



HAL
open science

Parametric inference and forecasting in continuously invertible volatility models

Olivier Wintenberger, Sixiang Cai

► **To cite this version:**

Olivier Wintenberger, Sixiang Cai. Parametric inference and forecasting in continuously invertible volatility models. 2011. hal-00597529v4

HAL Id: hal-00597529

<https://hal.science/hal-00597529v4>

Preprint submitted on 27 Jul 2011 (v4), last revised 3 Nov 2011 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PARAMETRIC INFERENCE AND FORECASTING FOR CONTINUOUSLY INVERTIBLE VOLATILITY MODELS

BY OLIVIER WINTENBERGER

Centre De Recherche en Mathématiques de la Décision, UMR CNRS 7534

Université de Paris-Dauphine

UNIVERSITÉ DE PARIS-DAUPHINE

owintenb@ceremade.dauphine.fr

AND

BY SIXIANG CAI

Université de Cergy-Pontoise, Département de Mathématiques, UMR CNRS

8088

UNIVERSITÉ DE CERGY-PONTOISE

sixiang.cai@u-cergy.fr

We introduce the notion of continuous invertibility on a compact set for volatility models driven by a Stochastic Recurrence Equation (SRE). We prove in this context the strong consistency and the asymptotic normality of the estimator based on the SRE given in [Straumann \(2005\)](#). We recover known results on univariate and multivariate GARCH type models where the estimator coincides with the classical QMLE. In EGARCH type models, our approach gives a strongly consistent and asymptotically normal estimator when the limiting covariance matrix exists. We provide a necessary and sufficient condition for the existence of this limiting covariance matrix in the EGARCH(1,1) model.

1. Introduction. Since the seminal paper of [Engle \(1982\)](#) and [Bollerslev \(1986\)](#), the General Autoregressive Conditional Heteroskedasticity (GARCH) type models have been successfully applied to volatility modeling. [Nelson \(1991\)](#) is the first attempt to introduce non linearity into volatility models with the Exponential-GARCH(1,1) type models. Since then, many other volatility models have been introduced: APGARCH of [Ding, Granger and Engle \(1993\)](#), GJR-GARCH of [Glosten, Jagannathan and Runkle \(1993\)](#), TGARCH of [Zakoïan \(1994\)](#), etc. Non linear volatility models have been used extensively in empirical researches (see [Brandt and Jones \(2006\)](#)) among many

AMS 2000 subject classifications: Primary 62F12; secondary 60H25, 62F10, 62M20, 62M10, 91B84

Keywords and phrases: Invertibility, volatility models, parametric estimation, strong consistency, asymptotic normality, asymmetric GARCH, exponential GARCH, stochastic recurrence equation, stationarity

others) and financial industry. Not surprisingly, theoretical investigations of EGARCH has attracted constant attention, see [He, Teräsvirta and Malmsten \(2002\)](#), [Harvey \(2010\)](#) and [Rodriguez and Ruiz \(2009\)](#). However, the validity of the estimation procedures used empirically in [Nelson \(1991\)](#) was not proved. Our study provides the first satisfactory answer to this open question for non linear volatility models including the EGARCH(1,1) model. We provide sufficient conditions for the estimator to be strongly consistent and asymptotically normal.

Consider a general volatility model of the form $X_t = \Sigma_t^{1/2} \cdot Z_t$ where Σ_t is the volatility and where the innovations Z_t are normalized, centered independent identical distributed (iid) random vectors. The natural filtration \mathcal{F}_t is generated by the past innovations (Z_t, Z_{t-1}, \dots) . It is assumed that a transformation of the volatility satisfies some (possibly non-linear) SRE, i.e. there exist a function h and some \mathcal{F}_{t-1} measurable random function ψ_t such that the following relation

$$(1) \quad (h(\Sigma_k))_{k \leq t} = \psi_t((h(\Sigma_k))_{k \leq t-1}, \theta_0)$$

holds. It is the case of all classical models of GARCH and EGARCH types, extensions of the simplest GARCH(1,1) and EGARCH(1,1) univariate models. To make our explanation clearer, let us recall the definitions of these models (in the univariate case, the volatility is denoted σ_t^2):

$$(2) \text{ GARCH(1,1):} \quad \sigma_t^2 = \alpha_0 + \beta_0 \sigma_{t-1}^2 + \gamma_0 X_{t-1}^2,$$

$$(3) \text{ EGARCH(1,1):} \quad \log(\sigma_t^2) = \alpha_0 + \beta_0 \log(\sigma_{t-1}^2) + (\gamma_0 Z_{t-1} + \delta_0 |Z_{t-1}|).$$

One can rewrite (2) as an SRE driven by the innovations: $\sigma_t^2 = \alpha_0 + (\beta_0 + \gamma_0 Z_{t-1}^2) \sigma_{t-1}^2$, i.e. $\psi_t(x, \theta) = \alpha + (\beta + \gamma Z_{t-1}^2)x$. This SRE is used by [Nelson \(1990\)](#) to obtain the Lyapunov condition $\mathbb{E}[\log(\beta_0 + \gamma_0 Z_0^2)] < 0$, necessary and sufficient for the stationarity. In general, the functional process (ψ_t) driving the SRE (1) is assumed to be a stationary ergodic process of Lipschitz functions. Such a SRE is said to be convergent when its solution is unique, non anticipative (i.e. function of \mathcal{F}_t at any time t) and its law does not depend on the initial values. This last property, also called "stability" of the SRE, ensures that the convergence of the SRE driven by (ψ_t) leads to the existence of the stationary process (X_t) . Sufficient conditions (also necessary in the linear case) for the convergence are the strict negativity of the Lyapunov coefficient and the existence of logarithmic moments, see [Bougerol and Picard \(1992\)](#) and [Bougerol \(1993\)](#).

Assume that the model have a non anticipative, stationary solution with invertible volatility matrices Σ_t . Using the relation $Z_t = \Sigma_t^{-1} \cdot X_t$ in ψ_t , it is possible to study a new SRE driven by the observations X_t

$$(4) \quad (h(\Sigma_k))_{k \leq t} = \phi_t((h(\Sigma_k))_{k \leq t-1}, \theta_0).$$

Here ϕ_t is an ergodic stationary process generated by (\mathcal{G}_{t-1}) , the sigma-field of the past values $(X_{t-1}, X_{t-2}, \dots)$. The convergence of this new SRE is closely related with the notion of invertibility (i.e. the existence of a non linear AR representation of the observations) see [Granger and Andersen \(1978\)](#), [Straumann \(2005\)](#) and [Straumann and Mikosch \(2006\)](#). Based on the result of [Bougerol \(1993\)](#), sufficient conditions on Lipschitz functions ϕ_t are the negativity of the Lyapunov coefficient and the existence of logarithmic moments. One easily check on (2) that $\phi_t(x, \theta) = \alpha + \beta x + \gamma X_{t-1}^2$ and the univariate GARCH(1,1) model is invertible as soon as $0 \leq \beta_0 < 1$. For the EGARCH(1,1) model, [Straumann \(2005\)](#) rewrite the SRE such that it is driven by the observations:

$$(5) \quad \log(\sigma_t^2) = \alpha_0 + \beta_0 \log(\sigma_{t-1}^2) + (\gamma_0 X_{t-1} + \delta_0 |X_{t-1}|) \exp(-\log(\sigma_{t-1}^2)/2).$$

A sufficient condition for invertibility follows under restrictions on the parameters $(\alpha_0, \beta_0, \gamma_0, \delta_0)$, see (13) below.

We introduce in this paper the notion of continuous invertibility **(CI)** for models that are invertible on each point θ of a given compact set Θ and a technical smoothness assumption on the function $\theta \mapsto \phi_0(x, \theta)$, see condition **(CI)** below for details. This technical assumption is automatically satisfied for all GARCH and EGARCH type models known by the authors. Thus the GARCH(1,1) model is continuously invertible on all compact sets of $\theta = (\alpha, \beta, \gamma) \in [0, \infty[^3$ satisfying $\beta < 1$ and the EGARCH(1,1) is continuously invertible on all compact sets satisfying (12). Under **(CI)**, one can compute recursively a forecast $\hat{g}_t(\theta)$ of the volatility Σ_t using only the past observations. The convergence of the SRE (4) ensures the stability of this recursive procedure with respect to the initial values chosen arbitrarily. Then one can reasonably study the error of forecasting using the Quasi-LIKelihood (QLIK) criteria:

$$(6) \quad n\hat{S}_n(\theta) = \sum_{t=1}^n \hat{s}_t(\theta) = \sum_{t=1}^n 2^{-1} \left(X_t^T \ell(\hat{g}_t(\theta))^{-1} X_t + \log(\det(\ell(\hat{g}_t(\theta)))) \right).$$

We study the M-estimator $\hat{\theta}_n$ of the unknown parameter θ_0

$$\hat{\theta}_n = \operatorname{argmin}_{\theta \in \Theta} \hat{S}_n(\theta).$$

We prove under **(CI)** and the identifiability of the model that $\hat{\theta}_n$ is a strongly consistent estimator of θ_0 (assumed to belong in Θ and to satisfy the stationary condition) and that the natural forecast $\hat{\xi}_t(\hat{\theta}_t)$ of the volatility Σ_t is also strongly consistent. We prove that $\hat{\theta}_n$ is asymptotically normal if moreover the limiting variance exists.

Continuous invertibility is proved to be a sufficient condition for inferring and forecasting volatility efficiently in GARCH and EGARCH type models. We apply our result to recover existing results in the GARCH case where the inference procedure coincides with the classical Quasi Maximum Likelihood Estimator (QMLE). In the GARCH(1,1) case, we recover the already known result: If Θ is a compact subset of $[0, \infty]^3$ with the constraint $\beta < 1$, the recursive forecast defined by (2) is stable and the inference is strongly consistent if $\theta_0 \in \Theta$ and $\mathbb{E}[\log(\beta_0 + \gamma Z_0^2)] < 0$. Remark that the invertibility problem does not matter since the stationary condition implies the invertibility one. For more general models, we recover the results of [Francq and Zakoian \(2004\)](#) that slightly refine [Berkes, Horvath and Kokoszka \(2003\)](#) for GARCH(p,q) models, we recover the results of [Francq and Zakoian \(2011\)](#) for CCC-GARCH(p,q) models and for AGARCH(p,q) models we refine the results of [Straumann and Mikosch \(2006\)](#). We also apply our results in the EGARCH(1,1) model and give the first proof of the strong consistency and the asymptotic normality of the estimator used in [Nelson \(1991\)](#). The theoretical validity of this commonly used procedure only holds under **(CI)**, i.e. when the SRE (5) is stable with respect to its initial values. On the contrary, it is shown in [Sorokin \(2011\)](#) that forecasting the volatility with SRE may be inconsistent when the SRE is unstable. To sum up, one can think of the following "equivalences":

Convergence of the SRE generated by the innovations (Z_t, Z_{t-1}, \dots)	\iff (A)	stationarity, ergodicity and log-moments
Convergence of the SRE generated by the observations (X_t, X_{t-1}, \dots)	\iff (B)	invertibility, forecasting and statistical inference

The equivalence (A) is crucial when studying existence of stationary solutions of volatility models. We want to emphasize the importance of the second equivalence (B) for the statistical inference and the volatility forecast. The inference procedure described here is the most commonly used in practice and the high generality of our model allows to apply our approach in almost all volatility models. The consequences of this work on empirical study is huge, see [Wintenberger and Cai \(2011\)](#) for the EGARCH(1,1) case (applications on other classical models are also in progress). Finally, notice

that the statistical inference of θ_0 is possible without assuming **(CI)**: it has been done by [Zaffaroni \(2009\)](#) using Whittle's estimator.

An outline of the paper can be given as follows. In Section 2, we discuss the standard notions of invertibility and introduce the continuous invertibility. In Section 3 our main results on the statistical inference based on the SRE are stated. We apply this results in some GARCH type models and in the EGARCH(1,1) model in Section 4. The Appendix contains the technical computation of the necessary and sufficient condition for the existence of the asymptotic variance of $\hat{\theta}_n$ in the EGARCH(1,1) model.

2. Continuously invertible volatility models.

2.1. The general volatility model. In this paper, (Z_t) is a stationary ergodic sequence of real vectors called the innovations. Let us denote \mathcal{F}_t the filtration generated by (Z_t, Z_{t-1}, \dots) . Consider the general volatility model $X_t = \Sigma_t^{1/2} \cdot Z_t$ where (1) is satisfied: $(h(\Sigma_k))_{k \leq t} = \psi_t((h(\Sigma_k))_{k \leq t-1}, \theta_0)$. The function h is injective from the space of real matrices of size $k \times k$ to an auxiliary separable metric space F . The random function $\psi_t(\cdot, \theta_0)$ is a \mathcal{F}_{t-1} adapted random function from the space of the sequences of elements in the image of h to itself. Let us denote ℓ the inverse of h (from the image of h to the space of real matrices of size $k \times k$) and call it the link function.

2.2. Convergent SRE and stationarity. A first question regarding this very general model is whether or not a stationary solution exists. As the sequence of the transformed volatilities $(h(\Sigma_k))_{k \leq t}$ is a solution of a fixed point problem, we recall the following result due to [Bougerol \(1993\)](#). Let (E, d) be a complete separable metric space. A map $f : E \rightarrow E$ is a Lipschitz map if $\Lambda(f) = \sup_{(x,y) \in E^2} d(f(x), f(y)) / d(x, y)$ is finite. For any sequence of random element in (E, d) , (X_t) is said to be exponential almost sure convergence to 0 $X_t \xrightarrow{\text{e.a.s.}} 0$ as $t \rightarrow \infty$ if for $X_t = o(e^{-Ct})$ a.s. for some $C > 0$.

THEOREM 1. *Let (Ψ_t) be a stationary ergodic sequence of Lipschitz maps from E to E . Suppose that $\mathbb{E}[\log^+(d(\Psi_0(x), x))] < \infty$ for some $x \in E$, that $\mathbb{E}[\log^+ \Lambda(\Psi_0)] < \infty$ and that for some integer $r \geq 1$,*

$$\mathbb{E}[\log \Lambda(\Psi_0^{(r)})] = \mathbb{E}[\log \Lambda(\Psi_0 \circ \dots \circ \Psi_{-r+1})] < 0.$$

Then the SRE $X_t = \Psi_t(X_{t-1})$ for all $t \in \mathbb{Z}$ is convergent: it admits a unique stationary solution $(Y_t)_{t \in \mathbb{Z}}$ which is ergodic and for any $y \in E$

$$Y_t = \lim_{m \rightarrow \infty} \Psi_t \circ \dots \circ \Psi_{t-m}(y), \quad t \in \mathbb{Z}.$$

The Y_t are measurable with respect to the $\sigma(\Psi_{t-k}, k \geq 0)$ and

$$d(\tilde{Y}_t, Y_t) \xrightarrow{e.a.s.} 0, \quad t \rightarrow \infty$$

such that $\tilde{Y}_t = \Psi_t(\tilde{Y}_{t-1})$ for all $t > 0$.

The sufficient Lyapunov assumptions $\mathbb{E}[\log \Lambda(\Psi_0^{(r)})] < 0$ is also necessary in the linear case, see [Bougerol and Picard \(1992\)](#). The logarithmic moments of the solution of a convergent SRE is proved in the following result that seems to be new:

THEOREM 2. *Under the assumptions of Theorem (1) and $\mathbb{E}[(\log^+ d(\Psi_0(x), x))^2] < \infty$ the unique stationary solution satisfies $\mathbb{E}[\log^+(d(Y_0, y))] < \infty$ for all $y \in E$.*

PROOF. The basic inequality $\log(1 + y + z) \leq \log(1 + y) + \log(1 + z)$ will be used several time. Remark also that for any r.v. $X \geq 0$ we have the equivalence $\mathbb{E}[\log(1 + X)] < \infty$ iff $\mathbb{E}[\log^+(X)] < \infty$. Thus $\mathbb{E}[\log(1 + d(\Psi_0(y), y))] < \infty$ for all $y \in E$ as $d(\Psi_0(y), y) \leq d(\Psi_0(x), x) + \Lambda(\Psi_0)d(x, y)$. For any $y \in E$, one denotes $\Psi^{(-m)} = \Psi_0 \circ \dots \circ \Psi_{-m}(y)$, $w = d(y, \Psi^{(1-m)}(y)) \geq 0$ and $z = \Lambda(\Psi^{(1-m)})d(\Psi_{-m}(y), y) \geq 0$. From the triangular inequality one obtains $d(y, \Psi^{(-m)}(y)) \leq w + z$ and using we derive that

$$\log(1 + d(y, \Psi^{(-m)}(y))) \leq \log(1 + w + z) \leq \log(1 + w) + \log(1 + z).$$

Eqn. 27 in [Bougerol and Picard \(1992\)](#) asserts the existence of $0 < \rho < 1$ and $\epsilon > 0$ satisfying

$$\overline{\lim}_{m \rightarrow \infty} \frac{1}{m} \log(\Lambda(\Psi^{(-m)})) \leq \log(\rho) - \epsilon \quad a.s.$$

Thus $\Lambda(\Psi^{(-m)}) \leq \rho^m$ a.s. for all m larger than some fixed $M > 0$. Writing $v_m = \log(1 + d(y, \Psi^{(-m)}(y)))$, for all $m \geq M$ we have:

$$v_m \leq v_{m-1} + \log(1 + \rho^{m-1}d(\Psi_{-m}(y), y)).$$

A straightforward recurrence leads to the following upper bound of all $(v_m)_{m \geq M}$

$$v_m \leq v_M + \sum_{j=1}^{\infty} \log(1 + \rho^{j-1}d(\Psi_{-j-M}(y), y)).$$

As $\log(1 + d(y, Y_0)) = \lim_{m \rightarrow \infty} v_m$ a.s., it remains to prove that the upper bound is integrable to conclude by the dominated integration Theorem. Using the stationarity of (v_m) , we know that $\mathbb{E}[\log(1 + \rho^{j-1}d(\Psi_{-j-M}(y), y))] =$

$\mathbb{E}[\log(1 + \rho^{j-1}d(\Psi_0(y), y))]$ for all $m \geq M$. Thus $\mathbb{E}[v_M] = \mathbb{E}[\log(1 + d(y, \Psi_0(y)))] < \infty$ by assumption. We conclude the proof by comparing the series with an integral:

$$\sum_{j \geq 1} \mathbb{E}[\log(1 + \rho^{j-1}d(\Psi_0(y), y))] \leq \frac{1}{1 - \rho} \int_0^1 \frac{\mathbb{E}[\log(1 + ud(\Psi_0(y), y))]}{u} du.$$

Let us prove that his integral converges as soon as $\mathbb{E}[(\log^+ \Lambda(\Psi_0))^2] < \infty$. Using that $\mathbb{E}[\log(1 + ud(\Psi_0(y), y))] = \int_0^\infty \mathbb{P}(\log(1 + ud(\Psi_0(y), y)) \geq t) dt$ and denoting $v = (e^t - 1)/u$ the integral becomes:

$$\int_0^1 \int_0^\infty \frac{\mathbb{P}(\log(1 + ud(\Psi_0(y), y)) \geq t)}{u} dt du = \int_0^1 \int_0^\infty \frac{\mathbb{P}(d(\Psi_0(y), y) \geq v)}{1 + uv} dv du.$$

Using Fubini's theorem and $\int_0^1 (1 + uv)^{-1} du = \log(v + 1)/v$ for all $v \geq 0$ we get an upper bound in term of

$$\int_0^\infty \frac{\log(1 + v)\mathbb{P}(d(\Psi_0(y), y) \geq v)}{v} dv.$$

This integral converges in $+\infty$ as

$$\int_0^\infty \frac{\log(1 + v)\mathbb{P}(d(\Psi_0(y), y) \geq v)}{1 + v} dv = \mathbb{E}[\log(1 + d(y, \Psi(y)))]$$

and the desired result follows. \square

In order to apply Theorem 1 in our case, let us denote by E the separable metric space of the sequences of elements in the image of h . Equipped with the metric $\sum_{j \geq 1} 2^{-j} d(x_j, y_j) / (1 + d(x_j, y_j))$, the space E is complete. A sufficient condition for stationarity of (X_t) is that the SRE driven by (ψ_t) converges in E . It simply expresses as the Lyapunov condition $\mathbb{E}[\log \Lambda(\psi_0^{(r)})] < 0$ for some integer $r \geq 1$ and some logarithmic moments. This assumption of stationarity is sufficient but not optimal in many cases:

REMARK 1. *The state space of the SRE (1), denoted E , in its most general form, is a space of infinite sequences. However in all classical models we can find a lag p such that $(h(\Sigma_k))_{t-p+1 \leq k \leq t} = \psi_t((h(\Sigma_k))_{t-p \leq k \leq t-1}, \theta_0)$. The state space E is now the finite product of p spaces. It can be equipped by unbounded metrics such that $p^{-1} \sum_{j=1}^p d(x_j, y_j)$ or $\sqrt{\sum_{j=1}^p d^2(x_j, y_j)}$. The product metric has to be carefully chosen as it changes the value of the Lipschitz coefficients of the ϕ_t . Yet, even if the products spaces are embedded, the smallest possible lag p in the SRE*

yields the sharpest Lyapunov condition. Finally, if E has a finite dimension and if the condition of convergence of the SRE expresses in term of the top Lyapunov coefficient, one can choose any metric induced by any norm, see [Bougerol \(1993\)](#) for details.

In view of Remark 1, instead of choosing a specific metric space (E, d) we prefer to work under the less explicit assumption

(ST) The process (X_t) satisfying (1) exists. It is a stationary, non anticipative and ergodic process with finite logarithmic moments.

In view of Theorem 2, it is reasonable to require that the solution has finite logarithmic moments. It is very useful when considering the invertibility of the general model, see Theorem 2 below.

2.3. *The invertibility and the observable invertibility.* Now that under **(ST)** the process (X_t) is stationary and ergodic, we investigate the question of invertibility of the general model (1). We want to emphasize that the notions of invertibility is linked with convergences of SREs, governed by Lyapunov conditions. Following [Tong \(1993\)](#), we say that a volatility model is invertible if the volatility can be expressed as a function of the past observed values:

DEFINITION 1. *The model is invertible if the sequence of the volatilities (Σ_t) is adapted to the filtration (\mathcal{G}_{t-1}) generated by $(X_{t-1}, X_{t-2}, \dots)$.*

It is natural to assume invertibility to be able to forecast the volatility. This notion of invertibility is very weak and consists in restricting the underlying filtration (\mathcal{F}_t) of the SRE to (\mathcal{G}_{t-1}) . Indeed, under **(ST)** then $\mathcal{G}_t \subseteq \mathcal{F}_t$ is well defined. If the volatility matrices are invertible, using $Z_t = \Sigma_t^{-1} \cdot X_t$ in ψ_t we can express (1) as (4): $(h(\Sigma_k))_{k \leq t} = \phi_t((h(\Sigma_k))_{k \leq t-1}, \theta_0)$, a SRE driven by the whole past of the observations. The sequence of random functions (ϕ_t) is an ergodic and stationary process adapted to (\mathcal{G}_{t-1}) . Using the sufficient conditions of convergence of SREs given in Theorem 1, the invertibility follows if the $\phi_t(\cdot, \theta_0)$ are Lipschitz maps such that for some $x \in E$ and $r > 0$,

$$(7) \quad \begin{aligned} \mathbb{E}[\log^+(d(x, \phi_0(x, \theta_0)))] < \infty, \quad \mathbb{E}[\log^+ \Lambda(\phi_0(\cdot, \theta_0))] < \infty \\ \text{and} \quad \mathbb{E}[\log \Lambda(\phi_0(\cdot, \theta_0)^{(r)})] < 0. \end{aligned}$$

The Remark 1 also holds for the SRE driven by (ϕ_t) : the metric space (E, d) must be chosen carefully. The conditions (7) (with the optimal metric space (E, d)) are called the conditions of invertibility.

PROPOSITION 1. *Under (ST) and (7), the general model (4) is invertible.*

Another notion of invertibility is introduced in [Straumann and Mikosch \(2006\)](#). We call it observable invertibility. Let us assume that there exists some approximations $\hat{\phi}_t$ of ϕ_t such that $\hat{\phi}_t$ is a measurable function of the past observations $(X_{t-1}, X_{t-2}, \dots, X_1)$ and not the whole past $(X_{t-1}, X_{t-2}, \dots)$.

DEFINITION 2. *The model is observably invertible if and only if the solution of the approximative SRE*

$$(8) \quad (h(\hat{\Sigma}_k))_{k \leq t} = \hat{\phi}_t((h(\hat{\Sigma}_k))_{k \leq t-1}, \theta_0)$$

is convergent, i.e. $\|\hat{\Sigma}_t - \Sigma_t\| \rightarrow 0$ in probability as $t \rightarrow \infty$.

Remark that in general the approximative SRE does not fit the conditions of Theorem 1 and in particular $(\hat{\phi}_t)$ is not necessarily stationary and ergodic. However, the Proposition below gives sufficient conditions for observable invertibility. It is a very useful result for the sequel of the paper, see Remark 2. Remark that the logarithmic moments are needed in the proof and that the only known effective way of checking it is given in our Theorem 2.

PROPOSITION 2. *If (ST) and (7) hold, if the link function ℓ is continuous and it exists $x \in E$ such that $d(\hat{\phi}_t(x), \phi_t(x)) \xrightarrow{\text{e.a.s.}} 0$ and $\Lambda(\hat{\phi}_t(\cdot, \theta_0) - \phi_t(\cdot, \theta_0)) \xrightarrow{\text{e.a.s.}} 0$ as $t \rightarrow \infty$, then the model is observably invertible.*

PROOF. One can extend the proof of Theorem 2.10 in [Straumann and Mikosch \(2006\)](#) written for Banach spaces to the case of the complete separable metric space (E, d) . That $d((h(\Sigma_k))_{k \leq t}, (h(\hat{\Sigma}_k))_{k \leq t}) \xrightarrow{\text{e.a.s.}} 0$ follows from the proof of Theorem 2.10 in [Straumann and Mikosch \(2006\)](#) under the following assumptions:

- S1** $\mathbb{E}[\log^+(d(x, \phi_0(x, \theta_0)))] < \infty$ for some $x \in E$,
- S2** $\mathbb{E}[\log \Lambda(\phi_0(\cdot, \theta_0))] < \infty$ and $\mathbb{E}[\log \Lambda(\phi_0(\cdot, \theta_0)^{(r)})] < 0$ for some $r > 0$,
- S2'** $\mathbb{E}[\log^+(d(y, (h(\Sigma_k))_{k \leq t}))] < \infty$ for all $y \in E$,
- S3** $d(\hat{\phi}_t(x), \phi_t(x)) \xrightarrow{\text{e.a.s.}} 0$ and $\Lambda(\hat{\phi}_t(\cdot, \theta_0) - \phi_t(\cdot, \theta_0)) \xrightarrow{\text{e.a.s.}} 0$ as $t \rightarrow \infty$.

The conditions **S1-S2** are equivalent to the invertibility conditions (7). **S3** holds from the assumptions in Proposition 2 and **S2'** follows from (ST). Finally, using the continuity of the projection on the first coordinate and the one of the link function ℓ , the desired result follows. \square

REMARK 2. *Classical models such that GARCH(p, q) or EGARCH(p, q) models satisfy an SRE for finite p lags $(h(\Sigma_k))_{t-p+1 \leq k \leq t} = \phi_t((h(\Sigma_k))_{t-p \leq k \leq t-1}, \theta_0)$ and for some ϕ_t generated by only a finite of past observation $(X_{t-1}, \dots, X_{t-q})$. In this context, the approximative SRE coincides with the initial ones, i.e. one can choose $\hat{\phi}_t = \phi_t$ for $t > q$. Therefore, conditions of Proposition 2 hold systematically; invertibility and observable invertibility are equivalent, i.e. they are induced by the same Lyapunov condition. As for any initial values of $\hat{\phi}_t$ (for $0 \leq t \leq q$) the conditions of Proposition 2 are satisfied, seeking simplicity we work in the sequel with $\hat{\phi}_t = \phi_t$ for all $t \geq 1$.*

2.4. *The continuous invertibility.* We have seen that the existing invertibility notions can be expressed in term of Lyapunov conditions. We introduce the notion of continuous invertibility in term of a Lyapunov condition and some smoothness assumption. Let us consider models with parametric functions having continuous Lipschitz coefficients:

(CL) For any metric spaces \mathcal{X}, \mathcal{Y} and \mathcal{Z} , a function $f : \mathcal{X} \times \mathcal{Y} \mapsto \mathcal{Z}$ satisfies (CL) if there exists a continuous function $\Lambda_f : \mathcal{Y} \mapsto \mathbb{R}^+$ such that $\Lambda(f(\cdot, y)) \leq \Lambda_f(y)$ for all $y \in \mathcal{Y}$.

As we want to infer the unknown parameter θ_0 , let us consider from Remark 2 the functional SRE of the form

$$(9) \quad (\hat{g}_k(\theta))_{t-p+1 \leq k \leq t} = \phi_t((\hat{g}_k(\theta))_{t-p \leq k \leq t-1}, \theta), \quad \forall t \geq 1,$$

on some θ with arbitrary initial values $(\hat{g}_k(\theta))_{1-p \leq k \leq 0}$. We introduce the condition of continuous invertibility on a compact set Θ where the SRE (9) is stable:

(CI) Assume that the SRE (9) holds with ϕ_t satisfying (CL) for stationary (Λ_{ϕ_t}) such that conditions $\mathbb{E}[\log \Lambda_{\phi_0}^{(r)}(\theta)] < 0$ on the compact set Θ . Assume moreover that $\mathbb{E}[\sup_{\Theta} \log^+ \Lambda_{\phi_0}^{(r)}(\theta)] < \infty$ and that there exists $y \in E$ such that $\mathbb{E}[\sup_{\Theta} \log^+(d(\phi_0(y, \theta), y))] < \infty$.

The condition (CI) implies the invertibility and the observable invertibility studied in Subsection 2.3 for all $\theta \in \Theta$. Then the law of the volatility approximation $\hat{g}_t(\theta)$ forgets exponentially fast their arbitrary initial values. It also implies the local uniform regularity of the solution $g_t(\cdot)$ of the functional SRE $(g_k(\cdot))_{t-p+1 \leq k \leq t} = \phi_t((g_k(\cdot))_{t-p \leq k \leq t-1}, \cdot)$ for all $t \in \mathbb{Z}$. On the contrary of Straumann (2005), the uniform continuity is not needed here. Our conditions are weaker than in Straumann (2005) and Straumann and Mikosch (2006).

THEOREM 3. *Assume that **(ST)** and **(CI)** hold. Then the functions $g_t(\cdot)$ are continuous for all $\theta \in \Theta$ and all $t \in \mathbb{Z}$. Moreover, for any $\theta \in \Theta$ there exists an $\epsilon > 0$ such that $\hat{g}_t(\theta)$ satisfying (9) satisfies*

$$(10) \quad \lim_{\rho \rightarrow 0} \sup_{\theta' \in \bar{B}(\theta, \epsilon) \cap \Theta} d(\hat{g}_t(\theta'), g_t(\theta')) \xrightarrow{e.a.s.} 0.$$

PROOF. For any $\rho > 0$, let us write $\Lambda_*^{(r)}(\theta, \rho) = \sup\{\Lambda_{\phi_0}^{(r)}(\theta'), \theta' \in \bar{B}(\theta, \rho) \cap \Theta\}$, where $\bar{B}(\theta, \rho)$ stands for the closed ball centered at θ with radius ρ . As $\mathbb{E}[\sup_{\Theta} \log \Lambda_{\phi_0}^{(r)}(\theta)] < \infty$, by the dominated convergence Theorem we obtain $\lim_{\rho \rightarrow 0} \mathbb{E}(\Lambda_*^{(r)}(\theta, \rho)) = \mathbb{E}(\lim_{\rho \rightarrow 0} \Lambda_*^{(r)}(\theta, \rho))$. But $\lim_{\rho \rightarrow 0} \Lambda_*^{(r)}(\theta, \rho) = \Lambda_{\phi_0}^{(r)}(\theta)$ by continuity of $\Lambda_{\phi_0}^{(r)}(\theta)$, we finally obtain

$$\lim_{\rho \rightarrow 0} \mathbb{E}(\Lambda_*^{(r)}(\theta, \rho)) = \mathbb{E}(\lim_{\rho \rightarrow 0} \sup_{\theta' \in \bar{B}(\theta, \rho) \cap \Theta} \Lambda_{\phi_0}^{(r)}(\theta')) = \mathbb{E}(\Lambda_{\phi_0}^{(r)}(\theta)) < \infty.$$

Thus, there exists an $\epsilon > 0$ such that $\mathbb{E}(\Lambda_*^{(r)}(\theta, \epsilon)) < \infty$.

Let us now work on $C(\bar{B}(\theta, \epsilon) \cap \Theta)$, the complete metric space of continuous functions from $\bar{B}(\theta, \epsilon) \cap \Theta$ to \mathbb{R} equipped with the supremum norm $d_{\infty} = \sup_{\bar{B}(\theta, \epsilon) \cap \Theta} d$. In this setting (\hat{g}_t) satisfy a functional SRE $(\hat{g}_k)_{k \leq t} = \phi_t((\hat{g}_k)_{k \leq t-1})$ with Lipschitz constants satisfying

$$\begin{aligned} \Lambda_{\infty}(\phi_t^{(r)}(\cdot)) &\leq \sup_{s_1, s_2 \in C(\bar{B}(\theta, \epsilon) \cap \Theta)} \frac{d_{\infty}(\phi_t^{(r)}(s_1), \phi_t^{(r)}(s_2))}{d_{\infty}(s_1, s_2)} \\ &\leq \sup_{s_1, s_2 \in C(\bar{B}(\theta, \epsilon) \cap \Theta)} \frac{\sup_{\bar{B}(\theta, \epsilon) \cap \Theta} d(\phi_t^{(r)}(s_1(\theta'), \theta'), \phi_t^{(r)}(s_2(\theta'), \theta'))}{d_{\infty}(s_1, s_2)} \\ &\leq \sup_{s_1, s_2 \in C(\bar{B}(\theta, \epsilon) \cap \Theta)} \frac{\sup_{\bar{B}(\theta, \epsilon) \cap \Theta} \Lambda(\phi_t^{(r)}(\cdot, \theta')) d(s_1(\theta'), s_2(\theta'))}{d_{\infty}(s_1, s_2)} \\ &\leq \sup_{s_1, s_2 \in C(\bar{B}(\theta, \epsilon) \cap \Theta)} \frac{\sup_{\bar{B}(\theta, \epsilon) \cap \Theta} \Lambda(\phi_t^{(r)}(\cdot, \theta')) d_{\infty}(s_1, s_2)}{d_{\infty}(s_1, s_2)} \\ &\leq \sup_{\bar{B}(\theta, \epsilon) \cap \Theta} \Lambda(\phi_t^{(r)}(\cdot, \theta')) \leq \sup_{\bar{B}(\theta, \epsilon) \cap \Theta} \Lambda_{\phi_t}^{(r)}(\theta') \leq \Lambda_*^{(r)}(\theta, \epsilon). \end{aligned}$$

As $\mathbb{E}[\sup_{\bar{B}(\theta, \epsilon) \cap \Theta} \log^+(d(\phi_t(y, \theta'), y))] \leq \mathbb{E}[\sup_{\Theta} \log^+(d(\phi_t(y, \theta), y))]$ is finite we can apply Theorem 1. By recurrence $\phi_t \circ \dots \circ \phi_{t-m}(\zeta_0) \in C(\bar{B}(\theta, \epsilon) \cap \Theta)$ is continuous in θ and so is g_t as the convergence holds uniformly on $\bar{B}(\theta, \epsilon) \cap \Theta$. It is true for any $\theta \in \Theta$ and the result follows. \square

3. Statistical inference under continuous invertibility.

3.1. *Statistical inference based on the SRE.* Here we describe the approach in [Straumann \(2005\)](#). Assume that (9) holds with θ_0 unknown and θ_0 belongs in the compact set Θ of stability, i.e. such that **(CI)** holds. Consider

$$\hat{\theta}_n = \operatorname{argmin}_{\theta \in \Theta} \hat{S}_n(\theta)$$

the M-estimator associated with the QLIK criteria (6) where (\hat{g}_t) is obtained from the approximative SRE (9).

REMARK 3. *This statistical procedure does not coincide with the Quasi Maximum Likelihood as it is not always efficient when the innovations are normally distributed. For a detailed discussion in the EGARCH(1,1) case see [Wintenberger and Cai \(2011\)](#).*

3.2. *Strong consistency of the parametric inference.* From now on, we assume that the innovations process (Z_t) is iid:

(IN) The Z_t are iid variables such that $\mathbb{E}[Z_0^T Z_0]$ is the identity matrix.

The next assumption implies that the volatility matrices are invertible and that the link function ℓ is continuous:

(IV) The functions ℓ^{-1} and $\log(\det(\ell))$ are Lipschitz satisfying $\det(\ell(g_0(\theta))) \geq C(\theta)$ for some continuous function $C : \Theta \mapsto (0, \infty)$.

REMARK 4. *The SRE criteria converges to the possibly degenerate limit*

$$S(\theta) = \mathbb{E}[s_0(\theta)] = 2^{-1} \mathbb{E} \left[X_0^T \ell(g_0(\theta))^{-1} X_0 + \log(\det[\ell(g_0(\theta))]) \right]$$

*Remark that $S(\theta_0) = 2^{-1} \mathbb{E}[Z_0^T Z_0 + \log(\det(\Sigma_0))]$ is finite under **(ST)**, **(IN)** and **(IV)** because $h(\Sigma_0)$ has logarithmic moments and $\log(\det(\ell))$ is Lipschitz. It is a considerable advantage of the QLIK criteria: it does not need moments of any order to be defined in θ_0 . Even if $S(\theta)$ may be equal to $+\infty$ for $\theta \neq \theta_0$, it does not interfere the statistical procedure by definition of $\hat{\theta}_n$ as a minimizer.*

If the model is identifiable, the estimator $\hat{\theta}_n$ is strongly consistent:

THEOREM 4. *Assume that **(ST)** and **(CI)** are satisfied on the compact set Θ . If **(IN)** and **(IV)** are satisfied and the model is identifiable, i.e. $g_0(\theta) = h(\Sigma_0)$ iff $\theta = \theta_0$, then $\hat{\theta}_n \rightarrow \theta_0$ a.s. for any $\theta_0 \in \Theta$.*

PROOF. First, remark that with no loss of generality we can always restrict (9) to θ satisfying $\det(\ell(\hat{\phi}_t(\cdot, \theta))) \geq C(\theta)$. We adapt the proof of [Jeantheau \(1993\)](#) and its notation $s_{*t}(\theta, \rho) = \inf\{s_t(\theta'), \theta' \in \bar{B}(\theta, \rho)\}$ and $\hat{s}_{*t}(\theta, \rho) = \inf\{\hat{s}_t(\theta'), \theta' \in \bar{B}(\theta, \rho)\}$. Let us recall Theorem 5.1 in [Jeantheau \(1993\)](#) : The M-estimator associated with the loss (6) is strongly consistent under the hypothesis **H1-H6**:

H1 Θ is compact.

H2 $\hat{S}_n(\theta) \rightarrow S(\theta)$ a.s. under the stationary law P_{θ_0} .

H3 $S(\theta)$ admits a unique minimum for $\theta = \theta_0$ in Θ . Moreover for any $\theta_1 \neq \theta_0$ we have:

$$\liminf_{\theta \rightarrow \theta_1} S(\theta) > S(\theta_1).$$

H4 $\forall \theta \in \Theta$ and sufficiently small $\rho > 0$ the process $(\hat{s}_{*t}(\theta, \rho))_t$ is ergodic.

H5 $\forall \theta \in \Theta, \mathbb{E}_{\theta_0}[s_{*1}(\theta, \rho)] > -\infty$.

H6 $\lim_{\rho \rightarrow 0} \mathbb{E}_{\theta_0}[s_{*1}(\theta, \rho)] = \mathbb{E}[s_{*1}(\theta)]$.

Let us check **H1-H6** in our case. **H1** is satisfied by assumption. **H2** is verified in two steps. First, by the e.a.s. convergence given by Theorem 3, arguments of [Straumann \(2005\)](#) and the Lipschitz properties of ℓ^{-1} and $\log(\det(\ell))$ we obtain

$$\frac{1}{n} \sum_{t=1}^n \sup_{\bar{B}(\theta, \varepsilon)} |\hat{s}_t(\theta') - s_t(\theta')| \rightarrow 0 \quad P_{\theta_0} - a.s.$$

Second we use that (s_t) is an ergodic sequence. Using Proposition 1.1 of [Jeantheau \(1993\)](#), $n^{-1} \sum_{t=1}^n s_t(\theta)$ converges P_{θ_0} -a.s. to $S(\theta)$ (taking values in $\mathbb{R} \cup \{+\infty\}$) as the s_t are bounded from below:

$$\frac{1}{n} \sum_{t=1}^n |s_t(\theta) - S(\theta)| \rightarrow 0 \quad P_{\theta_0} - a.s.$$

Combining this two steps leads to **H2**. The first part of **H3** is checked similarly than in (ii) p.2474 of [Straumann and Mikosch \(2006\)](#) and with the help of the Remark 4. Remark that S has a unique minimum iff $\mathbb{E}[\text{Tr}(\Sigma_0 \cdot \ell(g_0(\theta))^{-1}) - \log(\det(\Sigma_0 \cdot \ell(g_0(\theta))^{-1}))]$ has a unique minimum. As this criteria is the integrand of a sum of the $\lambda_i - \log(\lambda_i)$ where the λ_i are positive eigenvalues, we conclude under the identifiability condition from the property $x - \log(x) \geq 1$ for all $x > 0$ with equality iff $x = 1$. The second part is checked using the fact that

$$\liminf_{\theta \rightarrow \theta_1} S(\theta) \geq \mathbb{E}[\liminf_{\theta \rightarrow \theta_1} s_0(\theta)] = \mathbb{E}[s_0(\theta_1)] = S(\theta_1)$$

where the first inequality was already used for proving Theorem 3 and the first equality comes from the local continuity of g_0 and ℓ . **H4** is satisfied

from the ergodicity of (\hat{s}_t) . **H5** and **H6** follows from Theorem 3 that ensures the continuity of the function s_{*1} and by the lower bounded assumption on $\det(\ell)$, see Proposition 1.3 of [Jeantheau \(1993\)](#). \square

3.3. *Volatility forecasting.* From the inference of θ_0 , we deduce a natural forecast of the volatility $\hat{\Sigma}_t = \ell(\hat{g}_t(\hat{\theta}_t))$. It is strongly consistent:

THEOREM 5. *Under the conditions of Theorem 4 then $\|\hat{\Sigma}_t - \Sigma_t\| \rightarrow 0$ a.s. as $t \rightarrow \infty$.*

PROOF. It is a direct consequence of Theorems 3 and 4 that assert the a.s. convergence of $\hat{\theta}_t$ toward θ_0 and the local uniform convergence of \hat{g}_t toward g_t . Remark that for t sufficiently large such that $\hat{\theta}_t \in \bar{B}(\theta, \epsilon)$, a ball where the uniform Lyapunov condition $\mathbb{E}[\log \Lambda_\infty(\phi_t(\cdot))]$ < 0 is satisfied. Thus $\hat{g}_t(\hat{\theta}_t) - g_t(\theta_0) \rightarrow 0$ a.s. and by continuity of ℓ and g_t and from the identification $\Sigma_t = \ell(g_t(\theta_0))$ the result follows if $g_t(\hat{\theta}_t)$ converges to $g_t(\theta_0)$. For proving it, we use

$$d(g_t(\hat{\theta}_t), g_t(\theta_0)) \leq \Lambda_\infty(\phi_t(\cdot))d(g_{t-1}(\hat{\theta}_t), g_{t-1}(\theta_0)) + w_t(\hat{\theta}_t)$$

where $w_t(\hat{\theta}_t) = d(\phi_t(g_{t-1}(\theta_0), \hat{\theta}_t), \phi_t(g_{t-1}(\theta_0), \theta_0))$. The RHS term satisfies an SRE of linear stationary maps satisfying the Lyapunov condition. We apply Theorem 1 as for any $\hat{\theta}_t$, by assumption $\mathbb{E} \log^+(w_0(\hat{\theta}_t))$ is uniformly bounded by $\mathbb{E}[\sup_{\Theta} \log^+(2d(\phi_t(y, \theta), y))] < \infty$. We get

$$d(g_t(\hat{\theta}_t), g_t(\theta_0)) \leq \sum_{i=0}^{\infty} \Lambda_\infty(\phi_t(\cdot)) \cdots \Lambda_\infty(\phi_{t-i+1}(\cdot)) w_{t-i}(\hat{\theta}_t).$$

Conditioning on $(\hat{\theta}_t)$, the upper bound is a stationary normally convergent series of functions and

$$\begin{aligned} & \mathbb{P}\left(\sum_{i=0}^{\infty} \Lambda_\infty(\phi_t(\cdot)) \cdots \Lambda_\infty(\phi_{t-i+1}(\cdot)) w_{t-i}(\hat{\theta}_t) \rightarrow 0\right) \\ &= \mathbb{E}\left[\mathbb{P}\left(\sum_{i=0}^{\infty} \Lambda_\infty(\phi_t(\cdot)) \cdots \Lambda_\infty(\phi_{t-i+1}(\cdot)) w_{t-i}(\hat{\theta}_t) \rightarrow 0 \mid (\hat{\theta}_t)\right)\right] \\ &= \mathbb{E}\left[\mathbb{P}\left(\sum_{i=0}^{\infty} \Lambda_\infty(\phi_0(\cdot)) \cdots \Lambda_\infty(\phi_{-i+1}(\cdot)) w_{-i}(\hat{\theta}_t) \rightarrow 0 \mid (\hat{\theta}_t)\right)\right] \\ &= \mathbb{E}[1] = 1, \end{aligned}$$

the last inequalities following from the continuity of normally convergent series of functions, $\hat{\theta}_t \rightarrow \theta_0$ and $w_i(\hat{\theta}_t) \rightarrow w_i(\theta_0) = 0$ for all i a.s. as $t \rightarrow \infty$. \square

3.4. *Asymptotic normality of the parametric inference.* Classical computations show that if the M-estimator $\hat{\theta}_n$ is asymptotically normal then the asymptotic variance is given by the expression

$$\mathbf{V} = \mathbf{P}^{-1} \mathbf{Q} \mathbf{P}^{-1}$$

with $\mathbf{P} = \mathbb{E}[\mathbb{H}s_0(\theta_0)]$ and $\mathbf{Q} = \mathbb{E}[\nabla s_0(\theta_0) \nabla s_0(\theta_0)^T]$, where $\mathbb{H}s_0(\theta_0)$ is the Hessian matrix of $s_0(\theta_0)$.

(AV) Assume that $\mathbb{E}(\|Z_0 Z_0^T\|^2) < \infty$ and that the functions ℓ and ϕ_t are 2-times continuously differentiable on the compact set Θ that coincides with the closure of its interior.

The following moments assumptions ensure the existence of \mathbf{Q} and \mathbf{P} :

(MM) Assume that $\mathbb{E}[\|\nabla s_0(\theta_0)\|^2] < \infty$ and $\mathbb{E}[\|\mathbb{H}s_0(\theta_0)\|] < \infty$.

These moments assumptions holds only for $\theta = \theta_0$; they are simpler to verify than for the moment conditions for $\theta \neq \theta_0$ due to the specific form of the derivatives of the SRE criteria, see Remark 4 and computations in Bardet and Wintenberger (2009). The next assumption is classical and ensures to the existence of \mathbf{P}^{-1} :

(LI) The components of the vector $\nabla g_0(\theta_0)$ are linearly independent.

Let $\mathcal{V} = \bar{B}(\theta_0, \epsilon) \subset \Theta$ with $\theta_0 \in \overset{\circ}{\Theta}$ and $\epsilon > 0$ chosen in accordance with Theorem 3, i.e. such that $\mathbb{E}[\log(\sup_{\mathcal{V}} \Lambda_{\phi_0})] < 0$. The two next assumptions are specific to the SRE approach. They ensure that $\nabla \hat{s}_t(\theta)$ is a good approximation of $\nabla s_t(\theta)$ uniformly on the neighborhood \mathcal{V} of θ_0 :

(DL) The partial derivatives $\Phi_t = D_x(\phi_t), = D_\theta(\phi_t), = D_{x^2}^2(\phi_0), D_{\theta,x}^2(\phi_0)$ or $D_{\theta^2}^2(\phi_0)$ satisfy **(CL)** for stationary (Λ_{Φ_t}) with $\mathbb{E}[\sup_{\mathcal{V}} \log(\Lambda_{\Phi_0})] < \infty$. Assume there exists $y \in E$ such that $\mathbb{E}[\sup_{\mathcal{V}} (\log^+(d(\phi_0(y, \theta), y)))^2] < \infty$.

(LM) Assume that $y \rightarrow \nabla \ell^{-1}(y)$ and $y \rightarrow \nabla \log(\det(\ell(y)))$ are Lipschitz functions.

Now that \mathbf{V} is well defined in terms of derivatives of s_0 that are well approximated using the SRE approach, the procedure is asymptotically normal:

THEOREM 6. *Under the assumptions of Theorem 4, **(AV)**, **(MM)**, **(LI)**, **(DL)** and **(LM)** then the asymptotic variance \mathbf{V} is well defined and the statistical inference is asymptotically normal, i.e.*

$$\sqrt{n}(\hat{\theta}_n - \hat{\theta}_0) \rightarrow \mathcal{N}(0, \mathbf{V})$$

in distribution for any $\theta_0 \in \overset{\circ}{\Theta}$ with the asymptotic matrix \mathbf{V} that is invertible.

PROOF. Under **(CI)** and **(LM)** one applies our Theorem 2 to $(\sup_{\mathcal{V}} g_t(\theta))$. Then $\mathbb{E}[\log^+(\sup_{\mathcal{V}} g_0(\theta))] < \infty$ and from the Lipschitz condition in **(DL)** we have $\mathbb{E}[\sup_{\mathcal{V}} \log^+(\|\Phi_0(\theta)\|)] < \infty$ for $\Phi_0(\theta) = D_{\theta}(\phi_0)(g_0(\theta), \theta)$ or $D_{x^2}^2(\phi_0)(g_0(\theta), \theta)$ or $D_{\theta, x}^2(\phi_0)(g_0(\theta), \theta)$ or $D_{\theta^2}^2(\phi_0)(g_0(\theta), \theta)$. Using the existence of these logarithmic moments and the relation $\mathbb{E}[\log(\sup_{\mathcal{V}} \Lambda_{\phi_0})] < 0$, we apply recursively the Theorem 1 and prove the existence of continuous first and second derivatives of $(g_t(\theta))$ on \mathcal{V} as solutions of functional SRE. The asymptotic normality follows from a Taylor development on the first partial derivatives of S_n (see Section 5 of [Bardet and Wintenberger \(2009\)](#) for more details):

$$\nabla_i S_n(\hat{\theta}_n) - \nabla_i S_n(\theta_0) = \mathbb{H}S_n(\tilde{\theta}_{n,i})(\hat{\theta}_n - \theta_0)$$

Then the asymptotic normality follows from the following sufficient conditions:

1. $n^{-1/2} \nabla S_n(\theta_0) \rightarrow \mathcal{N}(0, \mathbf{Q})$,
2. $\|n^{-1} \mathbb{H}S_n(\tilde{\theta}_n) - \mathbf{P}\|$ converges a.s. to 0 for any sequence $(\tilde{\theta}_n)$ converging a.s. to θ_0 and \mathbf{P} is invertible,
3. $n^{-1/2} \|\nabla \hat{S}_n(\hat{\theta}_n) - \nabla S_n(\hat{\theta}_n)\|$ converges a.s. to 0.

Due to its specific expression and that (Z_t) is a normalized difference of martingales sequence with finite moments of order 4, $(\nabla S_n(\theta_0))$ is a martingale, see [Bardet and Wintenberger \(2009\)](#) for detailed computations. Under **(MM)**, the CLT for differences of martingale applied to $(\nabla S_n(\theta_0))$ leads to the first condition. The first part of the second condition are derived from similar arguments than in the proof of Theorem 5 and an application of the Cesaro mean theorem ensuring that $n^{-1} \|\mathbb{H}S_n(\tilde{\theta}_n) - \sum_{t=1}^n \mathbb{H}S_t(\theta_0)\| \rightarrow 0$ a.s. The ergodic Theorem on $(\mathbb{H}S_t(\theta_0))$ with **(MM)** leads to $\|n^{-1} \mathbb{H}S_n(\tilde{\theta}_n) - \mathbf{P}\| \rightarrow 0$ a.s. The fact that \mathbf{P} is invertible follows from **(LI)**, see [Bardet and Wintenberger \(2009\)](#) for detailed computations. Finally the third condition is obtained by applying Theorem 2.10 of [Straumann and Mikosch \(2006\)](#) to the SRE satisfied by (∇g_t) and its approximative SRE satisfied by $(\nabla \hat{g}_t)$ uniformly on \mathcal{V} . Thus $\sup_{\mathcal{V}} \|\nabla \hat{g}_t - \nabla g_t\| \xrightarrow{\text{e.a.s.}} 0$ as $t \rightarrow \infty$ and Lipschitz conditions on $\nabla \ell^{-1}$ and $\nabla \log(\det(\ell))$ in **(LM)** and arguments similar than in [Straumann \(2005\)](#) leads to the desired result. \square

4. Applications to GARCH and EGARCH type models.

4.1. *Some applications to GARCH type models.* In the GARCH type models, the stationarity assumption **(ST)** is crucial, whereas the continuous invertibility condition **(CI)** is automatically satisfied due to the form of the model. Sufficient conditions for **(ST)** have been extensively studied in the literature and they can be necessary in linear cases. The asymptotic properties of the QMLE (that coincides with the SRE based inference in these cases) follow from Theorems 4 and 6. Thus, we recover and slightly refine existing results in the AGARCH and CCC-GARCH models (we refer the reader to [Straumann \(2005\)](#) and [Francq and Zakoïan \(2011\)](#) respectively for details in these both cases).

First, let us consider the univariate APGARCH(p, q) model introduced in [Ding, Granger and Engle \(1993\)](#), [Zakoïan \(1994\)](#) and studied in [Straumann \(2005\)](#):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i (|X_{t-i}| - \gamma X_{t-i})^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{Z},$$

where $\alpha_0 > 0$, $\alpha_i, \beta_j \geq 0$ and $|\gamma| \leq 1$ (it coincides with the GARCH(p, q) model if $\gamma = 0$). Then we derive the strong consistency and the asymptotic normality directly from our Theorems 4 and 6. The conditions we obtained coincides with these of Theorem 5.5 and Theorem 8.1 of [Straumann and Mikosch \(2006\)](#) except their useless condition (8.1) as one does not need moments of any order.

Second, let us consider the multivariate CCC-GARCH(p, q) model introduced by [Bollerslev \(1990\)](#), first studied in [Jeantheau \(1998\)](#) and refined in [Francq and Zakoïan \(2011\)](#)

$$Diag(\Sigma_t^2) = A_0 + \sum_{i=1}^q A_i Diag(X_{t-i} X_{t-i}^T) + \sum_{i=1}^p B_i Diag(\Sigma_{t-i}^2)$$

and $(\Sigma_t^2)_{i,j} = \rho_{i,j} \sqrt{(\Sigma_t^2)_{i,i} (\Sigma_t^2)_{j,j}}$ for all (i, j) , where $Diag(M)$ is the vector of the diagonal elements of M . A necessary and sufficient conditions for **(ST)** is given in term of top Lyapunov condition in [Francq and Zakoïan \(2011\)](#). We recover the strong consistency and the asymptotic normality of [Francq and Zakoïan \(2011\)](#) directly from our Theorems 4 and 6.

4.2. *Application to the EGARCH(1,1) model.* Let (Z_t) be an iid sequence of random variables not concentrated on two points such that $\mathbb{E}(Z_0^2) = 1$.

The EGARCH(1, 1) model introduced by [Nelson \(1991\)](#) is an AR(1) model for $\log \sigma_t^2$,

$$X_t = \sigma_t Z_t \quad \text{with} \quad \log \sigma_t^2 = \alpha_0 + \beta_0 \log \sigma_{t-1}^2 + W_{t-1}(\theta_0)$$

where $W_t(\theta_0) = \gamma_0 Z_t + \delta_0 |Z_t|$ are the innovations of this AR(1) model. Let $\theta_0 = (\alpha_0, \beta_0, \gamma_0, \delta_0)$ be the unknown parameter. Assume that $|\beta_0| < 1$ such that there exists a stationary solution having a MA(∞) representation:

$$(11) \quad \log \sigma_t^2 = \alpha_0(1 - \beta_0)^{-1} + \sum_{k=1}^{\infty} \beta_0^{k-1} W_{t-k}(\theta_0).$$

The moments assumptions on Z_t ensures that the process $(\log \sigma_t^2)$ is ergodic, strongly and weakly stationary. Then the volatilities process (σ_t^2) is also ergodic and strongly stationary and **(ST)** holds. However, it does not necessarily have finite moment of any order.

The invertibility of the stationary solution of the EGARCH(1, 1) model does not hold in general. A sufficient condition for invertibility is given in [Straumann and Mikosch \(2006\)](#). Let us study its link with our condition **(CI)**. As $(\log \sigma_t^2)$ satisfies the SRE

$$\log \sigma_t^2 = \alpha_0 + \beta_0 \log \sigma_{t-1}^2 + (\delta_0 X_{t-1} + \gamma_0 |X_{t-1}|) \exp(-\log \sigma_{t-1}^2 / 2),$$

if it has a non anticipative solution the model is invertible. Keeping the notation of Section 2, the function h is now the logarithmic function and the SRE (8) holds with (ϕ_t) defined by

$$\phi_t(\cdot; \theta) : s \mapsto \alpha + \beta s + (\gamma X_{t-1} + \delta |X_{t-1}|) \exp(-s/2)$$

We check that the ϕ_t are random functions generated by \mathcal{G}_{t-1} . For any $\theta \in \mathbb{R} \times \mathbb{R}^+ \times \{\gamma \geq |\delta|\}$ we restrict $\phi_t(\cdot; \theta)$ on the complete separable metric space $[\alpha/(1 - \beta), \infty)$ equipped with $d(x, y) = |x - y|$. The process $(\phi_t(\cdot; \theta))$ is a stationary ergodic sequence of Lipschitz maps from $[\alpha/(1 - \beta), \infty)$ to $[\alpha/(1 - \beta), \infty)$ with Lipschitz coefficients

$$\Lambda(\phi_t(\cdot, \theta_0)) \leq \max\{\beta, 2^{-1}(\gamma X_{t-1} + \delta |X_{t-1}|) \exp(-2^{-1}\alpha/(1 - \beta)) - \beta\}.$$

The technical smoothness assumption **(CL)** is automatically satisfied as

$$(\Lambda_{\phi_t}(\theta))_t = (\max(\beta, 2^{-1}(\gamma X_{t-1} + \delta |X_{t-1}|) \exp(-2^{-1}\alpha/(1 - \beta)) - \beta))_t$$

is a stationary process of continuous functions of θ . The EGARCH(1,1) is continuously invertible on any compact set $\Theta \subset \mathbb{R} \times \mathbb{R}^+ \times \{\gamma \geq |\delta|\}$ such that

$$(12) \quad \mathbb{E}[\log(\max\{\beta, 2^{-1}(\gamma X_{t-1} + \delta |X_{t-1}|) \exp(-2^{-1}\alpha/(1-\beta)) - \beta\})] < 0.$$

This sufficient condition for continuous invertibility depends on the distribution of the observations (X_t) . Remark that if $\theta_0 \in \Theta$ then it satisfies the condition of stationarity $\beta_0 < 1$ and, from the MA(∞) representation (11) of $\log \sigma_t^2$, our condition (12) expressed at θ_0 implies that

$$(13) \quad \mathbb{E} \left[\log \left(\max \left\{ \beta_0, 2^{-1} \exp \left(2^{-1} \sum_{k=0}^{\infty} \beta_0^k (\gamma_0 Z_{-k-1} + \delta_0 |Z_{-k-1}|) \right) \right. \right. \right. \\ \left. \left. \left. \times (\gamma_0 Z_0 + \delta_0 |Z_0|) - \beta_0 \right\} \right) \right] < 0.$$

It is the condition of invertibility of the EGARCH(1,1) model given in [Straumann and Mikosch \(2006\)](#). Applying our results of [Theorem 4](#) and [Theorem 5](#), we obtain

THEOREM 7. *For any compact subset Θ of $\mathbb{R} \times \mathbb{R}^+ \times \{\gamma \geq |\delta|\}$ satisfying (12) then $\hat{\theta}_n \rightarrow \theta_0$ and $\hat{\sigma}_n^2 - \sigma_n^2 \rightarrow 0$ a.s. as $n \rightarrow \infty$ with $\hat{\sigma}_t^2 = \exp(\hat{g}_t(\hat{\theta}_n))$ if $\theta_0 \in \Theta$.*

PROOF. The condition **(CI)** follows from $\mathbb{E}[\log \Lambda(\phi_t(\cdot, \theta))] < 0$ by assumption of Θ and $\mathbb{E}[\sup_{\Theta} \log \Lambda(\phi_t(\cdot, \theta))] < \infty$ since $\mathbb{E} \log |X_{t-1}| = E(\log \sigma + \log |Z_{t-1}|) < \infty$ as $\log \sigma_t^2$ has a MA(∞) representation (11) and Z is integrable. Moreover as $\log^+(d(\phi_0(0, \theta), 0)) = \log^+ |\alpha + (\gamma X_{-1} + \delta |X_{-1}|)|$ then fixing $y = 0$ one has $\mathbb{E}[\sup_{\Theta} (\log^+(d(\phi_0(y, \theta), y))^2) < \infty$.

In the EGARCH(1,1) model the link function is the exponential function $\ell(x) = \exp(x)$ and since we have $\log \sigma_t^2 \geq \alpha/(1-\beta)$, $1/\ell(x) = \exp(-x)$ is a Lipschitz function ($\log(\det(\ell)) = id$ is also a Lipschitz function). Moreover the volatility process (σ_t^2) is bounded from below by $C(\theta) = \exp(\alpha/(1-\beta))$. Finally, the identifiability condition $g_0(\theta) = h(\theta_0)$ iff $\theta = \theta_0$ is checked in Section 5.1 of [Straumann and Mikosch \(2006\)](#). \square

As a corollary of [Theorem 6](#) we get the asymptotic normality of the statistical inference in the EGARCH(1,1) model. It holds under the following necessary and sufficient condition of the existence of the asymptotic variance \mathbf{V} :

$$(MM') \quad \mathbb{E}[Z_0^4] < \infty \text{ and } \mathbb{E}[(\beta_0 - 2^{-1}(\gamma_0 Z_0 + \delta_0 |Z_0|)^2) < 1.$$

THEOREM 8. *Assume that Θ is well chosen as in Corollary 7 and that **(MM')** holds then $\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow \mathcal{N}(0, \mathbf{V})$ in law with an invertible matrix \mathbf{V} for any $\theta_0 \in \overset{\circ}{\Theta}$.*

PROOF. By definition, (ϕ_t) is 2-times continuously differentiable and simple computations give $D_x(\phi_t)(x, \theta) = \beta - 2^{-1}(\gamma X_{t-1} + \delta |X_{t-1}|) \exp(-x/2)$, $D_\theta(\phi_t)(x, \theta) = (1, x, X_{t-1} \exp(-x/2), |X_{t-1}| \exp(-x/2))^T$, $D_{x^2}^2(\phi_t)(x, \theta) = 4^{-1}(\gamma X_{t-1} + \delta |X_{t-1}|) \exp(-x/2)$, $D_{x,\theta}^2(\phi_t)(x, \theta) = (0, 1, 2^{-1} X_{t-1} \exp(-x/2), 2^{-1} |X_{t-1}| \exp(-x/2))^T$ and $D_{\theta^2}^2(\phi_t)(x, \theta) = 0$. Moreover, as the link function is $\ell(x) = \exp(x)$ is also 2-times continuously differentiable, the last assertion of the condition **(AV)** of Theorem 6 holds. The fact that **(MM)** holds under the conditions $\mathbb{E}[Z_0^4] < \infty$ and $\beta_0^2 - \delta_0 \mathbb{E}|Z_0| + (\delta_0^2 + \gamma_0^2)/4 < 1$ is technical and postponed to the Appendix. The fact that **(LI)** holds if Z_0 is not concentrated on two points is classical, see for instance Lemma 8.2 of [Straumann and Mikosch \(2006\)](#). Assumption **(DL)** is satisfied from the expressions of the derivatives (that are Lipschitz functions) and as all the logarithmic moments are finite due to $\mathbb{E}[\log(X_{t-1}^2)] < \infty$. Finally **(LM)** is automatically satisfied due to the specific expression of the link function. Thus Theorem 6 applies. \square

Acknowledgments. One of the author would like to thank C. Francq and J.-M. Zakoïan for helpful discussions and for having pointing out some mistakes on an earlier version.

References.

- BARDET, J. M. and WINTENBERGER, O. (2009). Asymptotic normality of the quasi maximum likelihood estimator for multidimensional causal processes. *Ann. Statist.* **37** 2730-2759.
- BERKES, I., HORVATH, L. and KOKOSZKA, P. (2003). GARCH processes: structure and estimation. *Bernoulli* **9** 201-227.
- BOLLERSLEV, T. P. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31** 307-327.
- BOLLERSLEV, T. P. (1990). Modeling the coherence in short run nominal exchange rates: a multivariate generalized ARCH approach. *Review of Economics and Statistics* **72** 498-505.
- BOUGEROL, P. (1993). Kalman filtering with random coefficients and contractions. *SIAM J. Control and Optimization* **31** 942-959.
- BOUGEROL, P. and PICARD, N. (1992). Stationarity of GARCH processes and of some non-negative time series. *J. Econometrics* **52** 115-127.
- BRANDT, M. W. and JONES, C. S. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business & Economic Statistics* **24** 470-486.
- DEMOS, A. and KYRIAKOPOULOU, D. (2009). Asymptotic expansions of the QMLEs in the EGARCH(1,1) model. preprint.
- DING, Z., GRANGER, C. W. J. and ENGLE, R. (1993). A long memory property of stock market returns and a new model. *J. Empirical Finance* **1** 83-106.
- ENGLE, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* **50** 987-1007.

- FRANCO, C. and ZAKOÏAN, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **10** 605-637.
- FRANCO, C. and ZAKOÏAN, J. M. (2011). QML estimation of a class of multivariate asymmetric GARCH models. forthcoming in *Econometric Theory*.
- GLOSTEN, L. R., JAGANNATHAN, R. and RUNKLE, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* **48** 1779-1801.
- GRANGER, C. W. J. and ANDERSEN, A. (1978). On the invertibility on time series models. *Stochastic Processes and their Applications* **8** 87-92.
- HARVEY, A. (2010). Exponential conditional volatility models. Cambridge Working Papers in Economics report No. 1040, Faculty of Economics, University of Cambridge.
- HE, C., TERÄSVIRTA, T. and MALMSTEN, H. (2002). Moment structure of a family of first-order exponential GARCH models. *Econometric Theory* **18** 868-885.
- JEANTHEAU, T. (1993). Modèles autorégressifs à erreur conditionnellement hétéroscédastique. PhD thesis, Université Paris VII.
- JEANTHEAU, T. (1998). Strong consistency of estimation for multivariate ARCH models. *Econometric Theory* **14** 70-86.
- NELSON, D. B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* **6** 318-334.
- NELSON, D. B. (1991). Conditional Heteroskedasticity in Asset Returns : A New Approach. *Econometrica* **59** 347-370.
- RODRIGUEZ, M. J. and RUIZ, E. (2009). GARCH models with leverage effect: differences and similarities. Statistics and Econometrics Working Papers report No. ws090302, Universidad Carlos III, Departamento de Estadística y Econometría.
- SOROKIN, A. (2011). Non-invertibility in some heteroscedastic models. Arxiv preprint #1104.3318.
- STRAUMANN, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models. Lectures Notes in Statistics* **181**. Springer, New York.
- STRAUMANN, D. and MIKOSCH, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equation approach. *Ann. Statist.* **34** 2449-2495.
- TONG, H. (1993). *Non-Linear Time Series, A Dynamical System Approach. Oxford Statistical Science Series* **6**. Oxford Science Publications, Oxford.
- WINTENBERGER, O. and CAI, S. (2011). The EGARCH(1,1) process: structure, estimation and prediction. Working paper.
- ZAFFARONI, P. (2009). Whittle estimation of EGARCH and other exponential volatility models. *J. Econometrics* **151** 190-200.
- ZAKOÏAN, J. M. (1994). Threshold heteroscedastic models. *J. Econom. Dynam. Control* **18** 931-955.

Appendix: checking (MM) in the EGARCH(1,1) model under (MM').

To check (MM) is technical as we have to compute explicitly the diagonal terms the matrix $\mathbf{B} = \mathbb{E}[\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T]$. Similar computations have been done in [Demos and Kyriakopoulou \(2009\)](#). Remember that $\mathbf{V} = \mathbf{P}^{-1}\mathbf{Q}\mathbf{P}^{-1}$ with $\mathbf{P} = \mathbb{E}[\mathbb{H}s_0(\theta_0)]$ and $\mathbf{Q} = \mathbb{E}[\nabla s_0(\theta_0)\nabla s_0(\theta_0)^T]$. Let us first prove the three identities $\mathbf{P} = 2^{-1}\mathbf{B}$, $\mathbf{Q} = 4^{-1}(\mathbb{E}Z_0^4 - 1)\mathbf{B}$ and thus $\mathbf{V} =$

$(\mathbb{E}Z_0^4 - 1)\mathbf{B}^{-1}$. For the first identity, we compute

$$\begin{aligned}\mathbf{P} &= 2^{-1}\mathbb{E}\left[(\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T Z_0^2 + \mathbb{H}g_t(\theta_0)(1 - Z_0^2)\right] \\ &= 2^{-1}\mathbb{E}[\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T] = 2^{-1}\mathbf{B}.\end{aligned}$$

For the second identity, we compute

$$\begin{aligned}\mathbf{Q} &= \mathbb{E}\left[\frac{1}{4}\mathbb{E}\left[\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T(1 - Z_t^2)^2 \mid \mathcal{F}_{t-1}\right]\right] \\ &= 4^{-1}\mathbb{E}[(1 - Z_0^2)^2]\mathbb{E}[\nabla g_t(\theta_0)(\nabla g_t(\theta_0))^T] = 4^{-1}(\mathbb{E}Z_0^4 - 1)\mathbf{B}\end{aligned}$$

and the third identity follows the first ones. Thus, for checking the assumption **(MM)**, it is enough to check that diagonal coefficients \mathbf{B}_{ii} are well defined when $\mathbb{E}(Z_0^4) < \infty$. Let us denote $W_t = \gamma_0 Z_t + \delta_0 |Z_t|$, $U_t = (1, \log \sigma_t^2, Z_t, |Z_t|)$ and $V_t = \beta_0 - 2^{-1}(\gamma_0 Z_t + \delta_0 |Z_t|)$. Then $(\nabla g_t(\theta_0))$ is the solution of the linear SRE

$$\nabla g_t(\theta_0) = U_{t-1} + V_{t-1}\nabla g_{t-1}(\theta_0) = \sum_{l=1}^{\infty} \left(U_{t-l} \prod_{k=1}^{l-1} V_{t-k} \right).$$

Using the convention $\prod_{k=1}^0 V_{t-k} = 1$, we obtain the expression

$$\nabla g_t(\theta_0) = \sum_{l=1}^{\infty} \left(U_{t-l} \prod_{k=1}^{l-1} V_{t-k} \right).$$

More precisely, we have the expressions: then

$$\begin{aligned}\mathbf{B}_{11} &= \mathbb{E}\left(\frac{\partial g_t(\theta_0)}{\partial \theta_1}\right)^2 = \mathbb{E}\left[\sum_{l=1}^{\infty} \prod_{k=1}^{l-1} V_{t-k}\right]^2, \\ \mathbf{B}_{22} &= \mathbb{E}\left(\frac{\partial g_t(\theta_0)}{\partial \theta_2}\right)^2 = \mathbb{E}\left[\sum_{l=1}^{\infty} \log \sigma_{t-l}^2 \prod_{k=1}^{l-1} V_{t-k}\right]^2, \\ \mathbf{B}_{33} &= \mathbb{E}\left(\frac{\partial g_t(\theta_0)}{\partial \theta_3}\right)^2 = \mathbb{E}\left[\sum_{l=1}^{\infty} Z_{t-l} \prod_{k=1}^{l-1} V_{t-k}\right]^2, \\ \mathbf{B}_{44} &= \mathbb{E}\left(\frac{\partial g_t(\theta_0)}{\partial \theta_4}\right)^2 = \mathbb{E}\left[\sum_{l=1}^{\infty} |Z_{t-l}| \prod_{k=1}^{l-1} V_{t-k}\right]^2.\end{aligned}$$

To prove that condition **(MM)** is satisfied, i.e. that $\sum_{i=1}^4 \mathbf{B}_{ii} < \infty$, we use the following Lemma

LEMMA 1. $\sum_{i=1}^4 \mathbf{B}_{ii} < \infty$ iff $\mathbb{E}V_0^2 < 1$.

PROOF. That the first coefficient B_{11} is finite comes easily:

$$\begin{aligned}
\mathbf{B}_{11} &= \mathbb{E}\left(\sum_{l=1}^{\infty} \prod_{k=1}^{l-1} V_{t-k}\right)^2 = \mathbb{E}\left(\sum_{l=1}^{\infty} \sum_{l'=1}^{\infty} \prod_{k=1}^{l-1} V_{t-k} \prod_{k'=1}^{l'-1} V_{t-k'}\right) \\
&= \mathbb{E}\left(2 \sum_{l \geq 1} \left(\prod_{k=1}^{l-1} V_{t-k}\right)^2 \sum_{l' > l} \prod_{k'=l}^{l'-1} V_{t-k'}\right) + \mathbb{E} \sum_{l=1}^{\infty} \left(\prod_{k=1}^{l-1} V_{t-k}\right)^2 \\
&= 2 \sum_{l \geq 1} (\mathbb{E}V_0^2)^{l-1} \frac{\mathbb{E}V_0}{1 - \mathbb{E}V_0} + \frac{1}{1 - \mathbb{E}V_0^2} \\
&= 2 \frac{1}{1 - \mathbb{E}V_0^2} \times \frac{\mathbb{E}V_0}{1 - \mathbb{E}V_0} + \frac{1}{1 - \mathbb{E}V_0^2}.
\end{aligned}$$

For the second coefficient \mathbf{B}_{22} , it is more complicated. We need some preliminary work. We know that $W_t = \gamma_0 Z_t + \delta_0 |Z_t| = 2(\beta_0 - V_t)$, and

$$\log \sigma_t^2 = \frac{\alpha_0}{1 - \beta_0} + \sum_{k=1}^{\infty} \beta_0^{k-1} W_{t-k} = \frac{\alpha_0 + 2\beta_0}{1 - \beta_0} - 2 \sum_{k=1}^{\infty} \beta_0^{k-1} V_{t-k}$$

so, we decompose \mathbf{B}_{22} into three parts,

$$\begin{aligned}
\mathbf{B}_{22} &= \mathbb{E} \left[\sum_{l=1}^{\infty} \log \sigma_{t-l}^2 \prod_{k=1}^{l-1} V_{t-k} \right]^2 \\
&= \mathbb{E} \left[\sum_{l=1}^{\infty} \left(\frac{\alpha_0 + 2\beta_0}{1 - \beta_0} - 2 \sum_{k=1}^{\infty} \beta_0^{k-1} V_{t-l-k} \right) \prod_{k'=1}^{l-1} V_{t-k'} \right]^2 \\
&= \left(\frac{\alpha_0 + 2\beta_0}{1 - \beta_0} \right)^2 \mathbb{E} \left[\sum_{l=1}^{\infty} \prod_{k'=1}^{l-1} V_{t-k'} \right]^2 + 4 \mathbb{E} \left[\sum_{l=1}^{\infty} \sum_{k=1}^{\infty} \beta_0^{k-1} V_{t-l-k} \prod_{k'=1}^{l-1} V_{t-k'} \right]^2 \\
&\quad - 4 \times \frac{\alpha_0 + 2\beta_0}{1 - \beta_0} \mathbb{E} \left[\sum_{l=1}^{\infty} \sum_{k=1}^{\infty} \beta_0^{k-1} V_{t-l-k} \prod_{k'=1}^{l-1} V_{t-k'} \right].
\end{aligned}$$

That the first term of the sum is finite is already known. For the last term, it is straightforward from $\mathbb{E} \sum_{l=1}^{\infty} \sum_{k=1}^{\infty} \beta_0^{k-1} V_{t-l-k} \prod_{k'=1}^{l-1} V_{t-k'} = (1 - \beta_0)^{-1} \mathbb{E}V_0 / (1 - \mathbb{E}V_0)$. For the second term of the sum, we need an expan-

sion

$$\begin{aligned}
& \left[\sum_{l=1}^{\infty} \sum_{k=1}^{\infty} \beta_0^{k-1} V_{t-l-k} \prod_{k'=1}^{l-1} V_{t-k'} \right]^2 \\
&= 2 \times \sum_{1 \leq l < l' < \infty} \sum_{p, q=1}^{\infty} \beta_0^{p+q-2} V_{t-l-p} V_{t-l'-q} \prod_{p'=1}^{l-1} V_{t-p'}^2 \prod_{q'=l}^{l'-1} V_{t-q'} \\
&\quad + \sum_{l=1}^{\infty} \sum_{p, q=1}^{\infty} \beta_0^{p+q-2} V_{t-l-p} V_{t-l-q} \prod_{p'=1}^{l-1} V_{t-p'}^2 \\
&= 4 \times \sum_{1 \leq l < l' < \infty} \sum_{1 \leq p < q < \infty} \beta_0^{p+q-2} V_{t-l-p} V_{t-l'-q} \prod_{p'=1}^{l-1} V_{t-p'}^2 \prod_{q'=l}^{l'-1} V_{t-q'} \\
&\quad + 2 \times \sum_{1 \leq l < l' < \infty} \sum_{p=1}^{\infty} \beta_0^{2p-2} V_{t-l-p} V_{t-l'-p} \prod_{p'=1}^{l-1} V_{t-p'}^2 \prod_{q'=l}^{l'-1} V_{t-q'} \\
&\quad + 2 \sum_{l=1}^{\infty} \sum_{1 \leq p < q < \infty} \beta_0^{p+q-2} V_{t-l-p} V_{t-l-q} \prod_{p'=1}^{l-1} V_{t-p'}^2 + \sum_{l=1}^{\infty} \sum_{p=1}^{\infty} \beta_0^{2p-2} V_{t-l-p}^2 \prod_{p'=1}^{l-1} V_{t-p'}^2
\end{aligned}$$

and in expectation we obtain a bounded term if $\mathbb{E}V_0^2 < 1$:

$$\begin{aligned}
& \mathbb{E} \left[\sum_{l=1}^{\infty} \sum_{k=1}^{\infty} \beta_0^{k-1} V_{t-l-k} \prod_{k'=1}^{l-1} V_{t-k'} \right]^2 \\
&= 4 \times \frac{\mathbb{E}V_0^2}{1 - \mathbb{E}V_0^2} \left[\frac{\beta_0}{(1 - \beta_0)(1 - \beta_0^2)} \frac{\mathbb{E}V_0}{1 - \mathbb{E}V_0} - \frac{1}{(1 - \beta_0)(1 - \beta_0^2)} \frac{\mathbb{E}V_0 \beta}{1 - \beta_0^2 \mathbb{E}V_0} \right] \\
&\quad + 4 \times \frac{\beta_0 (\mathbb{E}V_0)^3}{(1 - \beta_0)(1 - \beta_0^2)(1 - \beta_0^2 \mathbb{E}V_0)} \frac{1}{1 - \mathbb{E}V_0^2} \\
&\quad + 2 \times \frac{1}{1 - \beta_0^2} \frac{\mathbb{E}V_0^2}{1 - \mathbb{E}V_0^2} \left[\frac{\mathbb{E}V_0}{1 - \mathbb{E}V_0} - \frac{\mathbb{E}V_0}{1 - \beta_0^2 \mathbb{E}V_0} \right] \\
&\quad + 2 \frac{1}{1 - \mathbb{E}V_0^2} \frac{(\mathbb{E}V_0)^3}{(1 - \beta_0^2)(1 - \beta_0^2 \mathbb{E}V_0)} \\
&\quad + 2 \frac{1}{1 - \mathbb{E}V_0^2} (\mathbb{E}V_0)^2 \frac{\beta_0}{(1 - \beta_0)(1 - \beta_0^2)} + \frac{\mathbb{E}V_0^2}{1 - \mathbb{E}V_0^2} \frac{1}{1 - \beta_0^2}.
\end{aligned}$$

That \mathbf{B}_{33} is finite under $\mathbb{E}V_0^2 < 1$ comes from

$$\begin{aligned}
\mathbf{B}_{33} &= \mathbb{E} \left[\sum_{l=1}^{\infty} Z_{t-l} \prod_{k=1}^{l-1} V_{t-k} \right]^2 \\
&= 2\mathbb{E} \sum_{l=1}^{\infty} \sum_{l'>l}^{\infty} Z_{t-l} Z_{t-l'} \prod_{k=1}^{l-1} V_{t-k} \prod_{k'=1}^{l'-1} V_{t-k'} + \mathbb{E} \sum_{l=1}^{\infty} Z_{t-l}^2 \left(\prod_{k=1}^{l-1} V_{t-k} \right)^2 \\
&= \mathbb{E} Z_0^2 \sum_{l=1}^{\infty} (\mathbb{E}V_0^2)^{l-1} = \frac{\mathbb{E}Z_0^2}{1 - \mathbb{E}V_0^2}.
\end{aligned}$$

That the last coefficient is also finite comes from the computation

$$\begin{aligned}
\mathbf{B}_{44} &= \mathbb{E} \left[\sum_{l=1}^{\infty} |Z_{t-l}| \prod_{k=1}^{l-1} V_{t-k} \right]^2 \\
&= 2\mathbb{E} \sum_{l=1}^{\infty} \sum_{l'>l}^{\infty} |Z_{t-l}| |Z_{t-l'}| \prod_{k=1}^{l-1} V_{t-k} \prod_{k'=1}^{l'-1} V_{t-k'} + \mathbb{E} \sum_{l=1}^{\infty} Z_{t-l}^2 \left(\prod_{k=1}^{l-1} V_{t-k} \right)^2 \\
&= 2 \sum_{l=1}^{\infty} \sum_{l'>l}^{\infty} \mathbb{E} |Z_{t-l'}| \mathbb{E} \left(\prod_{k=1}^{l-1} V_{t-k}^2 \right) \mathbb{E} \left(|Z_{t-l}| \prod_{k'=1}^{l'-1} V_{t-k'} \right) + \frac{\mathbb{E}Z_0^2}{1 - \mathbb{E}V_0^2} \\
&= 2 \sum_{l=1}^{\infty} \sum_{l'>l}^{\infty} (\mathbb{E} |Z_0|) (\mathbb{E}V_0^2)^{l-1} (\mathbb{E} |Z_0| V_0) \mathbb{E}V_0^{l'-l-1} + \frac{\mathbb{E}Z_0^2}{1 - \mathbb{E}V_0^2} \\
&= \frac{2\mathbb{E} |Z_0| (\mathbb{E} |Z_0| V_0)}{(1 - \mathbb{E}V_0)(1 - \mathbb{E}V_0^2)} + \frac{\mathbb{E}Z_0^2}{1 - \mathbb{E}V_0^2}.
\end{aligned}$$

□