and Its Applications, Springer, 2004.

[6] Nelson, D. B., "*Stationarity and Persistence in the GARCH(1,1) Model*", Econometric Theory, 6:318-334, 1990a.

[7] Pfanzagl, J. , "*On the Measurability and Consistency of Minimum Contrast Estimates*", *Metrika*, 14:249-272, 1969.

[8] Rockafellar, T.R., and Wetts J-B, "*Variational Analysis*", Springer-Verlag, 1997.

[9] Straumann, D., "*Estimation in Conditionally Heteroscedastic Time Series Models*", Springer, 2004.

[10] Surgailis, D., "*A Quadratic ARCH(∞) Model with Long Memory and Levy Stable Behavior of Squares*", *Adv. Appl. Prob.*, 40:1198-1222, 2008.

[11] Sentana, E., "*Quadratic ARCH Models*", *The Review of Economic Studies*, 62:639-661, 1995.

[12] van der Vaart, A.W. and J.A. Wellner, "*Weak convergence and empirical processes*" Springer, 2000.

[13] van der Vaart, A.W., "*Asymptotic Statistics*" CUP, 1997.

Appendix

This section contains a sequence of intermediate technical lemmas used in the proofs of the main results of the paper. They are grouped according to their initial use in the establishment of the basic results in the paper. Some of them are actually used in more than one parts of the asymptotic theory. Lemma 3.6 in the main body of the paper enables the partial (possibly one sided) differentiation of any of the volatility filters over Θ. We denote the relevant mathematical entities according to the notation introduced in paragraph 3.2. In the following K^* denotes an arbitrary non empty subset of Θ. Moreover the notation $\|\cdot\|_{K^*}$ denotes $\sup_{K^*}\|\cdot\|$. Notice that separability and compactness of K^* imply measurability of $\sup_{K^*}\|\cdot\|$ (apply the Theorem of Measurable Projections in van der Vaart and Wellner [12], example 1.7.5 p. 47). Furthermore for $\varepsilon > 0$, $\overline{B}(\theta, \varepsilon)$ denotes the closed ball in $\mathbb{R}^4$ of radius ε, centered at θ.

Lemmas enabling the establishment of the rate of convergence of the ergodic QMLE**.**

Lemma 4.1 $\mathbb{E}\left\|\frac{h_t'(\theta)}{h_t(\theta)}\right\|_{K^*}^{\delta} < \infty$ *for all* $\delta > 0$.

Proof. Remember that h_t depends on θ_0 through y_{t-1}, y_{t-2},... From lemma 3.1 we have that P a.s.

$$h_t'(\theta) = \left(\frac{1}{1-\beta},\ -\frac{\gamma^2}{4\alpha^2(1-\beta)} + \sum_{i=0}^{\infty}\beta^i y_{t-i-1}^2,\ \frac{\gamma}{2\alpha(1-\beta)} + \sum_{i=0}^{\infty}\beta^i y_{t-i-1},\right.$$
$$\left.\frac{\omega}{(1-\beta)^2} + \sum_{i=0}^{\infty} i\beta^{i-1}\alpha\left(y_{t-i-1} + \frac{\gamma}{2a}\right)^2\right)'$$

The first element of the vector $\frac{h_t'(\theta)}{h_t(\theta)}$ clearly uniformly bounded. For the second element note that $\forall\theta \in K$ there exist positive constants C_1, C_2 (possibly dependent on θ) so that $y_{t-i-1}^2 \leq C_1 + C_2\alpha\left(y_{t-i-1} + \frac{\gamma}{2\alpha}\right)^2$ for all i. This implies boundedness for the third element as well, observing that $|y_{t-i-1}| \leq 1 + y_{t-i-1}^2$. Due to compactness of K^* these bounds can be made uniform. As for the fourth element see Lemma 4.4. ■

Lemma 4.2 $\mathbb{E}\left\|\frac{h_t''(\theta)}{h_t(\theta)}\right\|_{K^*}^{\delta} < \infty$ *for all* $\delta > 0$.

Proof. Notice first that continuity of the second order (possibly one sided) derivatives implies Young's theorem (see the proof of Theorem 6 of Andrews [1]). The result follows readily from Lemma 4.1 and Lemma 4.4 as P a.s.

$$\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta'} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{(1-\beta)^2} \\ & \frac{\gamma^2}{2a^3(1-\beta)} & -\frac{\gamma}{2a^2(1-\beta)} & -\frac{\gamma^2}{4a^2(1-\beta)^2} + \frac{1}{1-\beta}\sum_{i=0}^{\infty}\beta^i y_{t-i-1}^2 \\ & & \frac{1}{2a(1-\beta)} & \frac{\gamma}{2a(1-\beta)^2} + \frac{1}{1-\beta}\sum_{i=0}^{\infty}\beta^i y_{t-i-1} \\ & & & \frac{2\omega}{(1-\beta)^3} + \sum_{i=0}^{\infty} i(i-1)\beta^{i-2}\alpha\left(y_{t-i-1} + \frac{\gamma}{2a}\right)^2 \end{pmatrix}$$

■

Lemma 4.3 $\mathbb{E}\left(z_0^2\right)^\gamma < \infty$*, for some* γ *and*

$$\lim_{t\to 0} t^{-\mu} P\left(z_0^2 \leq t\right) = 0, \text{ for some } \mu > 0$$

then for any $0 < \nu < \gamma$ *and some small enough* $\varepsilon > 0$

$$\mathbb{E}\left|\frac{\sigma_t^2}{h_t(\theta)}\right|^{\nu}_{\Theta\cap\overline{B}(\theta_0,\varepsilon)} < \infty.$$

Proof. Using $\gamma_0 y_{t-1} \leq -\gamma_0\left(1+y_{t-1}^2\right)$ we have

$$\begin{aligned}\frac{\sigma_t^2}{h_t(\theta)} &\leq \frac{\frac{\omega_0-\gamma_0+\frac{\gamma_0^2}{4a_0}}{\sigma_{t-1}^2} + (a_0-\gamma_0)\, z_{t-1}^2 + \beta_0}{\frac{\omega+\frac{\gamma^2}{4a}}{\sigma_{t-1}^2} + a z_{t-1}^2 + \frac{\gamma z_{t-1}}{\sigma_{t-1}} + \beta\frac{h_{t-1}(\theta)}{\sigma_{t-1}^2}} \\ &\leq \frac{C_0 + (a_0-\gamma_0)\, z_{t-1}^2}{\inf_{\sigma_{t-1}\in\left[\sqrt{\omega_0/(1-\beta_0)},+\infty\right)}\left(\frac{\omega+\frac{\gamma^2}{4a}}{\sigma_{t-1}^2} + \frac{\gamma z_{t-1}}{\sigma_{t-1}} + a z_{t-1}^2\right) + \beta\frac{h_{t-1}(\theta)}{\sigma_{t-1}^2}}\end{aligned}$$

where $C_0 = \frac{\omega_0-\gamma_0+\frac{\gamma_0^2}{4a_0}}{\inf\sigma_{t-1}^2} + \beta_0 = \frac{\omega_0-\gamma_0+\frac{\gamma_0^2}{4a_0}}{\omega_0/(1-\beta_0)} + \beta_0$. Also $\inf_{\sigma_{t-1}\in\left[\sqrt{\omega_0/(1-\beta_0)},+\infty\right)}\left(\frac{\omega+\frac{\gamma^2}{4a}}{\sigma_{t-1}^2} + \frac{\gamma z_{t-1}}{\sigma_{t-1}} + a z_{t-1}^2\right) = \inf_{x\in\left(0,\sqrt{(1-\beta_0)/\omega_0}\right)}\left[\left(\omega+\frac{\gamma^2}{4a}\right)x^2 + (\gamma z_{t-1})\, x + a z_{t-1}^2\right] = \frac{a z_{t-1}^2}{\left(1+\frac{\gamma^2}{4\omega a}\right)}$. Therefore we can write

$$\frac{\sigma_t^2}{h_t(\theta)} \leq \frac{C_0 + (a_0-\gamma_0)\, z_{t-1}^2}{C_1 z_{t-1}^2 + \beta\frac{h_{t-1}(\theta)}{\sigma_{t-1}^2}}$$

where $C_1 = \inf_{\theta\in K^*}\frac{a}{\left(1+\frac{\gamma^2}{4\omega a}\right)} > 0$. Now if $\beta_0 = 0$ choose ε arbitrarily and obtain that $\sup_{\theta\in K^*}\frac{\sigma_t^2}{h_t(\theta)} \leq \frac{C_0+(a_0-\gamma_0)z_{t-1}^2}{C_1 z_{t-1}^2}$ so then skip to equation 4 setting $M=1$ and the result follows using the same arguments. Now when $\beta_0 > 0$ choose ε small enough so that $\Theta\cap\overline{B}(\theta_0,\varepsilon)$ does not contain elements of Θ for which $\beta = 0$. This is always possible due to the fact that $\theta_0 \in K^*$ by assumption. Then there exists $K_1 > 0$ independent of θ so that

$$\frac{\sigma_t^2}{h_t(\theta)} \leq \frac{K_1\left(1+z_{t-1}^2\right)}{z_{t-1}^2 + \left(\frac{\sigma_{t-1}^2}{h_{t-1}(\theta)}\right)^{-1}}.$$

But

$$\left(\frac{\sigma_{t-1}^2}{h_{t-1}(\theta)}\right)^{-1} \geq \frac{z_{t-2}^2 + \left(\frac{\sigma_{t-2}^2}{h_{t-2}(\theta)}\right)^{-1}}{K_1\left(1+z_{t-1}^2\right)}.$$

So, substituting

$$\begin{aligned}\frac{\sigma_t^2}{h_t(\theta)} &\leq \frac{K_1^2\prod_{i=1}^{2}\left(1+z_{t-i}^2\right)}{K_1 z_{t-1}^2+K_1 z_{t-1}^2 z_{t-2}^2+z_{t-2}^2+\left(\frac{\sigma_{t-2}^2}{h_{t-2}(\theta)}\right)^{-1}} \\ &\leq \frac{K_1^2\prod_{i=1}^{2}\left(1+z_{t-i}^2\right)}{K_1 z_{t-1}^2+z_{t-2}^2+\left(\frac{\sigma_{t-2}^2}{h_{t-2}(\theta)}\right)^{-1}}\end{aligned}$$

Hence recursively

$$\frac{\sigma_t^2}{h_t(\theta)} \leq \frac{K_1^M\prod_{i=1}^{M}\left(1+z_{t-i}^2\right)}{K_2\sum_{i=1}^{M} z_{t-i}^2}, \text{ for any } M \geq 1 \tag{4}$$

where K_2 some constant dependent on K_1 and M. Finally following the remaining steps of Lemma 5.1 of Berkes et al. [2] we obtain the result. ■

Lemma 4.4 $\mathbb{E}\left\|\frac{\sum_{i=0}^{\infty} i^3\beta^i y_{t-i-1}^2}{1+\sum_{i=0}^{\infty}\beta^i y_{t-i-1}^2}\right\|_{K^*}^v < \infty$ *for any* $v > 0$.

Proof. First notice that the results of lemma 3.4 would hold even if the ergodic θ_0 did not belong to the examined compact subset of Θ. From this lemma there exists $\delta > 0$ so that $\mathbb{E}\left(y_0^2\right)^{\delta} < \infty$. Then, along the lines of the proof of Lemma 5.2 of Berkes et al. [2], replacing $c_i(u)$ with β^i and noting that we can choose $\beta < \rho_* < 1\ \forall\beta \in K^*$ and $M \geq M_0(\rho)$ large enough so that $i\beta^i < \rho^i$ for $i \geq M$. For any $M \geq 1$, the result follows. ■

Lemma 4.5 *For* ε *as in the previous lemma, if* $\mathbb{E}\left(z_0^2\right)^{\delta} < +\infty$ *for some* $\delta > 1$ *then for all* $\gamma_n \to 0$

$$\sup_{\theta\in\Theta\cap\overline{B}(\theta_0,\varepsilon):\|\theta-\theta_0\|\leq\gamma_n} \left\|c_n''(\theta) - \mathbb{E}\ell_0''(\theta)\right\| = o_p(1).$$

Proof. We have that with P probability one

$$c_n''(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left[\left(\frac{2y_t^2}{h_t(\theta)}-1\right)\frac{h_t'(\theta)\left[h_t'(\theta)\right]^T}{\left(h_t(\theta)\right)^2}+\left(1-\frac{y_t^2}{h_t(\theta)}\right)\frac{h_t''(\theta)}{h_t(\theta)}\right].$$

By Straumann [9] Propositions 5.2.12, 5.5.1 and 5.5.2 together with Proposition 2.1.1, $\left(c_n''\right)$ is a stationary ergodic sequence of random elements with values in $\mathbb{C}\left(\Theta\cap\overline{B}(\theta_0,\varepsilon), \mathbb{R}^{4\times 4}\right)$, with

$$\left\|\ell_0''(\theta)\right\|_{\Theta\cap\overline{B}(\theta_0,\varepsilon)} \leq \left\|\frac{h_0'}{h_0}\right\|_{\Theta\cap\overline{B}(\theta_0,\varepsilon)}^2\left(2\left|\frac{y_0^2}{h_0}\right|_{\Theta\cap\overline{B}(\theta_0,\varepsilon)}+1\right)+\left\|\frac{h_0''}{h_0}\right\|_{\Theta\cap\overline{B}(\theta_0,\varepsilon)}\left(1+\left|\frac{y_0^2}{h_0}\right|_{\Theta\cap\overline{B}(\theta_0,\varepsilon)}\right)$$

By an application of the Hölder inequality together with Lemmas 4.1, 4.2, 4.3 enabled by $\mathbb{E}\left(z_0^2\right)^{\delta} < +\infty$ we have that $\mathbb{E}\left\|\ell_0''(\theta)\right\|_{\Theta\cap\overline{B}(\theta_0,\varepsilon)} < +\infty$ and the fact that $\Theta\cap\overline{B}(\theta_0,\varepsilon)$ is compact. Then by the stationarity-ergodicity of $\left(\ell_t''(\theta)\right)_{t\in\mathbb{Z}}$, we apply Theorem 2.2.1 of Straumann [9] to obtain the result. ■

Lemma 4.6 *Under assumption A.2* $\mathbb{J}=\mathbb{E}\left[\ell_0''(\theta)\right]=\mathbb{E}\left[\frac{h_0'(\theta_0)\left(h_0'(\theta_0)\right)^T}{\sigma_0^4}\right]$ *is a positive definite matrix.*

Proof. Using Straumann's [9] Lemma 5.6.3, we require that the components of the vector $\left(1,-\frac{\gamma^2}{4a^2}+y,y_0,\sigma_0^2\right)'$ are linearly independent random variables. Suppose not, i.e. there exist $\lambda_1,\lambda_2,\lambda_3,\lambda_4\in\mathbb{R}$ where $(\lambda_1,\lambda_2,\lambda_3,\lambda_4)'\neq\mathbf{0}_{4\times1}$ such that:

$$\lambda_1-\frac{\lambda_2\gamma^2}{4a^2}+\lambda_2y_0^2+\lambda_3y_0+\lambda_4\sigma_0^2=0.$$

By Lemma 3.3, the only solution to this equation is the zero vector. ∎

Lemmas enabling the establishment of the asymptotic equivalence between the ergodic and the non-ergodic QMLE**.** In the following the symbol $\overset{e.a.s.}{\to}$ denotes experientially almost sure convergence (w.r.t. P) as defined in Section 2.5 of Straumann [9]. For economy of space we also denote with $\overset{a.s.}{\to}$ almost sure convergence (w.r.t. P).

Lemma 4.7 $\left\|h_t'(\theta)-h_t'^*(\theta)\right\|_{K^*}\overset{e.a.s.}{\to}0.$

Proof. We proceed elementwise. We identify the SRE that each (possibly one sided) partial derivatives satisfy and then use Proposition 5.2.12. of Straumann [9].

- For ω we have that $\frac{\partial h_t}{\partial\omega}=1+\beta\frac{\partial h_{t-1}}{\partial\omega}=\sum_{i=0}^{m}\beta^i+\beta^{m+1}\frac{\partial h_{t-m-1}}{\partial\omega}\overset{m\to\infty}{\to}\frac{1}{1-\beta}$ P a.s., hence the SRE $s_{t+1}=\Phi(s_t)$, $t\in\mathbb{Z}$ for which $[\Phi(s)](\beta)=1+\beta s$, $\log^+\Lambda(\Phi)=\log^+\beta<\infty$ and $\mathbb{E}\log\Lambda(\Phi(s))=\log\beta<0$, $\mathbb{E}\log^+\left\|\Phi(s)(\varsigma_0^2)\right\|_{K^*}=\mathbb{E}\log^+\left\|1-\beta\varsigma_0^2\right\|_{K^*}<\infty$. Then, by Proposition 5.2.12. of Straumann [9] we have that
  $\left\|\frac{\partial h_t^*(\theta)}{\partial\omega}-\frac{\partial h_t(\theta)}{\partial\omega}\right\|_{K^*}\overset{e.a.s.}{\to}0$ and $\frac{\partial h_t}{\partial\omega}=\frac{1}{1-\beta}$ P a.s.

- For a we have that $\frac{\partial h_t}{\partial a}=y_{t-1}^2-\frac{\gamma^2}{4a^2}+\beta\frac{\partial h_{t-1}}{\partial a}=\sum_{i=0}^{\infty}\beta^i\left(-\frac{\gamma^2}{4a^2}+y_{t-i-1}^2\right)=-\sum_{i=0}^{\infty}\beta^i\frac{\gamma^2}{4a^2}+\sum_{i=0}^{\infty}\beta^iy_{t-i-1}^2=-\frac{\gamma^2}{4a^2(1-\beta)}+\sum_{i=0}^{\infty}\beta^iy_{t-i-1}^2$
  P a.s., hence the SRE: $s_{t+1}=\Phi_t(s_t)$, where $[\Phi_t(s)](\beta)=y_{t-1}^2-\frac{\gamma^2}{4a^2}+\beta s$ and $\log^+\Lambda(\Phi_t)=\log^+\beta<\infty$, $\mathbb{E}\log\Lambda(\Phi_t)=\mathbb{E}\log\beta=\log\beta<0$, $\mathbb{E}\log^+\left\|\Phi_t(\varsigma_0^2)\right\|=\mathbb{E}\log^+\left\|y_{t-1}^2-\frac{\gamma^2}{4a^2}-\beta\varsigma_0^2\right\|_{K^*}<\infty$. So, by Proposition 5.2.12. of Straumann [9] we have that
  $\left\|\frac{\partial h_t^*(\theta)}{\partial a}-\frac{\partial h_t(\theta)}{\partial a}\right\|_{K^*}\overset{e.a.s.}{\to}0$ and $\frac{\partial h_t(\theta)}{\partial a}=-\frac{\gamma^2}{4a^2(1-\beta)}+\sum_{i=0}^{\infty}\beta^iy_{t-i-1}^2$ P a.s.

- For γ we have that $\frac{\partial h_t(\theta)}{\partial\gamma}=\frac{\gamma}{2a(1-\beta)}+\sum_{i=0}^{\infty}\beta^iy_{t-i-1}$ P a.s., hence the SRE: $s_{t+1}=\Phi_t(s_t)$, for which $\log^+\Lambda(\Phi_t)=\log^+\beta<\infty$ and $\mathbb{E}\log\Lambda(\Phi_t)=\mathbb{E}\log\beta=\log\beta<0$,

$\mathbb{E}\log^{+}|\Phi_t(\varsigma_0^2)| = \mathbb{E}\log^{+}\left|y_{t-1} - \frac{\gamma}{2a} - \beta\varsigma_0^2\right|_{K^*} \leq \log 2 + \mathbb{E}\log^{+}|y_{t-1}| + \mathbb{E}\log^{+}\left|\frac{\gamma}{2a} + \beta\varsigma_0^2\right|_{K^*} < \infty$. So, by Proposition 5.2.12. of Straumann [9] we have that $\left\|\frac{\partial h_t^*(\theta)}{\partial\gamma} - \frac{\partial h_t(\theta)}{\partial\gamma}\right\|_{K^*} \overset{e.a.s.}{\to} 0$ and $\frac{\partial h_t(\theta)}{\partial\gamma} = \frac{\gamma}{2a(1-\beta)} + \sum_{i=0}^{\infty}\beta^i y_{t-i-1}$ P a.s.

- For β we have that $\frac{\partial h_t(\theta)}{\partial\beta} = h_{t-1}(\theta) + \beta\frac{\partial h_t(\theta)}{\partial\beta^+} = \sum_{i=0}^{\infty}\beta^i h_{t-i-1}(\theta)$ P a.s., hence the SRE: $s_{t+1} = \Phi(s_t)$, for which $\log^{+}\Lambda(\Phi) = \log^{+}\beta < \infty$ and $\mathbb{E}\log\Lambda(\Phi) = \mathbb{E}\log\beta = \log\beta < 0$, $\mathbb{E}\log^{+}|\Phi(\varsigma_0^2)| = \mathbb{E}\log^{+}\left|h_{t-1}^*(\theta) + \beta\varsigma_0^2\right| \leq \log 2 + \mathbb{E}\log^{+}\left|h_{t-1}^*(\theta)\right| + \log^{+}\beta\varsigma_0^2 < \infty$. So, by Proposition 5.2.12. of Straumann [9] we have that $\left\|\frac{\partial h_t^*(\theta)}{\partial\beta} - \frac{\partial h_t(\theta)}{\partial\beta}\right\|_{K^*} \overset{e.a.s.}{\to} 0$ and $\frac{\partial h_t(\theta)}{\partial\beta} = h_{t-1}(\theta) + \beta\frac{\partial h_t(\theta)}{\partial\beta} = \sum_{i=0}^{\infty}\beta^i h_{t-i-1}(\theta)$ a.s. Then using Proposition 2.5.1 of Straumann [9], since $\beta^i \overset{e.a.s.}{\to} 0$ as $i\to\infty$ (choose e.g. $\gamma = \beta^{-\frac{1}{2}} > 1$ to see that $\gamma^i\beta^i \overset{a.s.}{\to} 0$) and y_{t-i-1}^2,y_{t-i-1}, $h_{t-i-1}(\theta)$ are ergodic with $\log^{+}y_{t-i-1}^2 < \infty$, $\log^{+}|y_{t-i-1}| < \infty$, $\log^{+}h_{t-i-1}(\theta) < \infty$ we have that $\sum_{i=0}^{\infty}\beta^i y_{t-i-1}^2$, $\sum_{i=0}^{\infty}\beta^i y_{t-i-1}$, $\sum_{i=0}^{\infty}\beta^i h_{t-i-1}(\theta)$ converge $e.a.s.$ Thus, $\frac{\partial h_t(\theta)}{\partial\theta}$ is well defined.

■

Lemma 4.8 $\|h_t''(\theta) - h_t''^*(\theta)\|_{K^*} \overset{e.a.s.}{\to} 0$

Proof. We proceed elementwise as in the previous proof. For economy of space we present only two cases. The others follow analogously. Consider first the $\frac{\partial^2 h_t(\theta)}{\partial a\partial\omega} = \beta\frac{\partial^2 h_{t-1}(\theta)}{\partial a\partial\omega}$. We have the SRE: $s_{t+1} = \Phi(s_t)$, for which $\log^{+}\Lambda(\Phi) = \log^{+}\beta < \infty$ and $\mathbb{E}\log\Lambda(\Phi) = \mathbb{E}\log\beta = \log\beta < 0$. $\mathbb{E}\log^{+}\|\Phi(\varsigma_0^2)\|_{K^*} = \mathbb{E}\log^{+}\|\beta\varsigma_0^2\|_{K} < \infty$. So, by Proposition 5.2.12. of Straumann we have that $\left\|\frac{\partial^2 h_t^*(\theta)}{\partial a\partial\omega} - \frac{\partial^2 h_t(\theta)}{\partial a\partial\omega}\right\|_{K^*} \overset{e.a.s.}{\to} 0$. Consider finally $\frac{\partial^2 h_t(\theta)}{\partial a\partial\beta}$. We obtain that

$$
\begin{aligned}
\left\|\frac{\partial^2 h_t^*(\theta)}{\partial a\partial\beta} - \frac{\partial^2 h_t(\theta)}{\partial a\partial\beta}\right\|_{K^*} &= \left\|\frac{\partial h_{t-1}^*(\theta)}{\partial a} - \frac{\partial h_{t-1}(\theta)}{\partial a} + \beta\frac{\partial h_{t-1}^*(\theta)}{\partial a\partial\beta^+} - \beta\frac{\partial h_{t-1}(\theta)}{\partial a\partial\beta^+}\right\| \\
&\leq \left\|\frac{\partial h_{t-1}^*(\theta)}{\partial a} - \frac{\partial h_{t-1}(\theta)}{\partial a}\right\| + \beta\left\|\frac{\partial h_{t-1}^*(\theta)}{\partial a\partial\beta^+} - \frac{\partial h_{t-1}(\theta)}{\partial a\partial\beta^+}\right\| \\
&\leq \sum_{i=0}^{\infty}\beta^i\left\|\frac{\partial h_{t-i-1}^*(\theta)}{\partial a} - \frac{\partial h_{t-i-1}(\theta)}{\partial a}\right\| \overset{e.a.s.}{\to} 0.
\end{aligned}
$$

■

Lemma 4.9 $\left\|h_t^{*\prime}(\theta)\,h_t^{*\prime}(\theta)^T - h_t'(\theta)\,h_t'(\theta)^T\right\|_{K^*} \overset{e.a.s.}{\to} 0$

Proof. Trivial. ■

Lemma 4.10 $n\|c_n'^*(\theta) - c_n'(\theta)\|_{K^*}$ *converges* P *a.s.*

Proof. By an application of the mean value theorem to the function $f(a,b) = \frac{a}{b}\left(1 - \frac{y_t^2}{b}\right)$, $a \in \mathbb{R},\ b > 0$, due to remark R.4, we obtain

$$\begin{aligned}\left\|\ell_t^{*\prime}(\theta) - \ell_t'(\theta)\right\|_{K^*} &= \left\|\left(1 - \frac{y_t^2}{h_t^*}\right)\frac{h_t^{*\prime}}{h_t^*} - \left(1 - \frac{y_t^2}{h_t}\right)\frac{h_t'}{h_t}\right\|_{K^*} \\ &\leq c\left(1 + y_t^2\right)\left[\|h_t - h_t^*\|_{K^*} + \|h_t' - h_t^{*\prime}\|_{K^*}\right]\end{aligned}$$

for some $c > 0$. Then, by Proposition 5.2.12 of Straumann [9], $\|h_t - h_t^*\|_K \overset{e.a.s.}{\to} 0$, and by lemma 4.7 $\|h_t' - h_t^{*\prime}\|_K \overset{e.a.s.}{\to} 0$ so $[\|h_t - h_t^*\|_K + \|h_t' - h_t^{*\prime}\|_K] \overset{e.a.s.}{\to} 0$. Then, lemma 3.4 implies that $E\log^+\left[c\left(1 + y_t^2\right)\right] < \infty$ by Lemma 2.5.3 of Straumann [9]. Then

$$n\left\|c_n'^*(\theta) - c_n'(\theta)\right\|_{K^*} \leq \sum_{t=1}^{\infty}\left\|\ell_t^{*\prime}(\theta) - \ell_t'(\theta)\right\|_{K^*} < \infty$$

and the result is obtained by an application of Proposition 2.5.1 of Straumann [9]. ■

Lemma 4.11 $\|c_n^{*\prime\prime}(\theta) - c_n''(\theta)\|_{K^*} \overset{a.s.}{\to} \mathbf{0}$.

Proof. Using the definitions of the second order (possibly one sided) derivatives, the triangle inequality By an application of the mean value theorem to the functions $f(a,b) = \frac{a}{b}\left(1 - \frac{y_t^2}{b}\right)$ and $g(a,b) = \left(\frac{2y_t^2}{a} - 1\right)\frac{b}{a^2}$, we obtain

$$\begin{aligned}&\|\ell_t^{*\prime\prime}(\theta) - \ell_t''(\theta)\|_{K^*} \\ \leq\ & \left\|\left(1 - \frac{y_t^2}{h_t^*}\right)\frac{h_t^{*\prime\prime}}{h_t^*} - \left(1 - \frac{y_t^2}{h_t}\right)\frac{h_t''}{h_t}\right\|_{K^*} \\ &+ \left\|\left(\frac{2y_t^2}{h_t^*} - 1\right)\frac{1}{(h_t^*)^2}h_t^{*\prime}\left(h_t^{*\prime}\right)^T - \left(\frac{2y_t^2}{h_t} - 1\right)\frac{1}{(h_t)^2}h_t'\left(h_t'\right)^T\right\|_{K^*} \\ \leq\ & c_1\left(1 + y_t^2\right)\left[\|h_t - h_t^*\|_{K^*} + \|h_t^{*\prime\prime} - h_t''\|_{K^*}\right] \\ &+ c_2\left(1 + y_t^2\right)\left[\|h_t - h_t^*\|_{K^*} + \left\|h_t^{*\prime}\left(h_t^{*\prime}\right)^T - h_t'\left(h_t'\right)^T\right\|_{K^*}\right]\end{aligned}$$

for some $c_1, c_2 > 0$ which exist due to compactness of K^* and the uniform boundedness of the volatility filters away from zero. Then, by Proposition 5.2.12 of Straumann [9], $\|h_t - h_t^*\|_{K^*} \overset{e.a.s.}{\to} 0$, by Lemma 4.8 $\|h_t'' - h_t^{*\prime\prime}\|_{K^*} \overset{e.a.s.}{\to} 0$ and by Lemma 4.9 $\left\|h_t^{*\prime}\left(h_t^{*\prime}\right)^T - h_t'\left(h_t'\right)^T\right\|_{K^*}$, so analogously to the proof of Lemma 4.10 we obtain

$$n\left\|c_n''(\theta) - c_n^{*\prime\prime}(\theta)\right\|_{K^*} \leq \sum_{t=1}^{\infty}\|\ell_t''(\theta) - \ell_t^{*\prime\prime}(\theta)\|_{K^*} < \infty.$$

■