ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Specification tests for GARCH processes with nuisance parameters on the boundary

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Cavaliere, G., Perera, I., Rahbek, A. (2023). Specification tests for GARCH processes with nuisance parameters on the boundary. JOURNAL OF BUSINESS & ECONOMIC STATISTICS, in press, 1-18 [10.1080/07350015.2023.2173206].

Availability:
This version is available at: https://hdl.handle.net/11585/915231 since: 2023-02-13

Published:
DOI: http://doi.org/10.1080/07350015.2023.2173206

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

12 October 2024

This is the final peer-reviewed accepted manuscript of:

Giuseppe Cavaliere, Indeewara Perera & Anders Rahbek (2023) Specification Tests for GARCH Processes with Nuisance Parameters on the Boundary, Journal of Business & Economic Statistics.

The final published version is available online at:

https://doi.org/10.1080/07350015.2023.2173206

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/)

When citing, please refer to the published version.

Specification tests for GARCH processes with nuisance parameters on the boundary

Giuseppe Cavaliere
Department of Economics, University of Exeter, UK and
Department of Economics, University of Bologna, Italy

Indeewara Perera
Department of Economics, The University of Sheffield, UK

Anders Rahbek*
Department of Economics, University of Copenhagen, Denmark

January 14, 2023

Abstract

This paper develops tests for the correct specification of the conditional variance function in GARCH models when the true parameter may lie on the boundary of the parameter space. The test statistics considered are of Kolmogorov-Smirnov and Cramér-von Mises type, and are based on empirical processes marked by centered squared residuals. The limiting distributions of the test statistics depend on unknown nuisance parameters in a non-trivial way, making the tests difficult to implement. We therefore introduce a novel bootstrap procedure which is shown to be asymptotically valid under general conditions, irrespective of the presence of nuisance parameters on the boundary. The proposed bootstrap approach is based on shrinking of the parameter estimates used to generate the bootstrap sample toward the boundary of the parameter space at a proper rate. It is simple to implement and fast in applications, as the associated test statistics have simple closed form expressions. Although the bootstrap test is designed for a data generating process with fixed parameters (i.e., independent of the sample size n), we also discuss how to obtain valid inference for sequences of DGPs with parameters approaching the boundary at the $n^{-1/2}$ rate. A simulation study demonstrates that the new tests: (i) have excellent finite sample behaviour in terms of empirical rejection probabilities under the null as well as under the alternative; (ii) provide a useful complement to existing procedures based on Ljung-Box type approaches. Two data examples illustrate the implementation of the proposed tests in applications.

Keywords: Bootstrap; Kolmogorov-Smirnov; Cramér-von Mises; Marked empirical process.

*Corresponding author: Anders Rahbek; email: anders.rahbek@econ.ku.dk.

1. Introduction

Generalized autoregressive conditionally heteroscedastic (GARCH) models introduced by Bollerslev (1986) are widely used for modelling various financial time series processes. The data generation mechanism of a GARCH model requires the conditional variance to be always strictly positive, which is generally obtained by imposing a strictly positive intercept and non-negative GARCH coefficients in the conditional variance equation. Consequently, in GARCH models, the admissible parameter space typically needs to be inequality restricted. This represents an important difference between GARCH and other popular time series models, such as AR and ARMA models. Although omnibus specification testing in GARCH type models against unspecified alternatives has attracted considerable attention in the recent literature, a crucial weakness in the current theory remains the exclusion of the presence of nuisance parameters on the boundary. This paper contributes towards addressing this issue by developing new statistical methodology for specification testing in GARCH models.

There are a number of different GARCH models available in the literature and many of them are nonnested models (see Francq and Zakoïan, 2010). Therefore, in many cases, a sensible way to proceed when testing a specification of a GARCH model is to leave the alternative model unspecified, or to test the lack-of-fit. This type of tests, also known as omnibus tests, have their roots in the seminal work of Kolmogorov (1933) on testing for a specific probability distribution function, and Grenander and Rosenblatt (1957) on testing the hypothesis of white noise dependence. Several omnibus specification tests in GARCH type models have been proposed in the literature. These include tests based on weighted empirical processes of standardized residuals (Koul and Ling, 2006; Escanciano, 2010), spectral distributions based tests in the frequency domain (Hidalgo and Zaffaroni, 2007; Escanciano, 2008), residual based tests for nonnegative valued processes (Fernandes and Grammig, 2005; Koul *et al.*, 2012), and Khmaladze type (Khmaladze, 1981) martingale transformations based tests (Bai 2003; Perera and Koul 2017), amongst others.

A key regularity condition imposed by the aforementioned specification tests is to restrict the true parameter to the interior of the null parameter space. Since the parameter space of a GARCH-type model is inequality restricted, this condition is not typically satisfied if some ARCH or GARCH coefficients are zero, because then the true parameter may lie on the boundary of the parameter space. Therefore, for the theory developed in the

above cited papers, the true parameter being an interior point is essential; for example, the limiting process obtained in Theorem 2.1 in Hidalgo and Zaffaroni (2007) would not be Gaussian if, for instance, a GARCH(p, q) model is estimated when the underlying true process is a GARCH$(p-1, q)$, or a GARCH$(p, q-1)$ process. Similarly, the asymptotic properties of the other aforementioned papers would also not hold when some nuisance parameters lie on the boundary.

In this paper we contribute towards the literature of specification testing in GARCH models by developing a new class of tests for the correct specification of the conditional variance function while allowing the null model to have an unknown number of nuisance parameters on the boundary of the parameter space. Our test statistics are Kolmogorov-Smirnov and Cramér-von Mises type statistics based on functionals of an empirical process marked by centered squared residuals and are easy to compute. The limiting distributions of the test statistics depend on (unknown) nuisance parameters in a non-trivial way. We propose a bootstrap method to implement the tests and show that it is asymptotically valid under general conditions, irrespective of the presence of nuisance parameters on the boundary. The proposed bootstrap approach is simple to implement, and is based on a method of shrinkage of the parameter estimates used to generate the bootstrap sample toward the boundary of the parameter space at an appropriate rate. This approach is similar to the related bootstrap scheme advocated in Cavaliere *et al.* (2022), in a different context, for bootstrapping likelihood ratio statistics, and it also has its roots in the modified bootstrap approach considered in Chatterjee and Lahiri (2011) for bootstrapping Lasso-type estimators. Our bootstrap tests are shown to be consistent against fixed alternatives. We also separately consider the case where the nuisance parameters lie in the interior of the parameter space, and show that the bootstrap implementations of these tests under standard residual based bootstrap are asymptotically valid and consistent. Our tests can be implemented easily because the test statistics have simple closed form expressions. A simulation study shows that the proposed tests have desirable finite sample properties. We illustrate the testing procedure by considering two real data examples.

The rest of this paper is structured as follows. Section 2 formulates the problem, defines the estimators and test statistics. Section 3 provides the results relating to the asymptotic validity and consistency of the bootstrap tests when the parameters are in the interior of the parameter space. Section 4 considers inference when some components of the true parameter lie on the boundary of the parameter space. Section 5 describes

a simulation study. Two empirical illustrations are discussed in Section 6. Section 7 concludes the paper. The proofs and additional simulation results are relegated to the online supplementary material.

2. Formulation of the Problem

Let $(Y_1, Y_2, \ldots, Y_n)$ be a realization of an observable stationary process $\{Y_t\}$ satisfying

$$Y_t = h_t^{1/2}\varepsilon_t, \quad t \in \mathbb{Z} := \{0, \pm 1, \pm 2, \cdots\}, \tag{1}$$

where the errors ε_t, $t \in \mathbb{Z}$, are independent and identically distributed (i.i.d.) random variables (r.v.'s) having zero mean and unit variance with common cumulative distribution function (c.d.f.) F_0, and $h_t = \mathrm{E}[Y_t^2 \mid \mathcal{H}_{t-1}]$, where $\mathcal{H}_t$ denotes the information available up to time t, $t \in \mathbb{Z}$.

As is well-known, a GARCH(p_1, p_2) model for h_t takes the form

$$h_t(\boldsymbol{\phi}) = \omega + \sum_{j=1}^{p_1} \alpha_j Y_{t-j}^2 + \sum_{k=1}^{p_2} \beta_k h_{t-k}(\boldsymbol{\phi}), \quad t \in \mathbb{Z}, \tag{2}$$

where the vector of parameters $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{p_1+p_2+1})' = (\omega, \alpha_1, \ldots, \alpha_{p_1}, \beta_1, \ldots, \beta_{p_2})'$ belongs to a compact parameter space

$$\Phi \subset (0, \infty) \times [0, \infty)^{p_1+p_2} \tag{3}$$

with $\omega > 0$, $\alpha_k \geq 0\,(k = 1, \ldots, p_1)$, $\beta_k \geq 0\,(k = 1, \ldots, p_2)$ and, in order to avoid well-known identification issues (see also Assumption (A3) below), it is assumed that $\sum_{k=1}^{p_1} \alpha_k \neq 0$ if $p_2 > 0$.

Suppose we wish to test the adequacy of the above GARCH(p_1, p_2) model for h_t, i.e., to test the null hypothesis

$$\begin{aligned} \mathsf{H}_0 : \quad & h_t = h_t(\boldsymbol{\phi}_0) = \omega_0 + \sum_{j=1}^{p_1} \alpha_{0j} Y_{t-j}^2 + \sum_{k=1}^{p_2} \beta_{0k} h_{t-k}(\boldsymbol{\phi}_0), \text{ a.s. for all } t, \text{ and} \\ & \text{for some } \boldsymbol{\phi}_0 = (\omega_0, \alpha_{01}, \ldots, \alpha_{0p_1}, \beta_{01}, \ldots, \beta_{0p_2})' \in \Phi, \end{aligned} \tag{4}$$

against the alternative $\mathsf{H}_1 : \mathsf{H}_0$ *is not true.*

Remark 1. *Note that in terms of notation, since some ARCH or GARCH coefficients may be zero, the null model (4) allows some components of $\boldsymbol{\phi}_0$ to be on the boundary of the parameter space Φ. For example, the ARCH($p_1 = 4$) given by $h_t = 0.2 + 0.3Y_{t-1}^2 + 0.6Y_{t-4}^2$ has $\boldsymbol{\phi}_0 = (0.2, 0.3, 0, 0, 0.6)'$ with $\alpha_{02} = \alpha_{03} = 0$, and hence α_{02} and α_{03} are boundary values.*

Let $\hat{\boldsymbol{\phi}}_n$ denote the Gaussian quasi maximum likelihood estimator [QMLE] defined by

$$\hat{\boldsymbol{\phi}}_n = \arg\min_{\boldsymbol{\phi}\in\boldsymbol{\Phi}} \sum_{t=1}^{n} \ell_t(\boldsymbol{\phi}), \quad \ell_t(\boldsymbol{\phi}) = \log h_t(\boldsymbol{\phi}) + [Y_t^2/h_t(\boldsymbol{\phi})], \tag{5}$$

with $h_t(\boldsymbol{\phi})$ being defined recursively by (2) for $t = 1, 2, \ldots, n$. To simplify the exposition, the vector of initial values, $\varsigma_0 = (Y_0, \ldots, Y_{1-p_1}, h_0, \ldots, h_{1-p_2})' \in \mathbb{R}^{p_1} \times [0, \infty)^{p_2}$, is assumed to be fixed for the statistical analysis. The asymptotic results do not change if ς_0 is replaced by an arbitrarily chosen vector (e.g., by setting $Y_t = 0$ and $h_t = 0$, all $t \leq 0$); see, for example, the discussions in Straumann and Mikosch (2006), Perera and Silvapulle (2021) and Jensen and Rahbek (2004).

We propose an omnibus test statistic based on the marked empirical process:

$$\mathcal{U}_n(y, \boldsymbol{\phi}) := n^{-1/2} \sum_{t=1}^{n} \left\{ \frac{Y_t^2}{h_t(\boldsymbol{\phi})} - 1 \right\} \mathbb{I}(Y_{t-1} \leq y), \quad y \in \mathbb{R},\ \boldsymbol{\phi} \in \Phi, \tag{6}$$

where $\mathbb{I}$ denotes the indicator function, evaluated at $\boldsymbol{\phi} = \hat{\boldsymbol{\phi}}_n$. We allow the domain of $\mathcal{U}_n(\cdot, \boldsymbol{\phi})$ to extend over the whole real line by letting $\mathcal{U}_n(-\infty, \boldsymbol{\phi}) := 0$ and $\mathcal{U}_n(\infty, \boldsymbol{\phi}) := n^{-1/2} \sum_{t=1}^{n} \{Y_t^2/h_t(\boldsymbol{\phi}) - 1\}$. Hence, $\mathcal{U}_n(\cdot, \boldsymbol{\phi})$ in (6) can be viewed as a process in the space of *cadlag* functions on $[-\infty, \infty]$, equipped with the uniform metric, which we denote by $\mathcal{D}(\mathbb{R})$. This process is an extension of the so-called cumulative sum process for the one sample setting to the current setup. Under H_0, $\mathrm{E}\,\mathcal{U}_n(y, \boldsymbol{\phi}_0) = 0$, for all y, but not under H_1. Hence, if H_0 is true, then we would expect $\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n)$ to be close to zero for all y, but not otherwise. Therefore, a suitable functional of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ can potentially be used as a test statistic for testing H_0 against H_1.

Example 1. (ARCH(1)) *As an example consider the ARCH(1) model as given by* (1) *with* $h_t = h_t(\boldsymbol{\phi}) = \omega + \alpha_1 Y_{t-1}^2$, *where* $\boldsymbol{\phi} = (\omega, \alpha_1)' \in \Phi \subset (0, \infty) \times [0, \infty)$. *In this case, the process in* (6) *is given by*

$$\mathcal{U}_n(y, \boldsymbol{\phi}) = n^{-1/2} \sum_{t=1}^{n} \left\{ Y_t^2 / \left(\omega + \alpha_1 Y_{t-1}^2\right) - 1 \right\} \mathbb{I}\left(Y_{t-1} \leq y\right).$$

With $\hat{\boldsymbol{\phi}}_n = (\hat{\omega}, \hat{\alpha}_1)'$ *the test statistic we discuss below is based on* $\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n)$, *which with* $\widehat{\varepsilon}_t = Y_t / \sqrt{\hat{\omega} + \hat{\alpha}_1 Y_{t-1}^2}$ *becomes* $\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n) = n^{-1/2} \sum_{t=1}^{n} \left\{ \widehat{\varepsilon}_t^2 - 1 \right\} \mathbb{I}\left(Y_{t-1} \leq y\right)$.

The use of cumulative sum processes for specification testing similar to $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ goes back to von Neumann (1941), who proposed a test of constant regression based on an analog of this process. Substantial developments have taken place of hypothesis testing

in time series models based on such empirical processes marked by certain *residuals*, see, e.g, Stute (1997), Koul and Stute (1999), Stute *et al.* (2006), Escanciano (2007) and Koul *et al.* (2012). More recently, analogs of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ have been used by several authors to propose asymptotically distribution free specification tests in related time series models; see, for example, Perera and Koul (2017) and Balakrishna *et al.* (2019). In the analyses presented in these papers certain tests based on analogs of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ have demonstrated desirable finite sample and asymptotic properties. Therefore, we find it of interest to develop specification tests based on similar statistics involving the process $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$. In particular, we consider the Kolmogorov-Smirnov (KS) and Cramér-von Mises (CvM) type statistics which can be defined in terms of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ as:

$$T_1 := \mathrm{KS} = \sup_y \left|\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n)\right|, \quad T_2 := \mathrm{CvM} = \int \mathcal{U}_n^2(y, \hat{\boldsymbol{\phi}}_n) dG_n(y), \tag{7}$$

where $G_n(y) := n^{-1} \sum_{t=1}^n \mathbb{I}(Y_{t-1} \leq y)$. Other suitable functionals of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ may also be considered as possible test statistics (see D'Agostino and Stephens, 1986).

3. Inference when the parameters are in the interior of the parameter space

Before moving to the general case which includes possible parameters on the boundary of the parameter space, we here consider the case the true parameter is in the interior of Φ.

The asymptotic distribution of $\mathcal{U}_n(\cdot, \boldsymbol{\phi}_0)$ under the null hypothesis H_0 can be derived by standard arguments, under the assumptions on the GARCH process discussed in the next subsection. Specifically, from a martingale central limit theorem (for example Hall and Heyde (1980), Corollary 3.1) and the Cramér-Wold device it follows that all finite dimensional distributions of $\mathcal{U}_n(\cdot, \boldsymbol{\phi}_0)$ converge weakly to a multivariate normal distribution with mean vector zero and covariance matrix given by the covariance function

$$K(x, y) := \mathrm{E}(\varepsilon_t^2 - 1)^2 \mathbb{I}(Y_{t-1} \leq x \wedge y) = (\kappa_\varepsilon - 1) G(x \wedge y), \quad x, y \in \mathbb{R}, \tag{8}$$

where G denotes the (unconditional) distribution function (d.f.) of Y_0, $\kappa_\varepsilon := \mathrm{E}\varepsilon_t^4 < \infty$ and $x \wedge y = \min(x, y)$. Under H_0, G may depend on $\boldsymbol{\phi}_0$, which is suppressed in the notation. Then, since the function $\pi(x) := K(x, x) = (\kappa_\varepsilon - 1) G(x)$ is nondecreasing and nonnegative, tightness of the process $\mathcal{U}_n(\cdot, \boldsymbol{\phi}_0)$ follows. Therefore, under H_0, $\mathcal{U}_n(\cdot, \boldsymbol{\phi}_0)$ converges weakly to the time-transformed Brownian motion $B \circ \pi(\cdot) := B(\pi(\cdot))$, in the space $\mathcal{D}(\mathbb{R})$ equipped with the uniform metric, where B is a standard Brownian motion on the positive real line.

However, since $\boldsymbol{\phi}_0$ is replaced by $\hat{\boldsymbol{\phi}}$, the weak limit of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}})$ is not of the form $B \circ \pi$; rather, it depends on $(\boldsymbol{\phi}_0, G)$. We derive this result in the next subsection, where weak convergence of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}})$ is derived for the case where the true value $\boldsymbol{\phi}_0$ lies in the interior of the parameter space.

3.1. Asymptotics for the original test statistics

First we introduce some notation to facilitate the presentation of the underlying assumptions for the asymptotic results. Let $\mathcal{A}_{\boldsymbol{\phi}}(z) = \sum_{i=1}^{p_1} \alpha_i z^i$ and $\mathcal{B}_{\boldsymbol{\phi}}(z) = 1 - \sum_{i=1}^{p_2} \beta_i z^i$ with $\mathcal{A}_{\boldsymbol{\phi}}(z) = 0$ if $p_1 = 0$ and $\mathcal{B}_{\boldsymbol{\phi}}(z) = 1$ if $p_2 = 0$. Furthermore, let

$$A_{0i} = \begin{pmatrix} \alpha_{01}\varepsilon_i^2 & \cdots & \alpha_{0p_1}\varepsilon_i^2 & \beta_{01}\varepsilon_i^2 & \cdots & \beta_{0p_2}\varepsilon_i^2 \\ & \boldsymbol{I}_{p_1-1} & \mathbf{0} & & \mathbf{0} & \\ \alpha_{01} & \cdots & \alpha_{0p_1} & \beta_{01} & \cdots & \beta_{0p_2} \\ & \mathbf{0} & & & \boldsymbol{I}_{p_2-1} & \mathbf{0} \end{pmatrix}, \quad i \geq 1,$$

with $\boldsymbol{I}_k$ denoting the $k \times k$ identity matrix.

In order to study the limiting behaviour of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ we make the following assumptions on the process $\{Y_t\}_{t\in\mathbb{Z}}$ which satisfies (1)–(2).

(A1). *The parameter space Φ in equation (3) is a compact subset of $(0,\infty) \times [0,\infty)^{p_1+p_2}$, and contains a hypercube $\bar{\Phi} := [\omega_L, \omega_U] \times [0, \epsilon]^{p_1+p_2}$, for some $\epsilon > 0$ and $\omega_U > \omega_L > 0$, which includes $\boldsymbol{\phi}_0$ with $\omega_L < \alpha_0 < \omega_U$ and $\max\{\alpha_{01}, \ldots, \alpha_{0p_1}, \beta_{01}, \ldots, \beta_{0p_2}\} < \epsilon$.*

(A2). *The sequence of matrices $\boldsymbol{A}_0 = (A_{01}, A_{02}, \ldots)$ has a strictly negative top Lyapunov exponent; i.e., $\gamma(\boldsymbol{A}_0) < 0$, where $\gamma(\boldsymbol{A}_0) := \lim_{i\to\infty} i^{-1} \log \|A_{0i}A_{0(i-1)} \ldots A_{01}\|$ a.s., and $\sum_{j=1}^{p_2} \beta_j < 1$, $\forall \boldsymbol{\phi} \in \Phi$.*

(A3). *If $p_2 > 0$, then assume that $\mathcal{A}_{\boldsymbol{\phi}_0}(1) \neq 0$, $\alpha_{0p_1} + \beta_{0p_2} \neq 0$, and the polynomials $\mathcal{A}_{\boldsymbol{\phi}_0}(z)$ and $\mathcal{B}_{\boldsymbol{\phi}_0}(z)$ have no common roots.*

(A4). *The errors $\varepsilon_t, t \in \mathbb{Z}$, are i.i.d. with zero mean and unit variance, ε_t^2 has a non-degenerate distribution, $\mathrm{E}|\varepsilon_t|^{4+d} < \infty$ for some $d > 0$.*

Assumption (A1) ensures that the true parameter $\boldsymbol{\phi}_0$ does not reach the upper bound of the hypercube $(0,\infty) \times [0,\infty)^{p_1+p_2}$ while allowing some components of $\boldsymbol{\phi}_0$ to be zero. The condition $\gamma(\boldsymbol{A}_0) < 0$ in (A2) ensures the existence of a unique strictly stationary solution $\{Y_t\}_{t\in\mathbb{Z}}$ to Model (1)–(2); see, e.g. Bougerol and Picard (1992a). Note that, in (A2), the

strict stationarity condition $\gamma(\boldsymbol{A}_0) < 0$ is imposed only on the true value $\boldsymbol{\phi}_0$, but for $\boldsymbol{\phi} \neq \boldsymbol{\phi}_0$ we only impose the weaker restriction $\sum_{j=1}^{p_2} \beta_j < 1$. In Assumption (A3), the condition $\mathcal{A}_{\boldsymbol{\phi}_0}(1) \neq 0$ rules out the case where all the α_{0i} are zero when $p_2 > 0$; hence, we do not allow the strictly stationary solution of (1)–(2) to be a strong white noise process when $p_2 > 0$. This restriction is required to avoid certain identifiability issues when estimating the GARCH parameters with $p_2 > 0$ (see Francq and Zakoïan, 2004). Note that, in the ARCH case (i.e. when $p_2 = 0$), Assumption (A3) is not required. For example, suppose that $p_2 = 0$ and we start with an ARCH(p_1) model. Then, Assumption (A3) does not apply, and hence it is possible to have $\mathcal{A}_{\boldsymbol{\phi}_0}(1) = 0$. Therefore, if $p_2 = 0$, our assumptions allow the strictly stationary solution of (1)–(2) to be a strong white noise process.

In the general GARCH case, when $p_2 > 0$, the condition $\alpha_{0p_1} + \beta_{0p_2} \neq 0$ in Assumption (A3) allows for overidentification of either the order of the ARCH parameters p_1 or the order of the GARCH parameters p_2, but not both. Here, *overidentification of the order* means having an order which is higher than what is required for the ARCH/GARCH parameters; e.g., specifying an ARCH(3) model when the DGP is ARCH(2). The condition $\mathrm{E}|\varepsilon_t|^{4+d} < \infty$ in Assumption (A4) is required for the existence of the variance of the score vector $\partial \ell_t(\boldsymbol{\phi}_0)/\partial \boldsymbol{\phi}$; this is necessary for establishing the limiting distribution of the QMLE. Note that we do not assume that the true parameter $\boldsymbol{\phi}_0$ is in the interior of Φ. Thus, the assumptions do not exclude the cases where some α_i or β_j are zero. Assumptions similar to (A1)–(A4) have previously been discussed in the literature for establishing asymptotic properties of the QMLE; see, e.g., Francq and Zakoïan (2010) and Cavaliere *et al.* (2022).

Let

$$J(y, \boldsymbol{\phi}) := \mathrm{E}[\tau_1(\boldsymbol{\phi})\mathbb{I}(Y_0 \leq y)], \quad \tau_t(\boldsymbol{\phi}) := \frac{(\partial/\partial\boldsymbol{\phi})h_t(\boldsymbol{\phi})}{h_t(\boldsymbol{\phi})}, \quad t \in \mathbb{Z},\ \boldsymbol{\phi} \in \Phi.$$

Note that when some components of $\boldsymbol{\phi}_0$ lie on the boundary of the parameter space Φ, we define $\tau_t(\boldsymbol{\phi})$ at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$ by using the right derivatives of $h_t(\boldsymbol{\phi})$ at $\boldsymbol{\phi}_0$, denoted by $(\partial/\partial\boldsymbol{\phi})h_t(\boldsymbol{\phi}_0) := (\partial h_t(\boldsymbol{\phi}_0)/\partial\phi_i)_{i=1,\ldots,p_1+p_2+1}$ the vector of partial derivatives of h_t at $\boldsymbol{\phi}_0 = (\phi_{01}, \ldots, \phi_{0(1+p_1+p_2)})'$ with the i-th derivative replaced by the right derivative when $\phi_{0i} = 0$. We use the same convention for the derivatives of $\tau_t(\boldsymbol{\phi})$ and $(\partial/\partial\boldsymbol{\phi})h_t(\boldsymbol{\phi})$ at $\boldsymbol{\phi}_0$.

The next lemma provides an asymptotic uniform expansion for $\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n)$. We make use of this expansion in the proof of establishing the weak convergence of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$.

Lemma 1. *Suppose that Assumptions (A1) and (A4) hold. Then, uniformly in $y \in \mathbb{R}$,*

$$\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n) = \mathcal{U}_n(y, \boldsymbol{\phi}_0) - n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)' J(y, \boldsymbol{\phi}_0) + o_p(1). \tag{9}$$

Unlike the process $\mathcal{U}_n(y, \boldsymbol{\phi}_0)$, the estimated process $\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n)$ does not converge weakly to a time transformed Brownian motion, because the term $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)' J(y, \boldsymbol{\phi}_0)$ in (9), is of order $O_p(1)$ and hence is not asymptotically negligible. In fact, if Assumptions (A1)–(A3) are satisfied, then $\hat{\boldsymbol{\phi}}_n$ converges to $\boldsymbol{\phi}_0$ almost surely (a.s.), and additionally, if (A4) also holds and $\boldsymbol{\phi}_0$ is an interior point in Φ, then $\hat{\boldsymbol{\phi}}_n$ is asymptotically linear and satisfies

$$n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) = -\Sigma_n^{-1}(\boldsymbol{\phi}_0) n^{-1/2} \sum_{t=1}^{n} (1 - \varepsilon_t^2) \tau_t(\boldsymbol{\phi}_0) + o_p(1), \tag{10}$$

where $\Sigma_n(\boldsymbol{\phi}) := n^{-1} \sum_{t=1}^{n} \tau_t(\boldsymbol{\phi}) \tau_t(\boldsymbol{\phi})', \ \boldsymbol{\phi} \in \Phi$; see, e.g. Berkes *et al.* (2003).

By using Lemma 1 and (10), when $\boldsymbol{\phi}_0$ is an interior point in Φ, we show that $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ converges weakly to a centred Gaussian process. This result is stated in the next theorem.

THEOREM 1. *Suppose that (A1)–(A4) are satisfied and $\boldsymbol{\phi}_0$ is an interior point in Φ. Let*

$$M_t(\boldsymbol{\phi}) := -\Sigma^{-1}(\boldsymbol{\phi})(1 - \varepsilon_t^2)\tau_t(\boldsymbol{\phi}), \quad \Sigma(\boldsymbol{\phi}) := \mathrm{E}\{\tau_1(\boldsymbol{\phi})\tau_1(\boldsymbol{\phi})'\}, \quad \boldsymbol{\phi} \in \Phi, \ t \in \mathbb{Z}.$$

Additionally, assume that $\{Y_t\}_{t \in \mathbb{Z}}$ is square integrable. Then, the process $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ converges weakly to $\mathcal{U}_0$ in $\mathcal{D}(\mathbb{R})$, where $\mathcal{U}_0$ is a centred Gaussian process with covariance kernel

$$\begin{aligned} \mathrm{Cov}\{\mathcal{U}_0(x), \mathcal{U}_0(y)\} &= K(x, y) + J'(x, \boldsymbol{\phi}_0)\mathrm{E}[M_1(\boldsymbol{\phi}_0)M_1'(\boldsymbol{\phi}_0)]J'(y, \boldsymbol{\phi}_0) \\ &\quad - J'(x, \boldsymbol{\phi}_0)\mathrm{E}[(\varepsilon_1^2 - 1)M_1(\boldsymbol{\phi}_0)\mathbb{I}(Y_0 \le y)] \\ &\quad - J'(y, \boldsymbol{\phi}_0)\mathrm{E}[(\varepsilon_1^2 - 1)M_1(\boldsymbol{\phi}_0)\mathbb{I}(Y_0 \le x)], \end{aligned}$$

where $K(x, y)$ is as in (8).

In view of Theorem 1, the limiting distributions of KS and CvM statistics defined in (7) depend on the unknown $(\boldsymbol{\phi}_0, G)$ in a non-trivial way, despite the fact the true parameter is in the interior of Φ. Consequently, it does not appear that it would be possible to find a transformation that would lead to an asymptotically distribution free test, for example as in Bai (2003); Koul *et al.* (2012); Perera and Koul (2017); Escanciano *et al.* (2018). Hence, we proceed by considering bootstrap implementations of the tests.

3.2. BOOTSTRAP IMPLEMENTATION

In this section, we propose a bootstrap procedure for computing the critical values for the KS and CvM statistics in (7). We perform the resampling scheme under the null hypothesis and derive the asymptotic properties of the bootstrap statistics, irrespective of whether or

not the data generating process satisfies the null hypothesis. To this end, we initially standardize the residuals $\widehat{\varepsilon}_t := Y_t/\{h_t(\hat{\boldsymbol{\phi}}_n)\}^{1/2},\ t = 1, \ldots, n$, as

$$\check{\varepsilon}_t := \left\{ n^{-1} \sum_{i=1}^{n} \bar{\varepsilon}_i^2 \right\}^{-1/2} \bar{\varepsilon}_t, \quad \bar{\varepsilon}_t := \widehat{\varepsilon}_t - n^{-1} \sum_{i=1}^{n} \widehat{\varepsilon}_i, \quad t = 1, \ldots, n, \tag{11}$$

and define the associated empirical distribution function (e.d.f.) of $\{\check{\varepsilon}_1, \ldots, \check{\varepsilon}_n\}$ as

$$\check{F}_n(x) := n^{-1} \sum_{t=1}^{n} \mathbb{I}(\check{\varepsilon}_t \leq x), \quad x \in \mathbb{R}. \tag{12}$$

By construction, $\int_{\mathbb{R}} u d\check{F}_n(u) = 0$ and $\int_{\mathbb{R}} u^2 d\check{F}_n(u) = 1$, hence a random variable with distribution function $\check{F}_n$ has zero mean and unit variance, therefore matching the first and second order moments of the error distribution F_0. From Lemma S.1 in the supplementary material we obtain that $\check{F}_n$ converges to F_0 with probability one under the null hypothesis.

We next outline the bootstrap algorithm.

Bootstrap Algorithm 1

Step 1: Compute $\{\hat{\boldsymbol{\phi}}_n, T_j\}$ on the original sample $\{Y_1, \ldots, Y_n\}$, where T_j is the test statistic defined in (7) ($j = 1, 2$);

Step 2: Compute $\check{\varepsilon}_t$, $t = 1, \ldots, n$ as in (11) and draw a random sample (with replacement) of size n, say $\{\varepsilon_1^*, \ldots, \varepsilon_n^*\}$, independent of the original data, from the e.d.f. $\check{F}_n(\cdot)$ in (12);

Step 3: Generate the bootstrap sample $\{Y_1^*, \ldots, Y_n^*\}$ recursively with bootstrap true values $(\hat{\boldsymbol{\phi}}_n, \check{F}_n)$ by

$$Y_t^* = \{h_t^*(\hat{\boldsymbol{\phi}}_n)\}^{1/2} \varepsilon_t^*, \quad h_t^*(\hat{\boldsymbol{\phi}}_n) = \hat{\omega} + \sum_{j=1}^{p_1} \hat{\alpha}_j (Y_{t-j}^*)^2 + \sum_{k=1}^{p_2} \hat{\beta}_k h_{t-k}^*(\hat{\boldsymbol{\phi}}_n), \quad t \geq 1,$$

initialized with $(Y_0^*, \ldots, Y_{1-q}^*, h_0^*(\hat{\boldsymbol{\phi}}_n), \ldots, h_{1-p}^*(\hat{\boldsymbol{\phi}}_n))' = \varsigma_0$, where ς_0 is an arbitrarily chosen vector (e.g. $Y_t^* = 0$ and $h_t^* = 0$, all $t \leq 0$);

Step 4: Using $\{Y_1^*, \ldots, Y_n^*\}$, compute $\hat{\boldsymbol{\phi}}_n^*$, the bootstrap analogue of $\hat{\boldsymbol{\phi}}_n$;

Step 5: Compute the bootstrap test statistic T_j^* as

$$T_1^* = \mathrm{KS}^* = \sup_y \left| \mathcal{U}_n^*(y, \hat{\boldsymbol{\phi}}_n^*) \right|, \quad T_2^* = \mathrm{CvM}^* = \int \left\{ \mathcal{U}_n^*(y, \hat{\boldsymbol{\phi}}_n^*) \right\}^2 dG_n^*(y), \tag{13}$$

where $G_n^*(y)$ and $\mathcal{U}_n^*(y, \boldsymbol{\phi})$ are the bootstrap analogs of $G_n(y)$ and $\mathcal{U}_n(y, \boldsymbol{\phi})$, respectively.

Step 6: The bootstrap p-value is then defined as

$$p_n^* := p_n^*(\hat{\boldsymbol{\phi}}_n) := P_n^*(T_j^* \geq T_j) \tag{14}$$

where P_n^* denotes the probability measure induced by the bootstrap (i.e., conditional on the original data). $\square$

The bootstrap test corresponds to the decision rule:

$$\text{Reject } \mathsf{H}_0 \text{ at the nominal level } \alpha \text{ if the bootstrap p-value } p_n^* \text{ is less than } \alpha. \tag{15}$$

As is standard, p_n^* of (14) is unknown. It can be approximated with arbitrary accuracy by repeating steps 2–5 a large number of times, say B, and then setting $p_n^{*(B)}$ to be the fraction of times T_j^* exceeds T_j.

The above bootstrap algorithm is designed to mimic the null data generating process (DGP) by replacing the unknown $(\boldsymbol{\phi}_0, F_0)$ by the estimators $(\hat{\boldsymbol{\phi}}_n, \check{F}_n)$. It can be shown that, under the null hypothesis H_0, the estimators $(\hat{\boldsymbol{\phi}}_n, \check{F}_n)$ converge to $(\boldsymbol{\phi}_0, F_0)$ almost surely. Therefore, if $\boldsymbol{\phi}_0$ is an interior point of the parameter space, then the arguments of the proof of Theorem 1 can be extended to a triangular array setup to show that the bootstrap test in (15) is asymptotically valid (see Theorem 2 below).

In the bootstrap setup, we define $G_n^*(y) := n^{-1}\sum_{t=1}^n \mathbb{I}(Y_{t-1}^* \le y)$, $y \in \mathbb{R}$. Similarly, the bootstrap analogue of the marked empirical process $\mathcal{U}_n(y, \boldsymbol{\phi})$ in (6) is defined by

$$\mathcal{U}_n^*(y, \boldsymbol{\phi}) := n^{-1/2}\sum_{t=1}^n \left\{\frac{(Y_t^*)^2}{h_t^*(\boldsymbol{\phi})} - 1\right\}\mathbb{I}(Y_{t-1}^* \le y), \quad y \in \mathbb{R},\ \boldsymbol{\phi} \in \Phi. \tag{16}$$

Let $O_p^*(1)$, in probability, $o_p^*(1)$, in probability, and E^* denote the usual stochastic orders of magnitude and expectation, respectively, with respect to P_n^* defined above. We denote convergence in distribution of bootstrap statistics as '$\xrightarrow{d^*}$'. That is, '$T_j^* \xrightarrow{d^*} g_j(\mathcal{U}_0)$ *in probability*' means that $P_n^*(T_j^* \le \cdot) \xrightarrow{p} P\{g_j(\mathcal{U}_0) \le \cdot\}$, at every continuity point of $P\{g_j(\mathcal{U}_0) \le \cdot\}$.

Theorem 2 below establishes the asymptotic validity of the bootstrap test (15).

THEOREM 2. *Suppose that Assumptions (A1)–(A4) and H_0 are satisfied and $\boldsymbol{\phi}_0$ is an interior point in Φ. Additionally, assume that the process $\{Y_t\}_{i\in\mathbb{Z}}$ is square integrable. Let $\mathcal{U}_0$ be the limit process appearing in Theorem 1. Then, conditional on $\{Y_1, \ldots, Y_n\}$,*

1. *$\mathcal{U}_n^*(\cdot, \hat{\boldsymbol{\phi}}_n^*)$ converges weakly to $\mathcal{U}_0$, in probability.*
2. *$T_1^* = \mathrm{KS}^* \xrightarrow{d^*} \sup_y \left|\mathcal{U}_0(y)\right|$ and $T_2^* = \mathrm{CvM}^* \xrightarrow{d^*} \int \mathcal{U}_0^2(y)dG(y)$, in probability.*

In view of Theorem 2, the bootstrap test (15) based on T_j is asymptotically valid ($j = 1, 2$). Theorem 3 in the next section shows that the bootstrap tests have non-trivial asymptotic power against any fixed alternative under H_1.

For the validity of our bootstrap tests, as stated in Theorem 2, the true parameter $\boldsymbol{\phi}_0$ needs to be an interior point of Φ under H_0. It is of interest to see whether the bootstrap implementation of $T_j (j = 1, 2)$ can be modified to obtain a consistent bootstrap test for the case $\boldsymbol{\phi}_0$ lies on the boundary of Φ under H_0. We consider this in the next section.

3.3. The behavior of the bootstrap test under the alternative hypothesis

Note that the bootstrap algorithm in Section 3.2 mimics the null data generating process by replacing the unknown $(\boldsymbol{\phi}_0, F_0)$ by the estimators $(\hat{\boldsymbol{\phi}}_n, \check{F}_n)$. Under the null hypothesis H_0, $(\hat{\boldsymbol{\phi}}_n, \check{F}_n)$ converges to $(\boldsymbol{\phi}_0, F_0)$ almost surely. This result is used in the proof of Theorem 2 to show that, when $\boldsymbol{\phi}_0$ is an interior point under H_0, the standard bootstrap test in (15) has correct asymptotic level. In order to establish that the standard bootstrap test also has non-trivial asymptotic power, we need to assume that the limiting behaviour of $\hat{\boldsymbol{\phi}}_n$ and $\check{F}_n$, under the null as well as the alternative satisfy the usual regularity conditions. To this end, let $(\boldsymbol{\phi}_0^*, F_0^*)$ be the probability limit of $(\hat{\boldsymbol{\phi}}_n, \check{F}_n)$, such that $\hat{\boldsymbol{\phi}}_n \xrightarrow{p} \boldsymbol{\phi}_0^*$ and $d_2(\check{F}_n, F_0^*) \xrightarrow{p} 0$ as $n \to \infty$, where $d_2(F_X, F_Y)$ is the Mallows metric for the distance between two probability distributions F_X and F_Y defined by $d_2(F_X, F_Y) = \inf\{\mathrm{E}|X - Y|^2\}^{1/2}$, where the infimum is over all square integrable random variables X and Y with marginal distributions F_X and F_Y. Since $(\boldsymbol{\phi}_0^*, F_0^*) := \mathrm{plim}(\hat{\boldsymbol{\phi}}_n, \check{F}_n)$, where 'plim' is the probability limit as $n \to \infty$, under the null hypothesis H_0, we have that $(\boldsymbol{\phi}_0^*, F_0^*) = (\boldsymbol{\phi}_0, F_0)$ and under the alternative hypothesis H_1, $(\boldsymbol{\phi}_0^*, F_0^*)$ denotes a pseudo-true value, when the plim exists.

We also need to introduce the following regularity assumption under H_1.

(B1). *Under the alternative hypothesis H_1, the process $\{Y_t\}_{t\in\mathbb{Z}}$ is second order stationary and obeys model (1); moreover, $\hat{\boldsymbol{\phi}}_n \xrightarrow{p} \boldsymbol{\phi}_0^*$ and $d_2(\check{F}_n, F_0^*) \xrightarrow{p} 0$, where F_0^* is a c.d.f. with mean 0 and variance 1. The pseudo-true parameter $\boldsymbol{\phi}_0^* = (\omega_0^*, \alpha_{01}^*, \ldots, \alpha_{0p_1}^*, \beta_{01}^*, \ldots, \beta_{0p_2}^*)'$ satisfies $\sum_{i=1}^{p_1} \alpha_{0i}^* + \sum_{j=1}^{p_2} \beta_{0j}^* < 1$. Furthermore, for some $\epsilon > 0$ and $\omega_U > \omega_L > 0$, the hypercube $\bar{\Phi} = [\omega_L, \omega_U] \times [0, \epsilon]^{p_1+p_2}$ given in Assumption (A1), contains $\boldsymbol{\phi}_0^*$.*

The strict stationarity of the process $\{Y_t : t \in \mathbb{Z}\}$ obeying (1)–(4), which follows from (A1), (A2) and (A3), and its square integrability (as assumed under Theorem 2), ensures that the true parameter $\boldsymbol{\phi}_0 = (\omega_0, \alpha_{01}, \ldots, \alpha_{0p_1}, \beta_{01}, \ldots, \beta_{0p_2})'$ under the null hypothesis H_0 satisfies $\sum_{i=1}^{p_1} \alpha_{0i} + \sum_{j=1}^{p_2} \beta_{0j} < 1$ (see Bougerol and Picard, 1992a,b). Assumption (B1) assumes that this continues to hold when $\boldsymbol{\phi}_0^*$ is the pseudo true value under the alternative.

Example 2. (ARCH(1), continued) *Consider again the ARCH(1) example. Under H_0, it is known that* $\mathrm{plim}\, \hat{\boldsymbol{\phi}}_n = \boldsymbol{\phi}_0 = (\omega_0, \alpha_{01})'$ *under Assumptions (A1)–(A3). In particular,*

$\boldsymbol{\phi}_0 \in [\omega_L, \omega_U] \times [0, \epsilon]$, $\alpha_{10} < \exp\{-\mathrm{E}\log(\varepsilon_t^2)\}$ *and* ε_t *is an i.i.d. sequence of r.v.'s. Consider now the alternative of an ARCH(2),* $h_t(\boldsymbol{\phi}_0) = \omega_0 + \alpha_{10} Y_{t-1}^2 + \alpha_{20} Y_{t-2}^2$, *with* $\alpha_{20} > 0$ *and* $\alpha_{10} + \alpha_{20} < 1$. *The restriction* $\alpha_{10} + \alpha_{20} < 1$ *is required to ensure that the observable process* $\{Y_t : t \in \mathbb{Z}\}$ *is square integrable. In this case Assumption (B1) concerns the large-sample behaviour of* $\hat{\boldsymbol{\phi}}_n$*; specifically,* $\operatorname{plim} \hat{\boldsymbol{\phi}}_n = \boldsymbol{\phi}_0^* = (\omega_0^*, \alpha_{01}^*)'$. *It can be shown that*

$$\operatorname{plim} \hat{\alpha}_1 = \alpha_{01}^* = \frac{\alpha_{10}}{1 - \alpha_{20}} \text{ and } \operatorname{plim} \hat{\omega} = \omega_0^* = \frac{\omega_0}{1 - \alpha_{20}}. \tag{17}$$

Note that, as required $0 < \alpha_{01}^* < 1$ *since* $\alpha_{10} + \alpha_{20} < 1$. *Furthermore, one can select* $\epsilon > 0$ *and* $\omega_U > \omega_L > 0$ *such that* $[\omega_L, \omega_U] \times [0, \epsilon]$ *also contains* $\boldsymbol{\phi}_0^* = (\omega_0^*, \alpha_{01}^*)'$ *as we have* $\omega_0^* > 0$.

To study the large sample behaviour of $\check{F}_n$ *(to keep things simple leave out the mean and scale correction here) consider* $\check{F}_n(x) = n^{-1} \sum_{t=1}^n \mathbb{I}(\widehat{\varepsilon}_t \le x)$, $\widehat{\varepsilon}_t = Y_t / \left(\hat{\omega} + \hat{\alpha}_1 Y_{t-1}^2\right)^{1/2}$. *Note that, under* H_1,

$$\widehat{\varepsilon}_t = \frac{Y_t}{\sqrt{\hat{\omega} + \hat{\alpha}_1 Y_{t-1}^2}} \simeq \frac{\sqrt{1 - \alpha_{20}} Y_t}{\sqrt{\omega_0 + \alpha_{10} Y_{t-1}^2}} =: \varepsilon_t^0.$$

Let F_0^* *be the c.d.f. of* ε_t^0 *and* $H_n^0(x) := n^{-1} \sum_{t=1}^n \mathbb{I}(\varepsilon_t^0 \le x)$. *From the triangular inequality* $d_2(\check{F}_n, F_0^*) \le d_2(\check{F}_n, H_n^0) + d_2(H_n^0, F_0^*)$. *Since* $d_2(H_n^0, F_0^*) \overset{a.s.}{\to} 0$ *as* $n \to \infty$ *(see, for example, Lemma 8.4 of Bickel and Freedman, 1981), to have* $d_2(\check{F}_n, F_0^*) \overset{p}{\to} 0$ *it suffices to show that* $d_2(\check{F}_n, H_n^0) \overset{p}{\to} 0$. *By arguing as in the proof of Lemma S.1 in the supplementary material,*

$$\begin{aligned} \{d_2(\check{F}_n, H_n^0)\}^2 &\le n^{-1} \sum_{t=1}^n (\varepsilon_t^0 - \widehat{\varepsilon}_t)^2 = n^{-1} \sum_{t=1}^n \left(\frac{\sqrt{1 - \alpha_{20}} Y_t}{\sqrt{\omega_0 + \alpha_{10} Y_{t-1}^2}} - \frac{Y_t}{\sqrt{\hat{\omega} + \hat{\alpha}_1 Y_{t-1}^2}} \right)^2 \\ &\le \omega_L^{-2} n^{-1} \sum_{t=1}^n \left(\sqrt{1 - \alpha_{20}} \sqrt{\hat{\omega} + \hat{\alpha}_1 Y_{t-1}^2} - \sqrt{\omega_0 + \alpha_{10} Y_{t-1}^2} \right) Y_t^2 = o_p(1) \end{aligned}$$

by (17) and because $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0^*) = O_p(1)$ *and* $\mathrm{E}(Y_t^2) < \infty$. *Therefore,* $d_2(\check{F}_n, F_0^*) \overset{p}{\to} 0$, *as required under Assumption (B1).*

In the general case, if H_1 holds under (B1), then regardless of whether $\boldsymbol{\phi}_0^*$ is in the interior of Φ or some components of $\boldsymbol{\phi}_0^*$ are zero, one can fix $\omega_U > \omega_L > 0$ and $\epsilon_k > 0$, $k = 1, \ldots, p_1 + p_2$, such that

$$\bar{\Phi}^* := [\omega_L, \omega_U] \times [0, \epsilon_1] \times [0, \epsilon_2] \times \cdots \times [0, \epsilon_{p_1+p_2}] \subset \bar{\Phi} \tag{18}$$

contains $\boldsymbol{\phi}_0^* = (\omega_0^*, \alpha_{01}^*, \ldots, \alpha_{0p_1}^*, \beta_{01}^*, \ldots, \beta_{0p_2}^*)'$ with $\omega_L < \omega_0^* < \omega_U$, $\mathbf{0} \le \boldsymbol{\phi}_0^* < \bar{\boldsymbol{\phi}}_U := (\omega_U, \epsilon_1, \ldots, \epsilon_{p_1+p_2})'$, $\sum_{i=1}^{p_1} \alpha_i + \sum_{j=1}^{p_2} \beta_j < 1$ for every $\boldsymbol{\phi} = (\omega, \alpha_1, \ldots, \alpha_{p_1}, \beta_1, \ldots, \beta_{p_2})' \in \bar{\Phi}^*$.

Fix such $\epsilon_1, \ldots, \epsilon_{p_1+p_2} > 0$ and $\omega_U > \omega_L > 0$. Then, for every $\boldsymbol{\phi} \in \bar{\Phi}^*$ and c.d.f. F (with mean 0 and variance 1), the model defined by

$$Y_t^{(\boldsymbol{\phi},F)} = \{h_t^{(\boldsymbol{\phi},F)}(\boldsymbol{\phi})\}^{1/2}\varepsilon_t^{(F)}, \quad h_t^{(\boldsymbol{\phi},F)}(\boldsymbol{\phi}) = \omega + \sum_{j=1}^{p_1}\alpha_j\{Y_{t-j}^{(\boldsymbol{\phi},F)}\}^2 + \sum_{j=1}^{p_2}\beta_j h_{t-j}^{(\boldsymbol{\phi},F)}(\boldsymbol{\phi}), \quad (19)$$

has a unique strictly stationary and ergodic solution with $\mathrm{E}[\{Y_0^{(\phi,F)}\}^2] < \infty$, where $\varepsilon_t^{(F)} = F^{-1}(U_t) := \inf\{y \in \mathbb{R} : F(y) \geq U_t\}$ and $\{U_t, t \in \mathbb{Z}\}$ are i.i.d. uniform(0,1) random variables; see, e.g. Theorem 2.1 of Chen and An (1998). If (A1)–(A3) and H_0 hold and $\{Y_t\}_{i\in\mathbb{Z}}$ is square integrable (as assumed under Theorem 2), $\sum_{i=1}^{p_1}\alpha_{0i} + \sum_{j=1}^{p_2}\beta_{0j} < 1$, and hence w.l.o.g. we assume that the set $\bar{\Phi}^*$ in (18) also contains $\boldsymbol{\phi}_0$ with $\omega_L < \omega_0 < \omega_U$ and $0 \leq \boldsymbol{\phi}_0 < \bar{\boldsymbol{\phi}}_U$. For $(\boldsymbol{\phi}, F) = (\boldsymbol{\phi}_0, F_0)$ the model (19) is equivalent to the null DGP: (1)–(4).

We also need to introduce the following additional assumption.

(B2). $\mathrm{E}(\mid F_0^{*-1}(U_t) \mid^{4+d}) < \infty$ *for some* $d > 0$.

This condition extends the moment restriction in Assumption (A4) to the model defined by the pseudo true values $(\boldsymbol{\phi}_0^*, F_0^*)$ under H_1.

Example 3. (ARCH(1), continued) *To understand the nature of condition (B2) consider again the ARCH(1) example. Suppose that the alternative model is ARCH(2), where* $h_t(\boldsymbol{\phi}_0) = \omega_0 + \alpha_{10}Y_{t-1}^2 + \alpha_{20}Y_{t-2}^2$, *with* $\alpha_{20} > 0$ *and* $\alpha_{10} + \alpha_{20} < 1$. *Then, under the alternative,* F_0^* *is the c.d.f. of* $\varepsilon_t^0 = Y_t\{1-\alpha_{20}\}^{1/2}\{\omega_0 + \alpha_{10}Y_{t-1}^2\}^{-1/2}$ *(see Example 2). Therefore, one can view Assumption (B2) essentially as a condition on the moments of* ε_t^0*; interestingly, this can be reverted into a (sufficient) condition on the moments of* Y_t *(as it usually happens when one analyzes the bootstrap under the alternative; see, e.g. Cavaliere* et al.*, 2017). More precisely, we obtain*

$$\mathrm{E}\left(\left|\frac{\sqrt{1-\alpha_{02}}Y_t}{\sqrt{\omega_0 + \alpha_{01}Y_{t-1}^2}}\right|^{4+d}\right) \leq \mathrm{E}\left(\left|\frac{\sqrt{1-\alpha_{02}}Y_t}{\sqrt{\omega_L}}\right|^{4+d}\right) = \left(\frac{1-\alpha_{02}}{\omega_L}\right)^{2+d/2}\mathrm{E}\left(|Y_t|^{4+d}\right).$$

Hence, the moment condition $\mathrm{E}(|Y_t|^{4+d}) < \infty$ *is sufficient for the Assumption (B2) to hold.*

The next assumption ensures that the moment condition (A6) introduced later is satisfied by the stationary solution of the model (19) for any given $\boldsymbol{\phi} \in \bar{\Phi}^*$ and c.d.f. F.

(A5). *For every* $\boldsymbol{\phi} \in \bar{\Phi}^*$ *and c.d.f.* F *(with mean 0 and variance 1), the unique stationary and ergodic solution* $\{Y_t^{(\boldsymbol{\phi},F)} : t \in \mathbb{Z}\}$ *of the model (19) satisfies* $\mathrm{E}|Y_t^{(\boldsymbol{\phi},F)}|^6 < \infty$.

Assumption (A5) allows us to extend the arguments of Lemma 2 when $(\boldsymbol{\phi}_0, F_0)$ is replaced by arbitrary but fixed $(\boldsymbol{\phi}, F)$, including $(\boldsymbol{\phi}_0^*, F_0^*)$ under H_1, with $\boldsymbol{\phi} \in \bar{\Phi}^*$. This condition is required in the proof of showing that, when some components of $\boldsymbol{\phi}_0^*$ are allowed to be zero, the limiting distribution of $n^{1/2}(\hat{\boldsymbol{\phi}}_n^* - \hat{\boldsymbol{\phi}}_n)$, conditional on $(Y_1, \ldots, Y_n)$, is of order $O_p^*(1)$, in probability, under H_1. This result is essential for establishing that the tests have non-trivial asymptotic power, when some components of $\boldsymbol{\phi}_0^*$ are allowed to be zero under H_1.

Theorem 3 below establishes the consistency of the bootstrap test (15).

THEOREM 3. *Suppose that* H_1 *holds and there exists a* $y \in \mathbb{R}$, *with* $h_t = \mathrm{E}(Y_t^2 \mid \mathcal{H}_{t-1})$, $t \in \mathbb{Z}$, *such that* $\mathrm{E}[\{h_1/h_1(\boldsymbol{\phi}_0^*) - 1\}\mathbb{I}(Y_0 \leq y)] \neq 0$. *Additionally, assume that Assumptions (B1) and (B2) hold and* $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0^*) = O_p(1)$. *Then, conditional on* $\{Y_1, \ldots, Y_n\}$, *(a) if* $\boldsymbol{\phi}_0^*$ *is an interior point in* Φ, *the bootstrap test (15) based on* T_j *has asymptotic power 1* $(j = 1, 2)$, *and (b) if additionally Assumption (A5) is also satisfied, then irrespective of whether* $\boldsymbol{\phi}_0^*$ *is an interior point of* Φ, *the bootstrap test (15) based on* T_j *has asymptotic power 1* $(j = 1, 2)$.

In view of Theorem 3, for the bootstrap test (15) to have asymptotic power against a given alternative, it is necessary to have a $y \in \mathbb{R}$ such that $\mathrm{E}[\{h_1/h_1(\boldsymbol{\phi}_0^*) - 1\}\mathbb{I}(Y_0 \leq y)] \neq 0$, where $h_t = \mathrm{E}(Y_t^2 \mid \mathcal{H}_{t-1})$, $t \in \mathbb{Z}$. Since h_t is not of the form $h_t(\boldsymbol{\phi})$ under H_1 and $(\boldsymbol{\phi}_0^*, F_0^*)$ is the pseudo-true value, the requirement $\mathrm{E}[\{h_1/h_1(\boldsymbol{\phi}_0^*) - 1\}\mathbb{I}(Y_0 \leq y)] \neq 0$ is not very restrictive under H_1. However, in finite samples, the power of the tests can be sensitive to the form of the discrepancy between h_t and $h_t(\hat{\boldsymbol{\phi}}_n)$. More precisely, if $h_t(\hat{\boldsymbol{\phi}}_n)$ is significantly different from h_t such that the magnitude of the process $n^{-1/2}\sum_{t=1}^n \{Y_t^2/h_t(\hat{\boldsymbol{\phi}}) - 1\}\mathbb{I}(Y_{t-1} \leq y)$ is 'large' for some y, then the KS and CvM functionals of $n^{-1/2}\sum_{t=1}^n \{Y_t^2/h_t(\hat{\boldsymbol{\phi}}) - 1\}\mathbb{I}(Y_{t-1} \leq y)$ are likely to be significantly large compared to realizations from the empirical distributions of KS* and CvM*, respectively, leading to finite sample power of the bootstrap tests. For example, if the true conditional variance h_t is non-linear while the null parametric form $h_t(\boldsymbol{\phi})$ is linear, then our tests are likely to have better finite sample power compared to the case where h_t and $h_t(\boldsymbol{\phi})$ are both linear and the misspecification is only in terms of some missing lags, because in the latter case the KS and CvM functionals of $n^{-1/2}\sum_{t=1}^n \{Y_t^2/h_t(\hat{\boldsymbol{\phi}}) - 1\}\mathbb{I}(Y_{t-1} \leq y)$ are likely to be smaller compared to the former.

4. INFERENCE WHEN THE TRUE VALUE MAY LIE ON THE BOUNDARY

Heuristic arguments suggest that T_1 and T_2 in (7) could serve as possible test statistics for testing H_0 against H_1 regardless of whether $\boldsymbol{\phi}_0$ lies in the interior or on the boundary of

the parameter space. In fact, from Lemma 1, under assumptions (A1) and (A4), we have

$$\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n) = \mathcal{U}_n(y, \boldsymbol{\phi}_0) - n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)' J(y, \boldsymbol{\phi}_0) + o_p(1), \tag{20}$$

uniformly in $y \in \mathbb{R}$, irrespective of whether $\boldsymbol{\phi}_0$ is in the interior or on the boundary of Φ, with $\mathcal{U}_n(\cdot, \boldsymbol{\phi}_0)$ converging weakly to a time transformed Brownian motion. Therefore, the weak limit of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$, and hence the limiting distributions of T_1 and T_2, depend on the asymptotic behaviour of $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)' J(\cdot, \boldsymbol{\phi}_0)$. Hence, to establish the limiting distributions of the test statistics it is essential to study the large sample properties of $\hat{\boldsymbol{\phi}}_n$ when $\boldsymbol{\phi}_0$ lies on the boundary of the parameter space. Several important results on this have already been obtained by Andrews (2001) and Francq and Zakoian (2007). For the ease of reference, in the next subsection, we summarize some of these results in the notation used in this paper.

4.1. Limiting distributions of the estimators

In this section, we summarize several technical results regarding the asymptotic behaviour of the QMLE $\hat{\boldsymbol{\phi}}_n$ in (5) when some components of $\boldsymbol{\phi}_0$ are allowed to be zero, and hence $\boldsymbol{\phi}_0$ could be on the boundary of Φ.

In Theorem 1, when $\boldsymbol{\phi}_0$ is an interior point of Φ, no moment condition of Y_t^2 is needed to establish the weak convergence of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$ and for the existence of $\Sigma(\boldsymbol{\phi}_0)$. This is possible because when $\boldsymbol{\phi}_0$ is an interior point of Φ, since $\tau_t(\boldsymbol{\phi})$ is obtained by dividing the score vector by h_t, the strict stationarity condition $\gamma(\boldsymbol{A}_0) < \infty$ is sufficient to obtain $\mathrm{E}\|\tau_t(\boldsymbol{\phi}_0)\| < \infty$, $\mathrm{E}\|\tau_t(\boldsymbol{\phi}_0)\tau_t(\boldsymbol{\phi}_0)'\| < \infty$ and $\mathrm{E}\|(\partial/\partial\boldsymbol{\phi})\tau_t(\boldsymbol{\phi}_0)'\| < \infty$; see, e.g. Francq and Zakoïan (2004). However, if some components of $\boldsymbol{\phi}_0$ are zero, then one requires additional assumptions to ensure the existence of these moments. To see this, consider the score for the ARCH(2), with $\boldsymbol{\phi} = (\omega, \alpha_1, \alpha_2)'$ and $\alpha_{20} = 0$ and $\ell_t(\boldsymbol{\phi})$ defined in (5); the score in the direction of α_2 is given by,

$$\sum_{t=1}^{n} \frac{\partial \ell_t(\boldsymbol{\phi})}{\partial \alpha_2}\bigg|_{\boldsymbol{\phi}=\boldsymbol{\phi}_0} = \sum_{t=1}^{n} (1 - \varepsilon_t^2) s_t, \quad s_t = \frac{Y_{t-2}^2}{\omega_0 + \alpha_{10} Y_{t-1}^2}.$$

Hence, for s_t to have finite variance (and the CLT to apply), we need $\mathrm{E}(Y_t^4) < \infty$.

In order to ensure that $\Sigma(\boldsymbol{\phi}_0)$ exists when some components of $\boldsymbol{\phi}_0$ are zero, we assume the following moment condition.

(A6). $\mathrm{E}Y_t^6 < \infty$.

Note that, in view of the above example, the moment condition $\mathrm{E}Y_t^4 < \infty$ appears to be necessary (for the existence of $\Sigma(\boldsymbol{\phi}_0)$) and is also probably sufficient. However, in the proof given in Francq and Zakoian (2007) (which we rely on), the stronger moment condition (A6) is required to establish the asymptotic distribution of $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)$ when some components of $\boldsymbol{\phi}_0$ are zero. Therefore, here we use the stronger moment assumption (A6).

Since the parameter $\boldsymbol{\phi}_0$ is allowed to contain zero components, by the assumption that Φ contains a hypercube (see (A1)), the space $n^{1/2}(\Phi - \boldsymbol{\phi}_0)$ increases to the convex cone

$$\Lambda = \Lambda(\boldsymbol{\phi}_0) = \Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_{p_1+p_2+1},$$

where $\Lambda_1 = \mathbb{R}$, and for each $i = 2, \ldots, p_1 + p_2 + 1$, denoting $\boldsymbol{\phi}_0 = (\phi_{01}, \ldots, \phi_{0(1+p_1+p_2)})'$, $\Lambda_i = \mathbb{R}$ if $\phi_{0i} \neq 0$ and $\Lambda_i = [0, \infty)$ if $\phi_{0i} = 0$. Next lemma shows that, under (A1)–(A6), the asymptotic distribution of $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)$ can be represented as the projection of a normal vector distribution onto Λ; for further details on the nature of this projection, see Section 4 in Francq and Zakoian (2007).

Lemma 2. *Suppose that Assumptions (A1)-(A3) are satisfied. Then, $\hat{\boldsymbol{\phi}}_n \overset{a.s.}{\to} \boldsymbol{\phi}_0$, as $n \to \infty$. Additionally, assume that Assumptions (A4) and (A6) are also satisfied. Then,*

$$n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) \overset{d}{\to} \lambda^{\Lambda} := \arg\inf_{\lambda \in \Lambda} (\lambda - Z)'\Sigma(\boldsymbol{\phi}_0)(\lambda - Z),$$

where $Z \sim \mathcal{N}(0, (\kappa_\varepsilon - 1)\Sigma^{-1}(\boldsymbol{\phi}_0))$, $\Sigma(\boldsymbol{\phi}) := \mathrm{E}\{\tau_1(\boldsymbol{\phi})\tau_1(\boldsymbol{\phi})'\}$, $\boldsymbol{\phi} \in \Phi$.

The proof of Lemma 2 follows from Francq and Zakoian (2007). If $\boldsymbol{\phi}_0$ is an interior point, then $\Lambda = \mathbb{R}^{p_1+p_2+1}$ and $\lambda^{\Lambda} = Z \sim \mathcal{N}(0, (\kappa_\varepsilon - 1)\Sigma^{-1}(\boldsymbol{\phi}_0))$, which is the same as the classical case (e.g., see Berkes and Horváth, 2004) as we also considered in the previous section.

4.2. Consistent bootstrap with parameters on the boundary

The bootstrap true parameter value, say $\boldsymbol{\phi}_n^*$, plays a crucial role in defining the properties of any bootstrap test. For the standard bootstrap test in Section 3.2 we set $\boldsymbol{\phi}_n^*$ equal to $\hat{\boldsymbol{\phi}}_n$. In the proof of Theorem 2, under Assumptions (A1)–(A4), we obtain that the limiting behaviour of $n^{1/2}(\hat{\boldsymbol{\phi}}_n^* - \boldsymbol{\phi}_n^*)$, conditional on $(Y_1, \ldots, Y_n)$, is the same as that of $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)$, under H_0, since $\boldsymbol{\phi}_0$ is an interior point and $\boldsymbol{\phi}_n^* = \hat{\boldsymbol{\phi}}_n$. This result plays a key role in the proof of establishing the validity of the bootstrap tests for the case the true parameter lies in the interior of the parameter space. Convergence results of this type have also been used in establishing the asymptotic validity of other bootstrap methods in similar contexts (see Hidalgo and Zaffaroni, 2007; Perera *et al.*, 2016; Perera and Silvapulle, 2022).

However, in the current setup, the parameter $\boldsymbol{\phi}_0$ is allowed to contain zero components, and hence we require additional conditions to ensure that the bootstrap tests are consistent. In particular, a crucial requirement for the validity of the bootstrap tests is to have the following rate of consistency for the bootstrap true value $\boldsymbol{\phi}_n^* = (\phi_{n1}^*, \ldots, \phi_{n(1+p_1+p_2)}^*)'$:

$$n^{1/2}(\phi_{ni}^* - \phi_{0i}) = \begin{cases} o_p(1), & \text{if } \phi_{0i} = 0 \\ O_p(1), & \text{if } \phi_{0i} > 0 \end{cases}, \quad i = 1, 2, \ldots, 1 + p_1 + p_2. \tag{21}$$

This requirement has previously been introduced in Cavaliere *et al.* (2022) for establishing the validity of a bootstrap based inference procedure in a different context. In the current setup, (21) ensures that the bootstrap method based on $\boldsymbol{\phi}_n^*$ replicates the unknown limiting distribution of T_j under the null, while being of order $O_p^*(1)$, in probability, under the alternative ($j = 1, 2$), as is established in Theorems 4 and 5 below. If we set $\boldsymbol{\phi}_n^* = \hat{\boldsymbol{\phi}}_n$, then it only holds that $n^{1/2}(\phi_{ni}^* - \phi_{0i}) = O_p(1)$ for $i = 1, 2, \ldots, 1 + p_1 + p_2$, and hence (21) is not satisfied. Therefore, the standard bootstrap test outlined in Section 3.2 is not valid when some parameters lie on the boundary of Φ under H_0. Hence, instead of the standard bootstrap, we propose a new bootstrap method based on using a different mechanism in choosing the bootstrap true values ϕ_{ni}^*, $i = 1, 2, \ldots, 1+p_1+p_2$. The modified bootstrap testing procedure is based on shrinking the parameter estimators in the bootstrap data generation. Specifically, instead of using $\boldsymbol{\phi}^* = \hat{\boldsymbol{\phi}}_n = (\hat{\phi}_{n1}, \ldots, \hat{\phi}_{n(1+p_1+p_2)})'$ as the true value in the bootstrap data generation, we use a transformed version of $\hat{\boldsymbol{\phi}}_n$, denoted $\hat{\boldsymbol{\phi}}_n^\dagger = (\hat{\phi}_{n1}^\dagger, \ldots, \hat{\phi}_{n(1+p_1+p_2)}^\dagger)'$ defined by $\hat{\phi}_{n1}^\dagger = \hat{\phi}_{n1}$ and

$$\hat{\phi}_{ni}^\dagger := \hat{\phi}_{ni}\mathbb{I}(\hat{\phi}_{ni} > c_n) \quad i = 2, \ldots, 1 + p_1 + p_2, \tag{22}$$

where c_n is a scalar sequence converging to zero at a rate satisfying:

$$c_n \to 0, \quad \text{and} \quad n^{1/2}c_n \to \infty \quad \text{as } n \to \infty. \tag{23}$$

This approach has its roots in the Hodges-Le Cam super-efficient type estimators, see e.g. Bickel *et al.* (1998), Chatterjee and Lahiri (2011) and Cavaliere *et al.* (2022).

In view of the parameter restrictions in (3), denoting $\boldsymbol{\phi}_0 = (\phi_{01}, \ldots, \phi_{0(1+p_1+p_2)})'$, we have that $\phi_{01} = \omega_0 > 0$, $\phi_{0i} = \alpha_{0(i-1)} \geq 0$ ($i = 2, \ldots, 1 + p_1$), and $\phi_{0i} = \beta_{0(i-1-p_1)} \geq 0$ ($i = 2 + p_1, \ldots, 1 + p_1 + p_2$). Thus, ϕ_{01} is always in the interior, and ϕ_{0j} is on the boundary of the parameter space only if $\phi_{0j} = 0$ for some $j \in \{2, 3, \ldots, 1 + p_1 + p_2\}$; i.e. some ARCH or GARCH coefficient is zero. Since $\hat{\boldsymbol{\phi}}_n$ is root-n consistent, the proposed shrinkage in

terms of the c_n sequence ensures that $P(\hat{\phi}^{\dagger}_{nj} = 0) \to 1$ as $n \to \infty$ whenever $\text{plim}\, \hat{\phi}_{nj} = 0$, $j \in \{2, 3, \ldots, 1 + p_1 + p_2\}$, where 'plim' is the probability limit as $n \to \infty$. Hence, unlike $\hat{\phi}_{nj}$, in large samples, the transformed estimator $\hat{\phi}^{\dagger}_{nj}$ lies on the boundary of the parameter space with large probability whenever ϕ_{0j} is on the boundary; i.e. $\phi_{0j} = 0$. Since $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) = O_p(1)$ and c_n converges at a rate slower than $n^{-1/2}$, this ensures that the requirement (21) is satisfied by the parameter $\hat{\boldsymbol{\phi}}^{\dagger}_n$ defined by (22)–(23). Hence, as established in Theorems 4 and 5 below, the bootstrap based on $\boldsymbol{\phi}^*_n = \hat{\boldsymbol{\phi}}^{\dagger}_n$ allows us to replicate the unknown limiting distributions of T_1 and T_2 under H_0, while being of order $O^*_p(1)$, in probability, under the alternative.

We next provide a step-by-step guide of the proposed modified bootstrap approach.

Bootstrap Algorithm 2

Step 1: Compute $\{\hat{\boldsymbol{\phi}}_n, T_j\}$ on the original sample $\{Y_1, \ldots, Y_n\}$; moreover, compute $\hat{\boldsymbol{\phi}}^{\dagger}_n$ using $\{\hat{\boldsymbol{\phi}}_n, T_j\}$ as in (22)–(23);

Step 2: Compute $\check{\varepsilon}_t$, $t = 1, \ldots, n$ as in (11) and draw a random sample (with replacement) of size n, say $\{\varepsilon^*_1, \ldots, \varepsilon^*_n\}$, independent of the original data, from $\check{F}_n(\cdot) = n^{-1} \sum_{t=1}^n \mathbb{I}(\check{\varepsilon}_t \leq \cdot)$;

Step 3: Generate the bootstrap sample $\{Y^*_1, \ldots, Y^*_n\}$ with bootstrap true values $(\hat{\boldsymbol{\phi}}^{\dagger}_n, \check{F}_n)$ as

$$Y^*_t = \{h^*_t(\hat{\boldsymbol{\phi}}^{\dagger}_n)\}^{1/2} \varepsilon^*_t, \quad h^*_t(\hat{\boldsymbol{\phi}}^{\dagger}_n) = \hat{\omega}^{\dagger} + \sum_{j=1}^{p_1} \hat{\alpha}^{\dagger}_j (Y^*_{t-j})^2 + \sum_{k=1}^{p_2} \hat{\beta}^{\dagger}_k h^*_{t-k}(\hat{\boldsymbol{\phi}}^{\dagger}_n), \quad t \geq 1$$

initialized with $(Y^*_0, \ldots, Y^*_{1-p_1}, h^*_0(\hat{\boldsymbol{\phi}}^{\dagger}_n), \ldots, h^*_{1-p_2}(\hat{\boldsymbol{\phi}}^{\dagger}_n))' = \varsigma_0$, where ς_0 is an arbitrarily chosen vector (e.g., set $Y^*_t = 0$ and $h^*_t = 0$, all $t \leq 0$);

Step 4: Using $\{Y^*_1, \ldots, Y^*_n\}$, compute $\{\hat{\boldsymbol{\phi}}^*_n, T^*_j\}$ the bootstrap analogs of $\{\hat{\boldsymbol{\phi}}_n, T_j\}$.

Step 5: Compute the bootstrap test statistic T^*_j as $T^*_1 = \text{KS}^* = \sup_y \left|\mathcal{U}^*_n(y, \hat{\boldsymbol{\phi}}^*_n)\right|$ and $T^*_2 = \text{CvM}^* = \int \left\{\mathcal{U}^*_n(y, \hat{\boldsymbol{\phi}}^*_n)\right\}^2 dG^*_n(y)$, where $G^*_n(y)$ and $\mathcal{U}^*_n(y, \boldsymbol{\phi})$ are the bootstrap analogs of $G_n(y)$ and $\mathcal{U}_n(y, \boldsymbol{\phi})$, respectively.

Step 6: The bootstrap p-value is defined as $p^*_n := p^*_n(\hat{\boldsymbol{\phi}}^{\dagger}_n) := P^*_n(T^*_j \geq T_j)$ where P^*_n is the probability measure induced by the bootstrap (i.e., conditional on the original data). □

As for the previous bootstrap algorithm, the bootstrap test rejects H_0 at the nominal level α if the bootstrap p-value p^*_n is less than α.

Note that, the limiting distribution of $n^{1/2}(\hat{\boldsymbol{\phi}}^{\dagger}_n - \boldsymbol{\phi}_0)$ is the same as that of $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)$ whenever $\boldsymbol{\phi}_0$ is in the interior of Φ. Hence, the bootstrap test collapses into the bootstrap method outlined in Section 3.2 as $n \to \infty$, whenever $\boldsymbol{\phi}_0$ is in the interior of Φ.

4.3. Asymptotic validity

In this section we establish the asymptotic validity of the bootstrap based on the shrinking parameter estimators approach introduced in the previous subsection.

The bootstrap analogue of the marked empirical process $\mathcal{U}_n(y, \boldsymbol{\phi})$ for the bootstrap test based on Algorithm 2 is defined as in (16), with

$$\mathcal{U}_n^*(y, \boldsymbol{\phi}) := n^{-1/2} \sum_{t=1}^{n} \left\{ \frac{(Y_t^*)^2}{h_t^*(\boldsymbol{\phi})} - 1 \right\} \mathbb{I}(Y_{t-1}^* \leq y), \quad y \in \mathbb{R},\ \boldsymbol{\phi} \in \Phi,$$

except that Y_t^* and $h_t^*(\boldsymbol{\phi})$ are now based on the bootstrap method outlined in Section 4.2.

Note that, Assumptions (A1)–(A6) correspond to the original data generating process, and hence the underlying true parameter value $\zeta_0 = (\boldsymbol{\phi}_0, F_0)$ is fixed. However, in the bootstrap data generation the true parameter $(\hat{\boldsymbol{\phi}}_n^{\dagger}, \check{F}_n)$ is not fixed but converges to $(\boldsymbol{\phi}_0^*, F_0^*)$ as $n \to \infty$. Therefore, it is not adequate to assume only (A1)–(A6) in order to establish the validity of the bootstrap tests. As mentioned earlier, Assumption (A5) allows us to extend the arguments of Lemma 2 to a triangular array setup to obtain that the limiting distribution of $n^{1/2}(\hat{\boldsymbol{\phi}}_n^* - \hat{\boldsymbol{\phi}}_n^{\dagger})$, conditional on $(Y_1, \ldots, Y_n)$, is the same as that of $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)$ under H_0, while being of order $O_p^*(1)$, in probability, under H_1. This result is required in the proofs of the asymptotic validity and consistency of the bootstrap tests.

The next theorem establishes the asymptotic validity of the bootstrap Algorithm 2.

Theorem 4. *Suppose that H_0 holds and Assumptions (A1)–(A5) are satisfied. Additionally, w.l.o.g., assume that the set $\bar{\Phi}^*$ in (18) contains $\boldsymbol{\phi}_0$ with $\omega_L < \omega_0 < \omega_U$ and $\mathbf{0} \leq \boldsymbol{\phi}_0 < \bar{\boldsymbol{\phi}}_U$. Then, the conditional weak limit of $\mathcal{U}_n^*(\cdot, \hat{\boldsymbol{\phi}}_n^*)$ is the same as that of $\mathcal{U}_n(\cdot, \hat{\boldsymbol{\phi}}_n)$, in probability, and hence the bootstrap test based on T_j is asymptotically valid ($j = 1, 2$).*

The next theorem establishes the consistency of the bootstrap test under the alternative.

Theorem 5. *Suppose that H_1 holds and Assumptions (B1), (B2) and (A5) are satisfied. Additionally, assume that $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0^*) = O_p(1)$ and there exists a $y \in \mathbb{R}$, with $h_t = \mathrm{E}(Y_t^2 \mid \mathcal{H}_{t-1})$, $t \in \mathbb{Z}$, such that $\mathrm{E}[\{h_1/h_1(\boldsymbol{\phi}_0^*) - 1\}\mathbb{I}(Y_0 \leq y)] \neq 0$. Then, conditional on $\{Y_1, \ldots, Y_n\}$, the bootstrap test based on T_j has asymptotic power 1 ($j = 1, 2$).*

Theorem 4 shows that the proposed shrinkage in terms of the c_n sequence, or more generally, the requirement (21) ensures that the bootstrap test statistics T_1^* and T_2^* based on Algorithm 2 replicate the unknown limiting distributions of T_1 and T_2 under the null. Theorem 5 establishes that T_1^* and T_2^* are of order $O_p^*(1)$, in probability, under the alternative;

that is, the proposed bootstrap method is also consistent even if it is unknown whether any of the nuisance parameters are on the boundary. Finally, since h_t is not of the form $h_t(\boldsymbol{\phi})$ under H_1, the requirement on $\mathrm{E}[\{h_1/h_1(\boldsymbol{\phi}_0^*) - 1\}\mathbb{I}(Y_0 \leq y)]$ is not very restrictive.

4.4. The case of parameters near the boundary

Throughout the section we have assumed that the parameter vector $\boldsymbol{\phi}$ is a fixed element of the parameter space. This assumption rules out the case of drifting sequences of parameters converging to the boundary at the usual $n^{-1/2}$ rate, which is the standard way of investigating the effect of parameter values which are close to, but not on the boundary; see e.g. Ketz (2018) and the references therein.

To shed some light on the implication of this setup, consider again the case where the true parameter $\boldsymbol{\phi}_0$ is fixed. Then, the classical bootstrap will fail if some of the elements in $\boldsymbol{\phi}_0$ are on the boundary, since the limiting bootstrap measure will be random in the limit; see the discussion in Section 4.3. A similar result has been obtained e.g. by Chatterjee and Lahiri (2011) for the standard bootstrap of lasso-type estimators, which fails when some elements of the parameter vector are zero. Nevertheless, our proposed bootstrap – similarly to the modified bootstrap in Chatterjee and Lahiri (2011, eq. (2.4)) – will be valid, for in large samples the shrinkage-based estimator $\hat{\boldsymbol{\phi}}_n^{\dagger}$, employed as bootstrap true value, converges to the (zero) true value at a rate faster than $n^{-1/2}$. As a consequence, the conditional distribution of our modified bootstrap statistic is not random in the limit and, instead, it converges to the asymptotic distribution of the original statistic.

However, should the true parameter be in a $n^{-1/2}$-neighborhood from the boundary, then our method would incorrectly place the parameter on the boundary, and hence would not be asymptotically valid. This is because $\hat{\boldsymbol{\phi}}_n^{\dagger}$ is still of order $o_p(n^{-1/2})$ when $\boldsymbol{\phi}_0 \sim n^{-1/2}$; see below. To conclude, the proposed test is valid for sequences, $\boldsymbol{\phi}_{0,n}$ say, close to boundary provided $\boldsymbol{\phi}_{0,n} \sim n^{-\delta}$ with $\delta > 1/2$, while for $\delta = 1/2$ it is invalid.

In Section 5.2 below we analyze by Monte Carlo simulation the effects of this incorrect classification on the performance of our tests, and we do not see particular evidence of a critical behaviour of our proposed approach near the boundary.

Should the lack of validity near the boundary be a concern, a modified bootstrap procedure, which would be asymptotically valid even if the true parameter is in a $n^{-1/2}$ neighborhood of the boundary, can be derived following the approach in Doko Tchatoka and Wang (2021). To describe this procedure, suppose for the sake of simplicity that $\boldsymbol{\phi} = \phi$

is one-dimensional and takes value in the parameter space $[0, \infty)$ (a simple example is the ARCH(1) model with known intercept $Y_t = (1 + \phi Y_{t-1}^2)^{1/2}\varepsilon_t$, where $\phi \geq 0$), such that if $\phi = 0$ the parameter is on the boundary. Let ϕ_n^* denote the bootstrap true value and let $p_n^*(\phi_n^*)$ denote the bootstrap p-value based on ϕ_n^*. The standard bootstrap, see Algorithm 1, is obtained by setting $\phi_n^* = \hat{\phi}_n$ (with associated bootstrap p-value given by $p_n^*(\hat{\phi}_n)$), while the shrinkage-based bootstrap, see Algorithm 2, requires $\phi_n^* = \hat{\phi}_n^\dagger := \hat{\phi}_n \mathbb{I}(\hat{\phi}_n > c_n)$ (with associated bootstrap p-value given by $p_n^*(\hat{\phi}_n^\dagger)$). When the true parameter is $\phi_0 \sim n^{-1/2}$, both bootstrap algorithms are invalid, as emphasized. Specifically, on the one hand the shrinkage-based bootstrap fails as $n^{1/2}\hat{\phi}_n^\dagger = o_p(1)$ and, asymptotically, the bootstrap p-value is equivalent to a bootstrap which sets $\phi_n^* = 0$; i.e., $p_n^*(\hat{\phi}_n^\dagger) = p_n^*(0) + o_p(1)$. The invalidity arises because $p_n^*(\hat{\phi}_n^\dagger)$ does not match the 'infeasible' bootstrap p-value based on the true parameter, $p_n^*(\phi_n)$. On the other hand, the classical bootstrap fails because $\phi_n^* = \hat{\phi}_n$ is consistent at the $n^{1/2}$ rate, which is too slow for the bootstrap to work (see e.g. the discussion in Andrews, 2000).

A conservative bootstrap, which allows to generate a bootstrap p-value p_n^* which controls the type-I error rejection probability asymptotically, can be designed as follows (see also the 'hybrid bootstrap' algorithm in Doko Tchatoka and Wang, 2021).

Bootstrap Algorithm 3

Step 1: Compute $\hat{\phi}_n$. If $\mathbb{I}(\hat{\phi}_n > c_n) = 1$, set $\phi_n^* = \hat{\phi}_n$ as in Algorithm 1 and set $p_n^* = p_n^*(\hat{\phi}_n)$; otherwise, proceed to Step 2;

Step 2: Create a fine grid on the interval $[0, c_n]$, and call it $\mathcal{C}_n$;

Step 3: For each $c \in \mathcal{C}_n$, compute the bootstrap p-value $p_n^*(c)$;

Step 4: The (conservative) bootstrap p-value is computed as $p_n^* := \sup_{c \in \mathcal{C}_n} p_n^*(c)$.

As for the other bootstrap algorithms, the null hypothesis is rejected if p_n^* is below the chosen nominal significance level. This bootstrap will deliver a valid conservative test for the null hypothesis H_0 of correct model specification; i.e., under H_0, $P(p_n^* \leq \eta) \leq \eta + o(1)$ as $n \to \infty$, with equality holding when $n^{1/2}\phi_n - c_n \to \infty$ (which includes the special case of fixed parameter, $\phi_n = \phi_0$). The proof follows by simply noticing that, if $\phi = \phi_n \sim n^{-1/2}$, then for n large enough $\phi_n \in [0, c_n]$; hence, p_n^* from Step 4, for $\mathcal{C}_n$ fine enough, satisfies $p_n^* \geq p_n(\phi_n)$, the right hand side being the p-value corresponding to the infeasible bootstrap based on the true parameters.

Remark 2. *Note that Algorithm 3 is specified for a single parameter. In general, the*

shrinking is applied to $\hat{\phi}_{ni}$*, for* $i = 2, 3, \ldots, p_1 + p_2 + 1$*, and hence the grid in step 2 accordingly should be defined as a cube* $[0, c_n]^d$*, with* d *the number of parameters for which* $\hat{\phi}_{ni} \leq c_n$*, that is,* $d = \sum_{i=2}^{p_1+p_2+1} \mathbb{I}(\hat{\phi}_{ni} \leq c_n)$*.*

Remark 3. *The bootstrap-based test described in Algorithm 3 is conservative; that is, it controls the type-I error probability asymptotically. Hence, the actual rejection probability could be below the nominal significance level. Since ours is a misspecification test, the practitioner might prefer to use a liberal test; i.e. such that the probability of rejecting the null* H_0 *of correct model specification when* H_0 *is false is* at least *equal to the nominal level. Should this be the case, then it suffices to define the (liberal) bootstrap p-value in Step 4 as* $p_n^* := \inf_{c \in \mathcal{C}_n} p_n^*(c)$*.*

Remark 4. *The test described in this section is computationally intensive in large dimensional models. However, the dimension of the grid in Step 3 depends on the parameters which are near or on the boundary, and not on the dimension of the whole parameter set. Hence, it is feasible in most applications.*

5. Numerical Study

In this section we carry out a Monte Carlo simulation study to evaluate the finite sample performance of the KS and CvM tests based on the shrinking-based bootstrap method proposed in Section 4.2. Our main focus is the case where the true parameter value $\boldsymbol{\phi}_0$ of the data generating process lies on the boundary of the parameter space. For comparison, we also consider the case where $\boldsymbol{\phi}_0$ is an interior point. Several data generating processes under the alternative hypothesis are also considered in order to investigate the finite sample power properties of the tests. Although there are several other tests that can be applied for testing the conditional variance specification in GARCH-type models, as mentioned in the introduction, the theory for their validity does not hold when the true parameter is on the boundary. Hence, in these simulations, we compare the proposed tests with the general purpose Ljung-Box Q test which tests the significance of the serial dependence of the squared residuals estimated from the fitted model. We denote the Ljung-Box Q test for a lag length ℓ by LBQ(ℓ). For comparison, we also consider the standard bootstrap (as proposed in Section 3.2), the hybrid bootstrap (as proposed by Algorithm 3 in Section 4.4), and 'm out of n' bootstrap (see Hall and Yao, 2003). The asymptotic validity of the 'm out of n' bootstrap implementation of KS and CvM tests does not directly follow from Hall and

Yao (2003). Nevertheless, for comparison, we use it as an alternative approach to estimate the finite sample null distributions of the test statistics.

5.1. The null hypothesis and the Monte Carlo design

We assume that the observable process $\{Y_t\}$ obeys model (1). The tests are evaluated when the parametric form $h_t(\boldsymbol{\phi})$ under H_0 is

$$h_t(\boldsymbol{\phi}) = \omega + \alpha_1 Y_{t-1}^2 + \alpha_2 Y_{t-2}^2 + \beta_1 h_{t-1}(\boldsymbol{\phi}) + \beta_2 h_{t-2}(\boldsymbol{\phi}), \quad \omega > 0, \alpha_1, \alpha_2, \beta_1, \beta_2 \geq 0, \tag{24}$$

where $\boldsymbol{\phi} = (\omega, \alpha_1, \alpha_2, \beta_1, \beta_2)'$.

The results are based on $20,000$ Monte Carlo replications. For each replication and data generating process, we first compute the QMLE $\hat{\boldsymbol{\phi}}_n$ and compute the test statistics KS, CvM, and LBQ(ℓ), $\ell = 3, 5, 10, 15, 20$. In the bootstrap implementations, to reduce the computational burden, we adopt the 'Warp-Speed' Monte Carlo method of Giacomini *et al.* (2013). The results of LBQ(ℓ) are presented for only $\ell = 15$; the patterns of the results for $\ell = 3, 5, 10$ and 20 are similar to those for $\ell = 15$ and hence are omitted.

5.2. Empirical rejection probabilities under the null hypothesis

The data generating process [DGP] is

$$Y_t = h_t^{1/2} \varepsilon_t, \quad h_t = \omega_0 + \alpha_{01} Y_{t-1}^2 + \alpha_{02} Y_{t-2}^2 + \beta_{01} h_{t-1} + \beta_{02} h_{t-2}.$$

For the error distribution, we consider the standard normal distribution. For the conditional variance h_t of the true DGP we consider the following cases:

$$(\omega_0, \alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02}) = \begin{cases} (0.3, & 0, & 0.4, & 0, & 0) & \text{[D0]} \\ (0.3, & 0.5, & 0.45, & 0, & 0) & \text{[D1]} \\ (0.3, & 0.2, & 0.3, & 0.45, & 0) & \text{[D2]} \\ (0.3, & 0, & 0.4, & 0, & 0.55) & \text{[D3]} \\ (0.3, & 0, & 0.4, & 0.55, & 0) & \text{[D4]} \\ (0.3, & 0.2, & 0.25, & 0.2, & 0.3) & \text{[D5]} \end{cases} \tag{25}$$

Thus, for the DGPs D0–D4, the true parameter $\boldsymbol{\phi}_0 = (\omega_0, \alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02})'$ has at least one component on the boundary of the parameter space, and for the DGP D5, the true parameter is an interior point of the parameter space. Note for D1-D5, the sixth order moment assumption is deliberately not satisfied in order to assess robustness of our tests.

As for the choice of the shrinkage sequence c_n, as in Cavaliere *et al.* (2022), we set $c_n = vn^{-\epsilon}$, with $\epsilon = 0.45$ and $v = 1.60$. For the 'm out of n' bootstrap implementation, we set the

size m_n of the bootstrap sample to be the integer value of $cn/\log(n)$, with $c = 1.5$. Different choices of c_n and m_n are discussed in Section S.4.3 in the supplementary material. For the hybrid bootstrap algorithm 3 we select the grid $\mathcal{C}_n = \{kc_n : k = 0, 0.1, 0.2, \ldots, 0.9, 1\}$.

Table 1 presents the results for empirical level. The sample sizes $n = 100, 600$ and 2000 are considered. All the tests are run at the nominal 10% significance level. In these results both shrinking based and standard bootstrap approaches exhibit excellent performance. The 'm out of n' bootstrap based tests are oversized for small samples, that is they have larger rejection probability than the nominal significance level. Although their performance becomes better as the sample size increases they do not perform as well as the tests based on the first two bootstrap methods. As expected, the tests based on hybrid bootstrap turn out to be conservative (that is they have smaller rejection probability than the nominal significance level) as they are designed to control the type-I error probability asymptotically. Hence, the actual rejection probability for the tests based on hybrid bootstrap could be below the nominal significance level. The LBQ test turns out to be significantly undersized.

5.2.1. Uniform size properties

To evaluate the (uniform) finite sample size properties of the tests when some component of the true parameter is near the boundary, we initially consider the DGPs:

$$\begin{aligned} \mathrm{D}_{n1}: \quad & Y_t = h_t^{1/2}\varepsilon_t, \quad h_t = 0.3 + f_n Y_{t-1}^2 + 0.25 Y_{t-2}^2 + 0.35 h_{t-1} + 0.25 h_{t-2}, \\ \mathrm{D}_{n2}: \quad & Y_t = h_t^{1/2}\varepsilon_t, \quad h_t = 0.3 + 0.25 Y_{t-1}^2 + 0.25 Y_{t-2}^2 + f_n h_{t-1} + 0.35 h_{t-2}, \end{aligned}$$

with $f_n = 1.4n^{-1/2}$ and a standard normal error distribution. We investigate the rejection probabilities of the null hypothesis H_0 specified by $h_t(\boldsymbol{\phi})$ in (24), against the above two DGPs, as n varies. Thus, for both D_{n1} and D_{n2}, one of the ARCH/GARCH coefficients converges to the boundary of the null parameter space at the $n^{-1/2}$ rate.

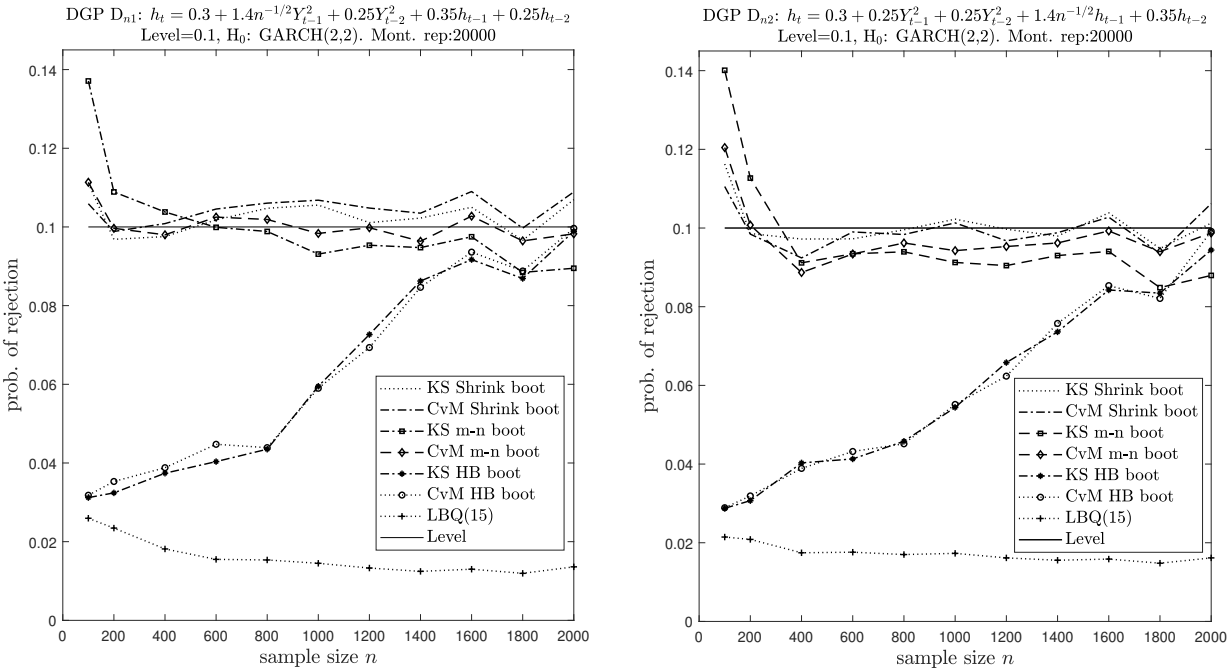
Figure 1 presents results on empirical level at 10% level of significance, for the DGPs D_{n1} and D_{n2}. Since the true parameter is in a $n^{-1/2}$ neighborhood of the boundary, the hybrid bootstrap approach based on Algorithm 3 is asymptotically valid, although it could be conservative in finite samples (see Remark 3 above). The results in Figure 1 are consistent with this because although the tests based on hybrid bootstrap exhibit significant size distortions in small samples they perform increasingly better as the sample size increases and f_n approaches the boundary. The shrinking based bootstrap exhibits the best overall performance followed by standard bootstrap (the results for the standard bootstrap are

Table 1: Empirical rejection probabilities for testing the null hypothesis $\mathsf{H}_0: \ h_t = h_t(\boldsymbol{\phi}) = \omega + \alpha_1 Y_{t-1}^2 + \alpha_2 Y_{t-2}^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2}$, for some $\omega > 0, \alpha_1, \alpha_2, \beta_1, \beta_2 \geq 0$.

	Shrinking based bootstrap		Standard bootstrap		'm out of n' bootstrap		Hybrid bootstrap		
n	KS	CvM	KS	CvM	KS	CvM	KS	CvM	LBQ(15)
	DGP D0: $(\omega_0, \alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02}) = (0.3, 0, 0.4, 0, 0)$								
100	0.088	0.093	0.087	0.089	0.099	0.090	0.025	0.023	0.018
600	0.094	0.095	0.093	0.093	0.094	0.098	0.079	0.079	0.021
2000	0.098	0.101	0.098	0.097	0.092	0.102	0.085	0.086	0.021
	DGP D1: $(\omega_0, \alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02}) = (0.3, 0.5, 0.45, 0, 0)$								
100	0.112	0.107	0.107	0.102	0.117	0.110	0.031	0.029	0.016
600	0.095	0.096	0.093	0.094	0.081	0.092	0.046	0.046	0.023
2000	0.094	0.096	0.094	0.092	0.078	0.086	0.051	0.050	0.023
	DGP D2: $(\omega_0, \alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02}) = (0.3, 0.2, 0.3, 0.45, 0)$								
100	0.108	0.104	0.100	0.095	0.119	0.105	0.030	0.028	0.014
600	0.097	0.096	0.096	0.095	0.086	0.088	0.044	0.042	0.019
2000	0.100	0.096	0.099	0.096	0.086	0.090	0.060	0.054	0.018
	DGP D3: $(\omega_0, \alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02}) = (0.3, 0, 0.4, 0, 0.55)$								
100	0.109	0.107	0.108	0.106	0.147	0.128	0.021	0.020	0.020
600	0.105	0.106	0.104	0.103	0.102	0.103	0.036	0.032	0.032
2000	0.103	0.103	0.105	0.102	0.099	0.103	0.042	0.039	0.036
	DGP D4: $(\omega_0, \alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02}) = (0.3, 0, 0.4, 0.55, 0)$								
100	0.099	0.098	0.091	0.089	0.124	0.102	0.027	0.027	0.022
600	0.101	0.102	0.097	0.095	0.098	0.097	0.040	0.039	0.026
2000	0.105	0.104	0.101	0.101	0.094	0.097	0.042	0.043	0.025
	DGP D5: $(\omega_0, \alpha_{01}, \alpha_{02}, \beta_{01}, \beta_{02}) = (0.3, 0.2, 0.25, 0.2, 0.3)$								
100	0.107	0.106	0.101	0.098	0.125	0.110	0.033	0.032	0.018
600	0.099	0.093	0.098	0.091	0.090	0.091	0.043	0.040	0.019
2000	0.102	0.099	0.101	0.099	0.092	0.097	0.069	0.066	0.017

Notes: The DGPs are of the form $Y_t = h_t^{1/2}\varepsilon_t$, $h_t = \omega_0 + \alpha_{01} Y_{t-1}^2 + \alpha_{02} Y_{t-2}^2 + \beta_{01} h_{t-1} + \beta_{02} h_{t-2}$. The nominal level is 10%. The simulation is based on 20,000 Monte Carlo replications.

Figure 1: Empirical size at the nominal 10% significance level for the DGPs D_{n1} and D_{n2}.



given in Table S1 in the supplementary material). The tests based on 'm out of n' bootstrap turn out to be oversized for small samples, although their performance becomes better as the sample size increases. The LBQ test continues to be significantly undersized for all n.

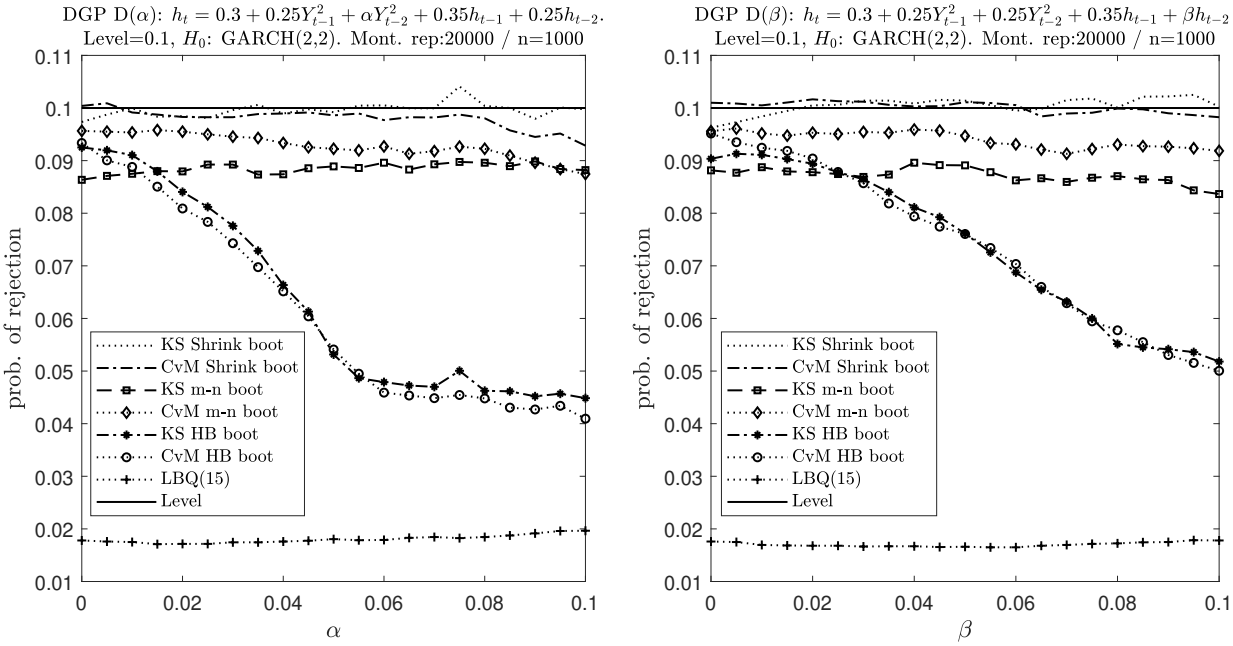
We end this subsection by investigating how well size is controlled both at, near and far away from the boundary. We do so by computing the empirical null rejection probabilities as a function of parameters that start at the boundary and move toward the interior of the parameter space, with the sample size being kept fixed. Specifically, we consider the following two DGPs:

$$\mathrm{D}(\alpha): \quad Y_t = h_t^{1/2}\varepsilon_t, \quad h_t = 0.3 + 0.25Y_{t-1}^2 + \alpha Y_{t-2}^2 + 0.35h_{t-1} + 0.25h_{t-2},$$
$$\mathrm{D}(\beta): \quad Y_t = h_t^{1/2}\varepsilon_t, \quad h_t = 0.3 + 0.25Y_{t-1}^2 + 0.25Y_{t-2}^2 + 0.35h_{t-1} + \beta h_{t-2},$$

with the standard normal error distribution. The tests are evaluated for testing the null hypothesis H_0 specified by $h_t(\boldsymbol{\phi})$ in (24). For the DGPs $\mathrm{D}(\alpha)$ and $\mathrm{D}(\beta)$, the parameters α and β are chosen in the set $\{0, 0.005, 0.01, 0.015, 0.02, 0.025, 0.03, \ldots, 0.085, 0.09, 0.095, 0.1\}$.

Figure 2 provides the empirical null rejection probabilities as a function of the parameters α and β for the DGPs $\mathrm{D}(\alpha)$ and $\mathrm{D}(\beta)$, respectively, at the nominal 10% significance level. The sample size is $n = 1000$. Since the shrinking based bootstrap and standard bootstrap are asymptotically equivalent for $\alpha, \beta > 0$, we do not include the standard boot-

Figure 2: Empirical size at the nominal 10% significance level for the DGPs D(α) and D(β).



strap in Figure 2. Detailed results including standard bootstrap and sample sizes $n = 600$ and 1000 are provided in Figure S1 in the supplementary material. When the tests are implemented using the 'm out of n' bootstrap, the CvM test exhibits better performance than the KS test; however, when they are implemented using either the shrinking based bootstrap or the standard bootstrap the KS test performs marginally better than the CvM test. The hybrid bootstrap approach appears to control size better (although not as well as the shrinking based bootstrap) at or near the boundary when $n = 1000$, however, as we move away from the boundary towards the interior of the parameter space the hybrid bootstrap tests exhibit smaller empirical rejection probabilities than the nominal significance level. The LBQ test appears to be significantly undersized.

5.3. Empirical power and the choice of the shrinkage sequence

In order to investigate empirical power properties of the tests we consider several DGPs under the alternative, including local alternatives (the details are provided in the supplementary material in Section S.4.2). In addition, we also consider DGPs that start at the boundary of the null parameter space and move further into the alternative space. These simulations are discussed in detail in Section S.4.2 in the supplementary material (see Table S2, and Figures S2 and S3). In these simulation experiments the shrinking based, standard and 'm out of n' bootstrap methods all perform very similarly and exhibit

excellent empirical power properties. As expected, the empirical null rejection probabilities of the tests increase as the DGP moves away from the boundary and further into the alternative space. The LBQ test does not exhibit any notable power.

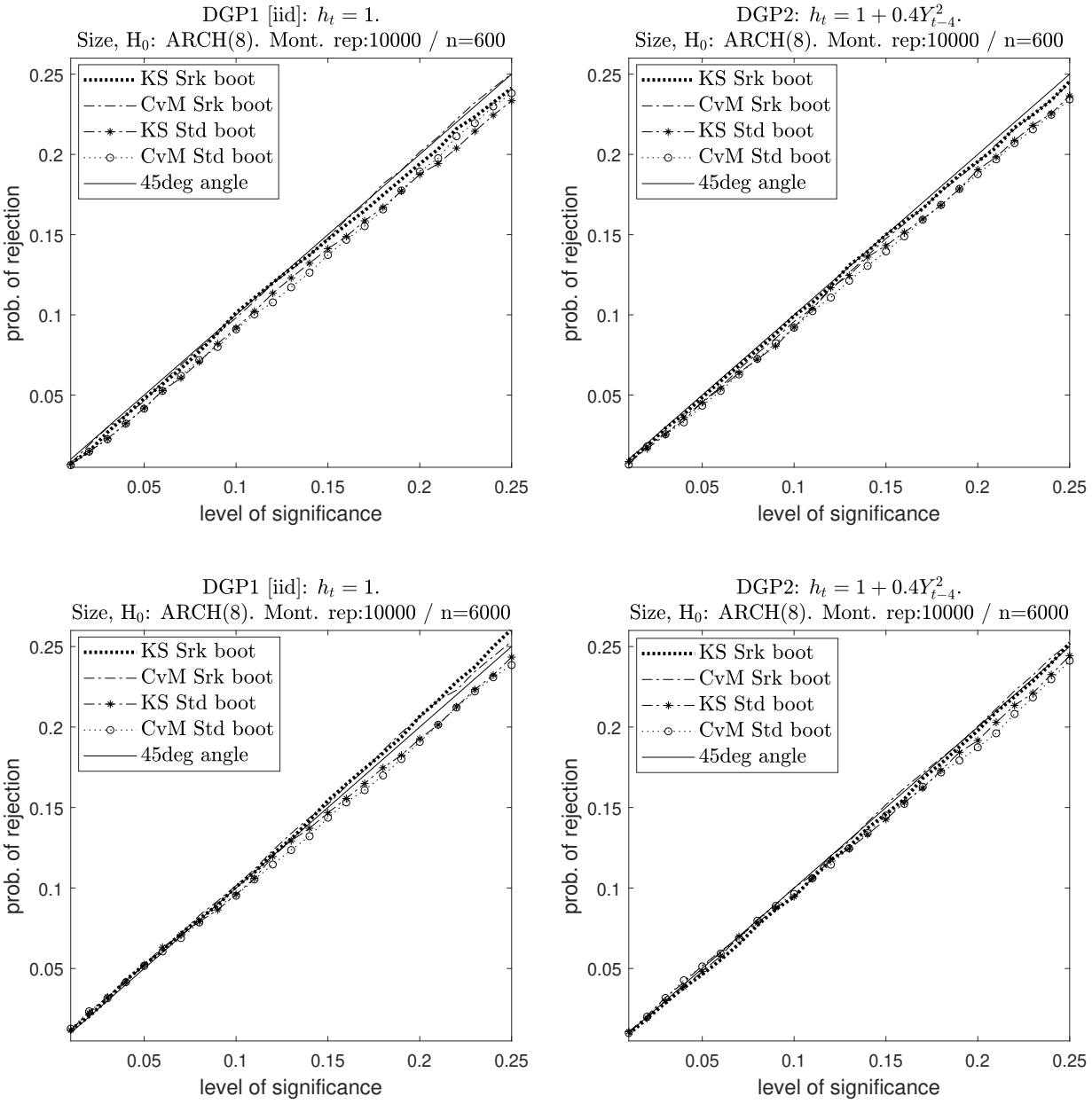
In order to investigate the effect of the choice of the shrinkage parameter c_n we evaluate the performance of the shrinking-based bootstrap tests while considering different choices for c_n. These simulations are discussed in Section S.4.3 in the supplementary material and the results are presented in Tables S3 and S4. In these simulations the shrinking-based bootstrap performs consistently well throughout with good empirical size properties for both KS and CvM tests, and the test results do not vary significantly on the choice of the shrinkage parameter c_n. By comparison, the test results for the 'm out of n' bootstrap exhibit some sensitivity to the choice of the tuning parameter m_n, particularly for small n.

5.4. Standard and shrinking-based bootstrap: further comparisons

As discussed earlier, the standard bootstrap is theoretically invalid when the true parameter lies on the boundary of the parameter space. However, the Monte Carlo simulation results reported in Section 5.2 show that the differences between the standard bootstrap and the (theoretically valid) shrinking-based bootstrap under the null are not always particularly substantial. In this section we investigate this finding further by means of a small additional Monte Carlo study as well as through some theoretical considerations.

Let us focus on the case where data are generated either as i.i.d., i.e. $Y_t = z_t$, z_t being i.i.d. $N(0,1)$, or as the ARCH(4), $Y_t = h_t^{1/2} z_t$, $h_t = 1 + 0.4 Y_{t-4}^2$. The null hypothesis is that the DGP belongs to the ARCH(8) class of models; hence, for both DGPs, the null hypothesis is satisfied; however, many of the unknown parameters (respectively, 8 for the i.i.d. case and 7 for the ARCH(4) case) lie on the boundary of the parameter space. In Figure 3 we report the empirical rejection probabilities of the standard and the shrinking-based bootstrap tests for nominal levels in the set $\{0.01, 0.02, ..., 0.25\}$. We consider samples of size $n \in \{600,\ 6000\}$; the number of Monte Carlo replications is set to $10,000$. From the inspection of the Figure 3 it can be noticed that indeed there are differences – albeit not large – between the standard bootstrap tests and the shrinking–based bootstrap tests, with the former being in terms of empirical size closer to the corresponding nominal level. Such differences, which as expected can be appreciated more at larger nominal levels, do not disappear as the sample size increases from $n = 600$ to $n = 6000$. Finally, a by-product result from this small exercise is that tests based on CvM tends to perform better than

Figure 3: Empirical type-I error rates for the bootstrap tests. In each panel, the Y-axis is the proportion of times the H_0 was rejected and the X-axis is the level of significance.



tests based on KS.

In terms of theory, in order to explain why the differences between the two bootstrap schemes are not dramatically large, we note the following. Recall that the large-sample behaviour of the original test statistics depends on the following expansion (see Lemma 1)

$$\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n) = \mathcal{U}_n(y, \boldsymbol{\phi}_0) - n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)' J(y, \boldsymbol{\phi}_0) + o_p(1). \tag{26}$$

The standard bootstrap fails because it is *unable* to mimic the term $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)$ on the right hand side when the true parameter is on the boundary. Specifically, while the asymptotic distribution of $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)$ is 'truncated normal' (see Lemma 2), its bootstrap analog $n^{1/2}(\hat{\boldsymbol{\phi}}_n^* - \hat{\boldsymbol{\phi}}_n)$ has a random limit. In contrast, when the shrinking bootstrap is employed, $n^{1/2}(\hat{\boldsymbol{\phi}}_n^* - \hat{\boldsymbol{\phi}}_n)$ approaches the correct limit. In general, the fact that the difference between the two bootstraps is not striking depends on the fact that the term $\mathcal{U}_n(\boldsymbol{\phi}_0; y)$ is of larger magnitude order than the second term, $n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)' J(y, \boldsymbol{\phi}_0)$. For instance, for the ARCH(1) DGP in Example 1, if the unknown parameter α lies on the boundary and z_t is i.i.d. $N(0,1)$, it is straightforward to see that (jointly)

$$\left(\mathcal{U}_n(y, \boldsymbol{\phi}_0), n^{1/2}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0)' J(y, \boldsymbol{\phi}_0)\right) \xrightarrow{d} \sqrt{2}\left(\sqrt{\Phi(y)} Z_1, \sqrt{1/3} J(y, \boldsymbol{\phi}_0) Z_2^+\right)$$

where $J(y, \boldsymbol{\phi}_0) = \mathrm{E}(z_t^2 \mathbb{I}(z_t \leq y))$, Φ is the $N(0,1)$ cdf, $Z_2^+ = \max\{0, Z_2\}$ and $Z_1, Z_2 \sim N(0,1)$. Hence, $\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n)$ has asymptotic representation

$$\mathcal{U}_n(y, \hat{\boldsymbol{\phi}}_n) \xrightarrow{d} \sqrt{2\Phi(y)} Z_1 + \sqrt{2/3} J(y, \boldsymbol{\phi}_0) Z_2^+.$$

Clearly, $\Phi(y)$ is larger than $J^2(y, \boldsymbol{\phi}_0)/3$ and, in this sense, $\mathcal{U}_n(y, \boldsymbol{\phi}_0)$ is the dominating term. Stated differently, the Z_2^+ term, which is not replicated by the standard bootstrap, has a small weight relative to the main term in (26), which is therefore the main driver of the asymptotic distributions of both the standard and the bootstrap test statistics. Because of this small weight, the invalidity of the standard bootstrap in simulations is often hard to see.

6. Empirical illustrations

To illustrate the bootstrap testing procedure, we briefly discuss two real data examples.

Application 1

We first consider a data example based on the daily log returns of the SPDR exchange-traded fund (ETF) for the S&P 500 index. This ETF is usually denoted by the tick symbol

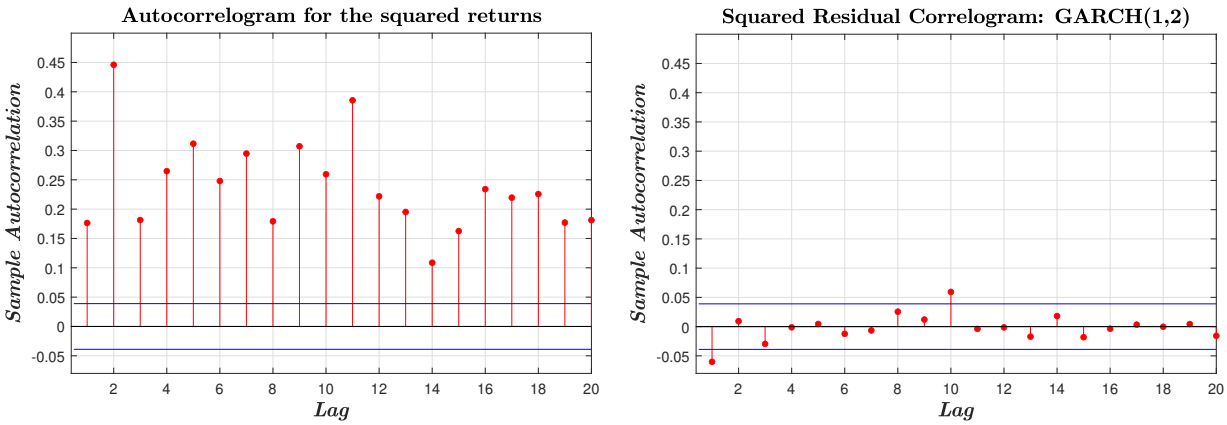
Table 2: P-values of the specification tests for testing GARCH(1,1) and GARCH(1,2) specifications for the conditional variance of the daily log-return of the SPDR ETF.

	Shrinking based bootstrap		Standard bootstrap		'm out of n' bootstrap		Hybrid bootstrap		
Null model	KS	CvM	KS	CvM	KS	CvM	KS	CvM	LBQ(20)
GARCH(1,1)	0.008	0.007	0.008	0.007	0.007	0.006	0.008	0.007	0.079
GARCH(1,2)	0.000	0.001	0.000	0.000	0.000	0.001	0.413	0.363	0.123

SPY. The data spans the period January 3, 2007 to June 30, 2017 and contains 2640 observations. A shorter version of this data set was previously studied by Tsay and Chen (2018), and by using some preliminary diagnostics, they concluded that a GJR-GARCH(1,1) model provides a good fit. In their empirical analysis, Tsay and Chen (2018) concluded that the leverage effect of the fitted GJR-GARCH(1,1) model is statistically significant at the 5% level. This indicates that if one specifies a GARCH(1,1) or a GARCH(1,2) model for the conditional variance then that may not provide a good fit for the data. In order to investigate this, in this empirical illustration, we employ the proposed KS and CvM bootstrap tests to test the adequacy of GARCH(1,1) and GARCH(1,2) specifications, expecting that the proposed tests would be able to detect a misspecification. For comparison the LBQ test considered in the simulations in the previous section is also considered.

In view of the results in Table 2, all the tests, except LBQ, convincingly reject the GARCH(1,1) specification with significantly small p-values. For the GARCH(1,2) specification, the shrinking based, standard, and 'm out of n' bootstrap p-values (of KS and CvM) are all zero up to 2 decimal places, whereas the p-values for the hybrid bootstrap based tests and LBQ(20) are larger than 0.1. Thus, the KS and CvM tests based on the shrinking based, standard, and 'm out of n' bootstrap methods clearly reject the GARCH(1,2) specification, but the hybrid bootstrap based tests and LBQ(20) fail to reject the GARCH(1,2) specification at the 10% level of significance. Since the hybrid bootstrap implementation is conservative and may produce larger p-values in finite samples (as evident from the simulations presented in the previous section), the results from the first three bootstrap approaches may be more reliable in this case. Note that the Ljung-Box Q test is designed to check the significance of the autocorrelations of the squared residuals at multiple lags

Figure 4: Autocorrelogram of the squared SPY log returns time series (first panel), and the squared residual correlogram for the fitted GARCH(1,2) model (second panel). The sample period is from January 3, 2007 to June 30, 2017.



jointly. Figure 4 shows the sample autocorrelations for both the squared values of the observed time series and the squared residuals estimated from the fitted GARCH(1,2). As expected, squared SPY log returns are significantly serially correlated, but the correlogram of squared residuals suggests no significant serial correlations except for some minor ones at lags 1 and 10. This explains the relatively large p-values of the LBQ test. However, squared residuals can be serially uncorrelated, but dependent, and hence it appears that the tests proposed in this paper are better suited than the Ljung-Box Q test in detecting the misspecification of the conditional variance specification in this case.

Application 2

In this illustrative example we consider a data set from the Caterpillar stock traded on the New York Stock Exchange. The variable of interest is the daily log return of the Caterpillar stock, defined by $Y_t = 100(\log P_t - \log P_{t-1})$ where P_t is the stock price at time t. The sample contains 2515 observations and spans the period Jan 02, 2001 to Dec 31, 2010. Tsay (2013) analyzed this data set by applying several diagnostic methods, and fitted a GARCH(1,1) model (see Table 5.1 in Tsay, 2013). When we fit a GARCH(1,2) model to this data set, the estimated GARCH(2) coefficient turns out to be statistically insignificant, practically at any level of significance. This indicates that, when testing the GARCH(1,2) specification, one component of the true parameter could potentially be a boundary point of the parameter space, whereas when the null model is GARCH(1,1) the true parameter could potentially be an interior point. Of course we do not have any certainty that this

Table 3: P-values of the specification tests for testing GARCH(1,1) and GARCH(1,2) specifications for the conditional variance of the daily log-return of the Caterpillar stock. The data spans the period Jan 02, 2001 to Dec 31, 2010.

	Shrinking based bootstrap		Standard bootstrap		'm out of n' bootstrap		Hybrid bootstrap		
Null model	KS	CvM	KS	CvM	KS	CvM	KS	CvM	LBQ(20)
GARCH(1,1)	0.362	0.314	0.362	0.314	0.219	0.227	0.362	0.314	0.999
GARCH(1,2)	0.499	0.416	0.499	0.416	0.312	0.289	0.500	0.461	0.999

is actually true. But, as an illustration, we employ the proposed KS and CvM bootstrap tests to test GARCH(1,1) and GARCH(1,2) specifications. For comparison the LBQ test is also considered. The p-values of the tests are given in Table 3. As expected the tests support both GARCH(1,1) and GARCH(1,2) specifications with large p-values. In the simulations in the previous section, the LBQ test was undersized when testing for the correct specification. Thus, the large p-values of the LBQ test in Table 3 are consistent with the simulation results reported in the previous section.

7. Conclusion

This paper contributes to advance the current statistical methodology for inference in GARCH models by developing bootstrap based omnibus specification tests while allowing parameters on the boundary of the parameter space. In particular, Kolmogorov-Smirnov and Cramér-von Mises type test statistics are proposed based on a certain empirical process marked by centered squared residuals. We first derive the asymptotic null distributions of the proposed test statistics when the true parameter is in the interior of the parameter space. Since the limiting distributions of the test statistics are not free from (unknown) nuisance parameters, we propose a bootstrap method to implement the tests and establish that the proposed bootstrap method is asymptotically valid and consistent. In view of our Monte Carlo simulations it appears that the finite sample performance of standard bootstrap is not strongly affected by the presence of parameters on the boundary. However, this bootstrap approach is not asymptotically valid if some components of the nuisance parameters lie on the boundary of the parameter space. Therefore, as an alternative, we

also propose a modified version of the bootstrap by employing a method of shrinkage of the parameter estimates in the bootstrap data generation. We show that the modified bootstrap procedure is asymptotically valid and consistent, regardless of the presence of nuisance parameters on the boundary. We also consider a hybrid version of this bootstrap approach that allows parameter values which are close to, but not on the boundary. Our bootstrap methods can be implemented easily under fairly general and easily verifiable assumptions and have desirable finite sample properties in terms of empirical size and power. In practical applications, they can be used as a first step in model selection, where validity of a general model is assessed via misspecification tests. Then, as a second step, the researcher can search for a more parsimonious model using, e.g., the tests in Cavaliere *et al.* (2022).

Our results can be extended in several directions. For instance, it is of interest to see if the methods we propose in this paper can be extended to models beyond the standard GARCH(p_1, p_2). To this end, consider the model $\mathcal{M}$ defined by

$$\mathcal{M}: \quad Y_t = h_t^{1/2}\varepsilon_t, \quad h_t = g_{\boldsymbol{\phi}}(Y_{t-1}, \cdots, Y_{t-p_1}, h_{t-1}, \cdots, h_{t-p_2}), \quad t \in \mathbb{Z}, \tag{27}$$

for some $\boldsymbol{\phi} \in \Phi \subset \mathbb{R}^{p_1+p_2}$, where $\{g_{\boldsymbol{\phi}}; \boldsymbol{\phi} \in \Phi\}$ is a parametric family of nonnegative functions on $\mathbb{R}^{p_1+p_2}$, and the error terms $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ are i.i.d. with zero mean and unit variance. Thus, $h_t = h_t(\boldsymbol{\phi}) = \mathrm{Var}(Y_t \mid \mathcal{H}_{t-1}),\ t \in \mathbb{Z}$. Consider the hypothesis testing problem

$$\mathsf{H}_0 : \text{Model } \mathcal{M} \text{ is correct} \qquad \text{vs} \qquad \mathsf{H}_1 : \text{Model } \mathcal{M} \text{ is not correct.} \tag{28}$$

The GARCH(p_1, p_2) model is a special case of $\mathcal{M}$. Another example is the asymmetric AGARCH(p_1, p_2) model defined by $h_t = h_t(\boldsymbol{\phi}) = \alpha_0 + \sum_{j=1}^{p_1} \alpha_j \big(|Y_{t-j}| - \gamma Y_{t-j}\big)^2 + \sum_{k=1}^{q} \beta_k h_{t-k}$, where $\boldsymbol{\phi} = (\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \gamma)'$, $\alpha_0 > 0, \alpha_j \geq 0$, $\beta_k \geq 0$ ($t \in \mathbb{Z}$, $1 \leq j \leq p$, $1 \leq k \leq q$). Similarly, several other extensions of the standard GARCH model can also be written in the general form (27).

Heuristic arguments suggest that the bootstrap tests proposed in this paper for ARCH(p) and GARCH(p_1, p_2) models can also be extended to this general setup. In fact, the bootstrap algorithm outlined in Section 4.2 can be readily applied to any model of the form (27), based on a suitable estimator for $\boldsymbol{\phi}$. However, the theory for QML parameter estimation, when the true value is on the boundary of the parameter space, in the family of models in (27), has not been developed yet. Therefore it is not a trivial task to extend the methods developed in this paper to a general setup of the form (27). One potential option would be to use penalized likelihood estimators as proposed by Fan and Li (2001), based on

smoothly clipped absolute deviation (SCAD) thresholding penalty functions. Alternatively, it may be possible to use Adaptive Lasso type estimators (see Zou, 2006). Furthermore, our testing procedures can also be potentially extended to Poisson autoregressions with exogenous covariates as considered in Agosto *et al.* (2016). All these extensions are left for future research.

Supplementary material

The supplementary material contains some additional simulation results, the proofs of the main results stated in the paper, as well as some auxiliary lemmas.

Acknowledgments

We thank the Co-Editor and two anonymous referees for their important comments which improved the quality of the paper. We also thank conference participants at the CFE 2022 conference at King's College London. This research was supported by the Danish Council for Independent Research (DSF Grant 015-00028B) and by the Italian Ministry of University and Research (PRIN 2020 Grant 2020B2AKFW). Rahbek gratefully acknowledges financial support from the Independent Research Fund Denmark (Grant no. 0133-00162B).

Disclosure statement

The authors report there are no competing interests to declare.

References

Agosto, A., Cavaliere, G., Kristensen, D., and Rahbek, A. (2016). Modeling corporate defaults: Poisson autoregressions with exogenous covariates (parx). *Journal of Empirical Finance*, **38**, 640–663.

Andrews, D. W. K. (2000). Inconsistency of the Bootstrap when a Parameter Is on the Boundary of the Parameter Space. *Econometrica*, **68**(2), 399–406.

Andrews, D. W. K. (2001). Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica*, **69**(3), 683–734.

Bai, J. (2003). Testing parametric conditional distributions of dynamic models. *The Review of Economics and Statistics*, **85**(3), 531–549.

Balakrishna, N., Koul, H. L., Sakhanenko, L., and Ossiander, M. (2019). Fitting a pth order parametric generalized linear autoregressive multiplicative error model. *Sankhya B*, **81**(1, suppl.), S103–S122.

Berkes, I. and Horváth, L. (2004). The efficiency of the estimators of the parameters in GARCH processes. *The Annals of Statistics*, **32**(2), 633–655.

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

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The Annals of Statistics*, **9**(6), 1196–1217.

Bickel, P. J., Klaassen, C. A. J., Ritov, Y., and Wellner, J. A. (1998). *Efficient and adaptive estimation for semiparametric models*. Springer-Verlag, New York. Reprint of the 1993 original.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**(3), 307–327.

Bougerol, P. and Picard, N. (1992a). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, **52**(1-2), 115–127.

Bougerol, P. and Picard, N. (1992b). Strict stationarity of generalized autoregressive processes. *The Annals of Probability*, **20**(4), 1714–1730.

Cavaliere, G., Nielsen, H. B., and Rahbek, A. (2017). On the consistency of bootstrap testing for a parameter on the boundary of the parameter space. *Journal of Time Series Analysis*, **38**(4), 513–534.

Cavaliere, G., Nielsen, H. B., Pedersen, R. S., and Rahbek, A. (2022). Bootstrap Inference On The Boundary Of The Parameter Space With Application To Conditional Volatility Models. *Journal of Econometrics*, **227**(1), 241–263.

Chatterjee, A. and Lahiri, S. N. (2011). Bootstrapping lasso estimators. *Journal of the American Statistical Association*, **106**(494), 608–625.

Chen, M. and An, H. Z. (1998). A note on the stationarity and the existence of moments of the GARCH model. *Statistica Sinica*, **8**(2), 505–510.

D'Agostino, R. B. and Stephens, M. A., editors (1986). *Goodness-of-fit techniques*, volume 68 of *Statistics: Textbooks and Monographs*. Marcel Dekker Inc., New York.

Doko Tchatoka, F. and Wang, W. (2021). Uniform Inference after Pretesting for Exogeneity with Heteroskedastic Data. MPRA Paper 106408, University Library of Munich, Germany.

Escanciano, J. C. (2007). Model checks using residual marked empirical processes. *Statistica Sinica*, **17**(1), 115–138.

Escanciano, J. C. (2008). Joint and marginal specification tests for conditional mean and variance models. *Journal of Econometrics*, **143**(1), 74–87.

Escanciano, J. C. (2010). Asymptotic distribution-free diagnostic tests for heteroskedastic time series models. *Econometric Theory*, **26**(3), 744–773.

Escanciano, J. C., Pardo-Fernández, J. C., and Keilegom, I. V. (2018). Asymptotic distribution-free tests for semiparametric regressions with dependent data. *The Annals of Statistics*, **46**(3), 1167 – 1196.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, **96**(456), 1348–1360.

Fernandes, M. and Grammig, J. (2005). Nonparametric specification tests for conditional duration models. *Journal of Econometrics*, **127**(1), 35–68.

Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli. Official Journal of the Bernoulli Society for Mathematical Statistics and Probability*, **10**(4), 605–637.

Francq, C. and Zakoian, J.-M. (2007). Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications*, **117**(9), 1265–1284.

Francq, C. and Zakoïan, J.-M. (2010). *GARCH models: structure, statistical inference and financial applications*. Wiley. John Wiley & Sons Ltd.

Giacomini, R., Politis, D. N., and White, H. (2013). A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory*, **29**(3), 567–589.

Grenander, U. and Rosenblatt, M. (1957). *Statistical analysis of stationary time series.* John Wiley & Sons, New York; Almqvist & Wiksell, Stockholm.

Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its application.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics.

Hall, P. and Yao, Q. (2003). Inference in arch and garch models with heavy-tailed errors. *Econometrica*, **71**, 285–317.

Hidalgo, J. and Zaffaroni, P. (2007). A goodness-of-fit test for ARCH(∞) models. *Journal of Econometrics*, **141**(2), 835–875.

Jensen, S. T. and Rahbek, A. (2004). Asymptotic inference for nonstationary GARCH. *Econometric Theory*, **20**(6), 1203–1226.

Ketz, P. (2018). Subvector inference when the true parameter vector may be near or at the boundary. *Journal of Econometrics*, **207**(2), 285–306.

Khmaladze, È. V. (1981). A martingale approach in the theory of goodness-of-fit tests. *Theory of Probability and Its Applications*, **26**(2), 240–257.

Kolmogorov, A. (1933). Sulla determinazione empirica di una lgge di distribuzione. *Inst. Ital. Attuari, Giorn.*, **4**, 83–91.

Koul, H. L. and Ling, S. (2006). Fitting an error distribution in some heteroscedastic time series models. *The Annals of Statistics*, **34**(2), 994–1012.

Koul, H. L. and Stute, W. (1999). Nonparametric model checks for time series. *The Annals of Statistics*, **27**(1), 204–236.

Koul, H. L., Perera, I., and Silvapulle, M. J. (2012). Lack-of-fit testing of the conditional mean function in a class of Markov multiplicative error models. *Econometric Theory*, **28**(6), 1283–1312.

Perera, I. and Koul, H. L. (2017). Fitting a two phase threshold multiplicative error model. *Journal of Econometrics*, **197**(2), 348–367.

Perera, I. and Silvapulle, M. J. (2021). Bootstrap based probability forecasting in multiplicative error models. *Journal of Econometrics*, **221**(1), 1–24.

Perera, I. and Silvapulle, M. J. (2022). Bootstrap specification tests for dynamic conditional distribution models. *Journal of Econometrics*, (in press).

Perera, I., Hidalgo, J., and Silvapulle, M. J. (2016). A goodness-of-fit test for a class of autoregressive conditional duration models. *Econometric Reviews*, **35**(6), 1111–1141.

Straumann, D. and Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equations approach. *The Annals of Statistics*, **34**(5), 2449–2495.

Stute, W. (1997). Nonparametric model checks for regression. *The Annals of Statistics*, **25**(2), 613–641.

Stute, W., Presedo Quindimil, M., González Manteiga, W., and Koul, H. L. (2006). Model checks of higher order time series. *Statistics & Probability Letters*, **76**(13), 1385–1396.

Tsay, R. and Chen, R. (2018). *Nonlinear Time Series Analysis*. Wiley Series in Probability and Statistics. Wiley.

Tsay, R. S. (2013). *An introduction to analysis of financial data with R*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ.

von Neumann, J. (1941). Distribution of the ratio of the mean square successive difference to the variance. *Annals of Mathematical Statistics*, **12**, 367–395.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, **101**(476), 1418–1429.