Journal of Multivariate Analysis 162 (2017) 32–50

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Inference for the autocovariance of a functional time series under conditional heteroscedasticity

Piotr Kokoszka [a,*], Gregory Rice [b], Han Lin Shang [c]

[a] Department of Statistics, Colorado State University, United States
[b] Department of Statistics and Actuarial Science, University of Waterloo, Canada
[c] Research School of Finance, Actuarial Studies and Statistics, Australian National University, Australia

ARTICLE INFO

Article history:
Received 6 June 2017
Available online 11 September 2017

AMS subject classifications:
62M10

Keywords:
Autocovariance
Conditional heteroskedasticity
Functional data

ABSTRACT

Most methods for analyzing functional time series rely on the estimation of lagged autocovariance operators or surfaces. As in univariate time series analysis, testing whether or not such operators are zero is an important diagnostic step that is well understood when the data, or model residuals, form a strong white noise. When functional data are constructed from dense records of, for example, asset prices or returns, a weak white noise model allowing for conditional heteroscedasticity is often more realistic. Applying inferential procedures for the autocovariance based on a strong white noise to such data often leads to the erroneous conclusion that the data exhibit significant autocorrelation. We develop methods for performing inference for the lagged autocovariance operators of stationary functional time series that are valid under general conditional heteroscedasticity conditions. These include a portmanteau test to assess the cumulative significance of empirical autocovariance operators up to a user selected maximum lag, as well as methods for obtaining confidence bands for a functional version of the autocorrelation that are useful in model selection/validation. We analyze the efficacy of these methods through a simulation study, and apply them to functional time series derived from asset price data of several representative assets. In this application, we found that strong white noise tests often suggest that such series exhibit significant autocorrelation, whereas our tests, which account for functional conditional heteroscedasticity, show that these data are in fact uncorrelated in a function space.

1. Introduction

Conditional heteroscedasticity is a frequently encountered feature of financial and economic time series that, since the seminal work of [6,10], is often modeled within the framework of generalized autoregressive conditionally heteroscedastic (GARCH) models. GARCH models have since been extended in many ways to multivariate observations in order to jointly model volatility across several return series, see [3,12,14], to name just a few references to univariate and multivariate GARCH models.

When evaluating such models for a given univariate or multivariate time series, one often begins by computing and plotting sample autocorrelations and/or autocovariances of the series with corresponding confidence intervals. The most common practice is to examine residual autocorrelations with the confidence bands computed under the assumption of a strong white noise. Such plots are useful in identifying models for the conditional mean, such as ARMA models, as well

* Corresponding author.
E-mail address: piotr.kokoszka@colostate.edu (P. Kokoszka).

http://dx.doi.org/10.1016/j.jmva.2017.08.004
0047-259X/

as for assessing whether a residual sequence is plausibly a white noise. It is well known that the standard error of sample autocorrelation estimates is strongly affected by conditional heteroscedasticity. In general, one should expect to see larger (in magnitude) autocorrelation estimates when considering a GARCH process, or a weak white noise exhibiting conditional heteroscedasticity, as compared to a strong white noise. If this effect is not properly accounted for, it may lead to an erroneous conclusion that a sequence exhibits autocorrelation. Several authors have developed diagnostic tests and model identification procedures for univariate and multivariate time series exhibiting conditional heteroscedasticity, including portmanteau tests, to measure the cumulative significance of the first K empirical autocovariances/autocorrelations; see, e.g., [9,11,28,32,36]. Many of these diagnostics are summarized in the monograph of Li [26].

The work referred to above presents methods applicable to multivariate time series of a relatively small dimension. Even for series of dimension 5–10, additional restrictions on the GARCH structure must be imposed to reduce the number of parameters to be estimated. In many financial and other applications, the data are very high dimensional multivariate observations (dozens or hundreds of coordinates), with a strong dependence between consecutive coordinates. The currently available multivariate goodness-of-fit techniques are not suitable for such data. We illustrate with data that motivate this work and will be further studied below. On consecutive days $i \in \{1, \ldots, T\}$, observations of the S&P 500 index are available at intraday times u, measured at a 1-min (or finer) resolution. These data may then be represented by a sequence of discretely observed curves or functions $\{P_i(u) : 1 \le i \le T, u \in [0, S]\}$, with S denoting the length of the trading day. Transformations of these functions that are of interest include the horizon h log returns, $R_i(u) = \ln P_i(u) - \ln P_i(u-h)$, where h is some given length of time, e.g., five minutes. For a fixed h, on any given trading day i we thus observe a high-dimensional multivariate vector, which can be viewed as a noisy function. We thus observe one function per day, i.e., a functional time series.

The scope of methodology for analyzing functional time series such as these has grown substantially in the last decade; we refer to [7], Chapters 13–16 of [17] and Chapter 8 of [24] for summaries of advances in the field. To date, most methods in this direction are still based on non-parametric and non-likelihood approaches that rely fundamentally on the estimation of autocovariance operators, see, e.g., [20,21,30,41]. Several authors have considered portmanteau tests for these operators; see [13,23,40]. In the presence of functional conditional heteroscedasticity, diagnostic tests for measuring the significance of autocovariance functions and general checks for model adequacy are needed. Specific functional conditionally heteroskedastic models have been proposed by [2,15]. Our methodology is more broadly applicable. It assumes a flexible nonparametric quantification of conditional heteroskedasticity.

In this paper, we develop diagnostic tests and visualization tools based on the empirical autocovariance functions of stationary functional time series exhibiting conditional heteroscedasticity. In particular, we derive a portmanteau test to measure the cumulative significance of the norms of the first K empirical autocovariance functions under a general conditional heteroscedasticity assumption. This test may be used to evaluate the adequacy of functional GARCH-type models for observed or residual curves. Building upon this theory, we further develop methods to construct confidence bands for empirical functional autocorrelation estimates that are useful for identifying correlation at specific lags and informing further modeling. As a demonstration, we apply the proposed methods to functional time series derived from the intraday returns on several densely observed asset prices. Our analysis suggests that the level of autocorrelation observed in these series is typically more than would be expected from a strong white noise model, but is in fact in accordance with functional conditional heteroscedasticity.

The paper is organized as follows. Its main methodological contributions are presented in Section 2. After formulating a suitable mathematical framework, we state the assumptions, including a general definition of functional conditional heteroscedasticity. We then define two test statistics, one to measure the autocovariance at a specified lag and another, a portmanteau statistic, to measure the cumulative significance of the first K autocovariance functions. We establish their asymptotic properties and explain how to perform the tests in practice. We also introduce confidence bands for quantities akin to autocorrelations of scalar time series. The results of a simulation study of the proposed methods are presented in Section 3, which also contains their application to intraday returns on several types of assets. Proofs of the asymptotic results and some technical calculations are presented in Appendix.

2. Problem statement and testing procedures

2.1. Mathematical framework and notation

We consider a discretely observed functional time series $\{X_i(u) : 1 \le i \le T, u \in (0, 1]\}$, of length T. It is convenient to think of i as denoting the day. The "intraday" parameter u is rescaled to be in the interval $(0, 1]$, without loss of generality. We write f for a function $f(u)$, $u \in (0, 1]$ and further use $\int$ to denote $\int_0^1$ when it does not cause confusion, and use (X_i) to denote the sequence $(X_i)_{i\in\mathbb{Z}}$. Before proceeding, we describe a convenient mathematical framework for the data: we treat each X_i as an element of the Hilbert space of real-valued square integrable functions defined on the interval $(0, 1]$ equipped with a nonnegative measure μ on its Borel subsets. We define the inner product between two functions f and g by

$$\langle f, g\rangle = \int f(u)g(u)\mu(du). \tag{1}$$

We then define $L^2(\mu)$ to be the Hilbert space of functions f for which $\|f\| = \langle f, f\rangle^{1/2} < \infty$, and we assume that each X_i is an element of $L^2(\mu)$.

There are several advantages to considering this general framework. For instance, if the data consist of price curves $P_i(u)$ observed at J intraday times $\mathcal{U}_J = \{u_j = j/J,\ j = 1, \ldots, J\} \subset (0, 1]$, it is then convenient to work in the Hilbert space $L^2(\mu_J)$, where μ_J is the counting measure on $\mathcal{U}_J$ divided by J. The inner product in $L^2(\mu_J)$ is then given by

$$\langle f, g\rangle = \int f(u)g(u)\mu_J(du) = \frac{1}{J}\sum_{j=1}^{J} f\,(j/J)\,g\,(j/J)\,.$$

On a typical trading day, if 1-min resolution price data are available, then $J = 390$. For data observed at different frequencies, like tick data, it might make more sense to take μ to be standard Lebesgue measure, and estimate the inner-products (1) with Riemann sums as the data allows. We use the same notation as above to denote the standard inner product and norm on $L^2(\mu^{\otimes d})$, with the dimension d being clear based on the input function, and employ tensor notation; see, e.g., Section 10.5 of [24]. To illustrate, in our context, the condition $\mathrm{E}\langle X_i \otimes X_j \otimes X_k \otimes X_\ell,\, w\rangle = 0$, which appears below, can be written explicitly as

$$\mathrm{E}\iiiint X_i(t)X_j(s)X_k(u)X_\ell(v)w(t, s, u, v)\mu(dt)\mu(ds)\mu(du)\mu(dv) = 0.$$

2.2. Problem formulation and assumptions

When modeling functional time series, one frequently wishes to determine whether or not a given series $\{X_i\}$, which might consist of the original observations or of the residuals of a model fit, is plausibly a weak white noise, i.e., if for $h > 0$, $\gamma_h(t, s) = 0$, where

$$\gamma_h(t, s) = \mathrm{E}\left[\{X_0(t) - \mu_X(t)\}\{X_h(s) - \mu_X(s)\}\right], \quad \mu_X(t) = \mathrm{E}X_0(t). \tag{2}$$

This would be the case if (X_i) were in fact a strong white noise, i.e., a sequence of independent and identically distributed random functions, but it would also hold if (X_i) followed a functional GARCH-type model. To illustrate, consider the functional GARCH(1,1) model

$$X_i = \sigma_i\varepsilon_i, \quad \sigma_i^2 = \delta + \alpha(X_{i-1}^2) + \beta(\sigma_{i-1}^2), \tag{3}$$

where (ε_i) is a strong white noise sequence and α and β are linear operators mapping $L^2(\mu)$ to $L^2(\mu)$. This is precisely the model put forward by [2]. It is a weak functional white noise, but not a strong white noise. As for the strong white noise, its lag h autocovariance functions are zero for all $h > 0$. However, the variability of the estimates of the functions $\gamma_h(\cdot, \cdot)$ is different in case of the strong white noise and GARCH-type functional time series. This has to be taken into account in inferential goodness-of-fit procedures. Developing suitable tests and confidence intervals is the objective of this paper.

In the first direction, we aim to develop methods to assess the validity of the hypotheses

$$\mathcal{H}_{0,h}:\ \gamma_h(t, s) = 0 \quad \text{and} \quad \mathcal{H}'_{0,K}: \forall_{j\in\{1,\ldots,K\}} \quad \gamma_j(t, s) = 0$$

that are consistent under functional conditional heteroscedasticity, which will be defined in the following. The null hypothesis $\mathcal{H}_{0,h}$ is to be tested for a fixed $h > 0$. We will then also introduce an approach akin to confidence bands for sample autocorrelations of scalar time series.

The key assumptions involve a quantification of the decay of temporal dependence, which may be present even if autocorrelations vanish, with lag separation, and vanishing second and fourth order moments, as implied by a GARCH structure. To specify the dependence structure, we formulate the following definition:

Definition 1. A sequence (X_i) is said to be L^p-m-approximable if each X_i admits the Bernoulli shift representation

$$X_i = g\,(\varepsilon_i, \varepsilon_{i-1}, \ldots)\,, \tag{4}$$

where the (ε_i) are i.i.d. elements taking values in a measurable space S, and g is a measurable function $g : S^\infty \to L^2(\mu)$. Moreover, if (ε_i') is an independent copy of the sequence (ε_i) defined on the same probability space, then

$$\sum_{m=0}^{\infty}(\mathrm{E}\|X_i - X_i^{(m)}\|^p)^{1/p} < \infty, \tag{5}$$

where

$$X_i^{(m)} = g(\varepsilon_i, \varepsilon_{i-1}, \ldots, \varepsilon_{i-m+1}, \varepsilon'_{i-m}, \varepsilon'_{i-m-1}, \ldots). \tag{6}$$

The gist of Definition 1 is that the dependence of g in (4) on the innovations far in the past decays so fast that these innovations can be replaced by their independent copies. Such a replacement is asymptotically negligible in the sense quantified by (5), which also implies that $\mathrm{E}\|X_i\|^p < \infty$. Representation (4) implies that the sequence (X_i) is stationary and ergodic. Assumptions similar to Definition 1 have been used extensively in recent theoretical work, as all stationary time

series models in practical use can be represented as Bernoulli shifts, see [1,4,16,22,34,38,40], among other contributions. Unlike a linear moving average, representation (4) admits heteroskedastic time series.

With this background, we state our first assumption.

Assumption 1. The time series (X_i) is L^4-m-approximable.

The next assumption requires that the sequence (X_i) possesses the first, second, and fourth order moment characteristics of a functional GARCH-type sequence.

Assumption 2. The sequence (X_i) satisfies

(i) if $u \in L^2(\mu)$, $\mathrm{E}\,\langle X_i, u\rangle = 0$ for all i;
(ii) if $v \in L^2(\mu \otimes \mu)$, and $i \neq j$, $\mathrm{E}\langle X_i \otimes X_j, v\rangle = 0$;
(iii) if $w \in L^2(\mu^{\otimes 4})$, and if the indices $i, j, k, \ell \in \mathbb{Z}$ have a unique maximum, then $\mathrm{E}\langle X_i \otimes X_j \otimes X_k \otimes X_\ell, w\rangle = 0$.

Assumptions 1 and 2 hold under general conditions for sequences satisfying functional GARCH equations, for instance for the functional GARCH(1,1) in (3). Although only standard Lebesgue measure is considered in [2], extensions of their conditions to an arbitrary measure μ are trivial, as they are formulated in terms of Hilbert space norms. Sufficient conditions for Assumption 1 to hold under the model (3) are established in Theorem 2.1 and Corollary 2.1 in [2], which state that if α and β in (3) are integral operators with kernels $a(t,s)$ and $b(t,s)$, respectively, then there exists a nonanticipative solution (X_i) to the equations in (3) if

$$\mathrm{E}\left[\ln\left[\left\{\iint r_0^2(t,s)\mu(dt)\mu(ds)\right\}^{1/2}\right]\right] < 0,$$

and

$$\mathrm{E}\left[\left\{\iint r_0^2(t,s)\mu(dt)\mu(ds)\right\}^{\nu/2}\right] < 1,$$

for some $\nu \geq 4$, where $r_0(t,s) = a(t,s)\varepsilon_0^2(s) + b(t,s)$. Moreover, this solution satisfies Assumptions 1 and 2.

2.3. Test statistics and their limit distributions

The autocovariance function γ_h is estimated by

$$\hat{\gamma}_h(t,s) = \frac{1}{T}\sum_{j=1}^{T-h}\{X_j(t) - \bar{X}(t)\}\{X_{j+h}(s) - \bar{X}(s)\}, \quad \bar{X}(t) = \frac{1}{T}\sum_{j=1}^{T} X_j(t),$$

and through these estimates we define test statistics for $\mathcal{H}_{0,h}$ and $\mathcal{H}'_{0,K}$ by

$$Q_{T,h} = T\|\hat{\gamma}_h\|^2 \quad \text{and} \quad V_{T,K} = T\sum_{h=1}^{K}\|\hat{\gamma}_h\|^2,$$

respectively. In this section we state large-sample properties of $Q_{T,h}$ and $V_{T,K}$. Even though our work is motivated by the functional GARCH models of [2,15], the asymptotic distribution of $Q_{T,h}$ and $V_{T,K}$ may be established for a larger class of nonlinear functional time series satisfying Assumptions 1 and 2.

Under Assumption 1, the autocovariance γ_h is an element of $L^2(\mu \otimes \mu)$, since

$$\iint \gamma_h^2(t,s)\mu(dt)\mu(ds) \leq \int \mathrm{E}\{X_1^2(t)\mu(dt)\}\int \mathrm{E}\{X_{1+h}^2(s)\mu(ds)\} = \{\mathrm{E}\,\|X_1\|^2\}^2.$$

A slightly longer, but similar, argument shows that

$$\mathrm{E}\iint \hat{\gamma}_h^2(t,s)\mu(dt)\mu(ds) < \infty.$$

Beginning with the test of $\mathcal{H}_{0,h}$, we assume in the following theorem that the lag $h \geq 1$ is fixed.

Theorem 1. *Suppose the functional time series (X_i) satisfies Assumptions 1 and 2. Then, there are mean zero Gaussian processes $\Gamma_{h,T}$, $T \geq 1$, in $L^2(\mu \otimes \mu)$, defined on the same, perhaps enlarged, probability space as (X_i), such that*

$$\iint \{\sqrt{T}\hat{\gamma}_h(t,s) - \Gamma_{h,T}(t,s)\}^2\mu(ds)\mu(dt) = o_P(1).$$

Their covariances do not depend on T, and for t, s, t', s' in the support of μ, are given by

$$c_{h,h}(t,s,t',s') = \mathrm{E}\left\{\Gamma_{h,T}(t,s)\Gamma_{h,T}(t',s')\right\} = \mathrm{E}\left\{X_0(t)X_0(t')X_h(s)X_h(s')\right\}. \tag{7}$$

Corollary 1. *Under the assumptions of Theorem 1,* $Q_{T,h} \rightsquigarrow \|\Gamma_{h,0}\|^2$.

(Throughout the paper, $\rightsquigarrow$ denotes convergence in distribution.) Theorem 1 is proven in Appendix. Corollary 1 is a simple consequence, but, for completeness, is also proven in Appendix. It shows that an asymptotic size α test of $\mathcal{H}_{0,h}$ is to reject when $Q_{T,h} > \Xi_{h,1-\alpha}$, where $\Xi_{h,1-\alpha}$ is the $1-\alpha$ quantile of the distribution of $\|\Gamma_{h,0}\|^2$. In Section 2.4, we explain how to obtain a feasible approximation to this distribution.

We now turn to the test of $\mathcal{H}'_{0,K}$ and the asymptotic properties of $V_{T,K}$.

Theorem 2. *Suppose the functional time series* (X_i) *satisfies Assumptions 1* and *2, and* K *is a positive integer. Then, there exist positive real constants* $(\xi_{\ell,K})_{\ell=1}^{\infty}$*, depending on the functions defined, for each* $i, j \in \{1, \ldots, K\}$*, by*

$$\mathrm{E}\{X_{-i}(t)X_0(s)X_{-j}(u)X_0(v)\}, \tag{8}$$

and satisfying $\sum_{i=1}^{\infty}\xi_{i,K} < \infty$*, such that*

$$V_{T,K} \rightsquigarrow V_K \stackrel{D}{=} \sum_{\ell=1}^{\infty}\xi_{\ell,K}\mathcal{N}_\ell^2, \tag{9}$$

where $(\mathcal{N}_\ell)_{\ell=1}^{\infty}$ *are independent and identically distributed standard normal random variables.*

The coefficients $\xi_{\ell,K}$ are defined by (A.6) and, in short, are the eigenvalues of a quite complicated covariance operator derived from the functions in (8). The direct estimation of these coefficients becomes infeasible even for small values of K, and so in order to test $\mathcal{H}'_{0,K}$ in practice, we propose to approximate the distribution of V_K on the right hand side of (9) using a Welch–Satterthwaite type χ^2 approximation, as explained in Section 2.4.

For a series (X_i) that has nonzero autocorrelation at lag h, it follows fairly directly from the theory developed to prove Theorems 1 and 2 that $Q_{T,h} \stackrel{P}{\to} \infty$ and that $V_{T,K} \stackrel{P}{\to} \infty$ for all $h \in \{1, \ldots, K\}$. We quantify nonzero autocorrelation at lag h by the following assumption:

Assumption 3. There exists a function $a_h \in L^2(\mu \otimes \mu)$ such that $\mathrm{E}\langle X_0 \otimes X_h, a_h\rangle \neq 0$.

Theorem 3. *If the functional time series satisfies Assumptions 1, 2*(i)*, and 3*(h)*, then* $Q_{T,h} \stackrel{P}{\to} \infty$*, as* $T \to \infty$*. As long as* $h \in \{1, \ldots, K\}$*, one also has* $V_{T,K} \stackrel{P}{\to} \infty$*, as* $T \to \infty$*.*

Local alternatives can be defined by introducing functional triangular arrays such that $\mathrm{E}\{X_{0,T}(t)X_{h,T}(s)\} = T^{-1/2}a_h(t,s)$ $\mu \otimes \mu$ almost everywhere. The remaining assumptions must be suitably modified for such arrays, and it can then be shown that the tests are consistent under such local alternatives. To save space, details are not presented.

2.4. Feasible approximations to the limit distributions

In order to make the results of Section 2.3 applicable to price data, we now detail how the tests are carried out when the price curves are observed on the regularly spaced grid $\mathcal{U}_J$ with the norm computed using the normalized counting measure μ_J. In this case, $Q_{T,h}$ reduces to

$$Q_{T,h}^{(J)} = \frac{T}{J^2}\sum_{j=1}^{J}\sum_{k=1}^{J}\hat{\gamma}_h^2(u_j, u_k).$$

Similarly,

$$\|\Gamma_{h,0}\|^2 = \iint \Gamma_{h,0}^2(u,s)\mu_J(du)\mu_J(ds) = \frac{1}{J^2}\sum_{j=1}^{J}\sum_{k=1}^{J}\Gamma_{h,0}^2(u_j, u_k),$$

and the covariance of the process $\{\Gamma_{h,0}(u_j, u_k),\ u_j, u_k \in \mathcal{U}_J\}$ is given by

$$c_{h,h}(u_j, u_k, u_{j'}, u_{k'}) = \mathrm{E}\{\Gamma_{h,0}(u_j, u_k)\Gamma_{h,0}(u_{j'}, u_{k'})\} = \mathrm{E}\{X_0(u_j)X_0(u_{j'})X_h(u_k)X_h(u_{k'})\}.$$

If $\lambda_{h,\ell}$, $\ell \geq 1$, are the eigenvalues of this 4D tensor of dimension J^4 defined by

$$\frac{1}{J^2}\sum_{j',k'=1}^{J} c_{h,h}(u_j, u_k, u_{j'}, u_{k'})\varphi_{\ell,h}(u_{j'}, u_{k'}) = \lambda_{\ell,h}\varphi_{\ell,h}(u_j, u_k), \tag{10}$$

then it follows from the Karhunen–Loève Theorem (see p. 25 of [7]) that

$$\frac{1}{J^2}\sum_{j=1}^{J}\sum_{k=1}^{J}\Gamma_{h,0}^2(u_j, u_k) \stackrel{D}{=} \sum_{\ell=1}^{\infty}\lambda_{h,\ell}\mathcal{N}_\ell^2,$$

where $(\mathcal{N}_\ell)_{\ell=1}^\infty$ are independent and identically distributed standard normal random variables. This leads to the conclusion that, as $T \to \infty$,

$$Q_{T,h}^{(J)} \rightsquigarrow \sum_{\ell=1}^{\infty} \lambda_{h,\ell} \mathcal{N}_\ell^2,$$

where the $\lambda_{h,\ell}$ are the eigenvalues defined in (10).

In practice, the eigenvalues $\lambda_{\ell,h}$ are estimated by $\hat{\lambda}_{\ell,h}$ satisfying

$$\frac{1}{J^2} \sum_{j',k'=1}^{J} \hat{c}_{h,h}(u_j, u_k, u_{j'}, u_{k'}) \hat{\varphi}_{\ell,h}(u_{j'}, u_{k'}) = \hat{\lambda}_{\ell,h} \hat{\varphi}_{\ell,h}(u_j, u_k), \tag{11}$$

where

$$\hat{c}_{i,j}(t, s, u, v) = \frac{1}{T} \sum_{k=1+\max\{i,j\}}^{T} X_{k-i}^c(t) X_k^c(s) X_{k-j}^c(u) X_k^c(v), \quad X_i^c(t) = X_i(t) - \bar{X}(t). \tag{12}$$

With functional data observed on $\mathcal{U}_J$, the tensor $\hat{c}_{h,h}(u_j, u_k, u_{j'}, u_{k'})$ can readily be constructed from the data, and eigenvalues $\hat{\lambda}_{\ell,h}$ may be computed using the function `svd.tensor` in the R package `tensorA` [37]. One may then estimate $\Xi_{h,1-\alpha}$ with $\hat{\Xi}_{h,1-\alpha}$, the $1-\alpha$ quantile of

$$\sum_{\ell=1}^{J^2} \hat{\lambda}_{h,\ell} \mathcal{N}_\ell^2, \tag{13}$$

which can be calculated using either Monte Carlo simulation, or directly using the function `imhoff` in R. We then reject $\mathcal{H}_{0,h}$ if $Q_{T,h} > \hat{\Xi}_{h,1-\alpha}$. The finite-sample properties of this test are studied in Section 3.

We now turn to the approximation to the limit V_K in Theorem 2. As noted in Section 2.3, we use the Welch–Satterthwaite approximation, see [25,39]. The idea is to approximate the limiting distribution V_K by a random variable $R_K \sim \beta \chi_\nu^2$, where χ_ν^2 denotes a χ^2 random variable with ν degrees of freedom, and β and ν are estimated so that the distribution of R_K has the same first two moments as the distribution of V_K. With $\mu_{V,K} = \mathrm{E}(V_K)$ and $\sigma_{V,K}^2 = \mathrm{var}(V_K)$, we then take

$$\beta = \frac{\sigma_{V,K}^2}{2\mu_{V,K}} \quad \text{and} \quad \nu = \frac{2\mu_{V,K}^2}{\sigma_{V,K}^2}. \tag{14}$$

We verify in Appendix that

$$\mu_{V,K} = \sum_{i=1}^{K} \iint \mathrm{E}\left\{X_0^2(t) X_i^2(s) \mu(dt)\mu(ds)\right\} \quad \text{and} \quad \sigma_{V,K}^2 = \sum_{1 \le i,j \le K} \eta_{i,j}, \tag{15}$$

where

$$\eta_{i,j} = 2 \iiiint [\mathrm{E}\{X_{-i}(t) X_0(s) X_{-j}(u) X_0(v)\}]^2 \mu(dt)\mu(ds)\mu(du)\mu(dv).$$

These parameters can be consistently estimated with simple plug-in estimators: we estimate $\mu_{V,K}$ and $\sigma_{V,K}^2$ with

$$\hat{\mu}_{V,K} = \sum_{i=1}^{K} \iint \hat{c}_{i,i}(t, s, t, s) \mu(dt)\mu(ds) \quad \text{and} \quad \hat{\sigma}_{V,K}^2 = \sum_{1 \le i,j \le K} \hat{\eta}_{i,j},$$

where

$$\hat{\eta}_{i,j} = 2 \iiiint \hat{c}_{i,j}^2(t, s, u, v) \mu(dt)\mu(ds)\mu(du)\mu(dv),$$

and where $\hat{c}_{i,j}$ is defined in (12). Estimates $\hat{\beta}$ and $\hat{\nu}$ are then defined with these estimates as in (14). An approximate size α test of $\mathcal{H}'_{0,K}$ is to reject when $V_{T,K} > \hat{\beta} \chi_{\hat{\nu},1-\alpha}^2$, where $\chi_{\nu,1-\alpha}^2$ is the $1-\alpha$ quantile of the χ^2 distribution with ν degrees of freedom.

The evaluation of the estimates $\hat{\eta}_{i,j}$ in order to carry out the two-moment χ^2 approximation for the distribution of $V_{T,K}$ requires the calculation of a four-variate integral on the 4-dimensional unit hypercube of $\hat{c}_{i,j}^2$. When the functions are observed on $\mathcal{U}_J$, this integral can in principle be approximated by the simple Riemann sum

$$\hat{\eta}_{i,j} = \frac{2}{J^4} \sum_{i,j,k,r=1}^{J} \hat{c}_{i,j}^2(u_i, u_j, u_k, u_r). \tag{16}$$

This calculation can become computationally intensive if J is large. To increase the speed of this calculation at the cost of a small loss in accuracy, we propose two Monte Carlo integration techniques to evaluate $\hat{\eta}_{i,j}$. The first is based on applying a trapezoidal rule to a sparse, randomly selected, regular grid, which we abbreviate by "rTrap" below. For each $i \in \{1, \ldots, R\}$, let v_i denote an ordered (increasing) random sample without replacement from $\mathcal{U}_J$. We may then estimate $\eta_{i,j}$ with

$$\hat{\eta}_{i,j} = 2 \sum_{p,q,k,r=1}^{R} \left[\frac{1}{2} \left\{ \hat{c}_{i,j}^2(v_p, v_q, v_k, v_r) + \hat{c}_{i,j}^2(v_{p-1}, v_{q-1}, v_{k-1}, v_{r-1}) \right\} \right] \left\{ \prod_{\ell \in \{p,q,k,r\}} (v_\ell - v_{\ell-1}) \right\},$$

where $v_0 = 0$. This method employs a sparser grid of points as compared to (16), which makes it simpler to compute, as well as a trapezoidal rule that averages the value of the function $\hat{c}_{i,j}^2$ over the two most extremal points of the grid. The second method, which we abbreviate "MCint", is based on standard Monte Carlo integration. In this case, let $(v_{k,1}, \ldots, v_{k,4})$ with $k \in \{1, \ldots, M\}$ denote points selected uniformly at random and with replacement from $\mathcal{U}_J \otimes \mathcal{U}_J \otimes \mathcal{U}_J \otimes \mathcal{U}_J$. We may then estimate $\eta_{i,j}$ with

$$\hat{\eta}_{i,j} = \frac{2}{M} \sum_{k=1}^{M} \hat{c}_{i,j}^2(v_{k,1}, v_{k,2}, v_{k,3}, v_{k,4}).$$

Both methods cut down substantially on the computational time required to provide an estimate of $\eta_{i,j}$, mainly because they each do not require the calculation of the entire function $\hat{c}_{i,j}$. The relative advantages of one method over the other, even asymptotically, are unclear since they would depend on the properties of the unknown function $c_{i,j}$. We compared both methods below, and found that they produced very similar results for the data generating processes we considered.

Choice of R **and** M: In practice one should choose R and M to be as large as time and computational resources allow. In our experience taking $R = 25$ and $M = 2000$ produces reliable results, and with these settings it takes typically less than one minute to calculate all integrals needed to perform the test based on $V_{T,10}$ for $T \leq 500$ on a modern laptop computer.

2.5. Confidence bands for functional autocorrelation measures

Following [18], we define the functional autocorrelation coefficient at lag h to be

$$\rho_h = \frac{\|\gamma_h\|}{\int \gamma_0(t,t)\mu(dt)}, \quad \|\gamma_h\|^2 = \iint \gamma_h^2(t,s)\mu(dt)\mu(ds),$$

where $\gamma_h(t,s)$ is defined by (2). One can readily verify using the Cauchy–Schwarz inequality that $\rho_h \in [0, 1]$. We estimate ρ_h with

$$\hat{\rho}_h = \frac{\|\hat{\gamma}_h\|}{\int \hat{\gamma}_0(t,t)\mu(dt)} = \frac{\sqrt{Q_{T,h}}}{\sqrt{T} \int \hat{\gamma}_0(t,t)\mu(dt)}. \tag{17}$$

It follows then, with $\hat{\Xi}_{h,1-\alpha}$ again denoting the estimated asymptotic $1-\alpha$ quantile of $Q_{T,h}$, that

$$\widehat{B}_h(1-\alpha) = \frac{\sqrt{\hat{\Xi}_{h,1-\alpha}}}{\sqrt{T} \int \hat{\gamma}_0(t,t)\mu(dt)}$$

is an asymptotic upper $1-\alpha$ confidence bound for ρ_h. We can similarly compute such a bound under the assumption that the sequence (X_i) forms a strong white noise. In this case the limiting covariance defined in (7) does not depend on the lag h, and is of the form

$$c_0^\star(t,s,u,v) = \mathrm{E}\left[\{X_0(t)X_0(u)\}\,\mathrm{E}\{X_0(s)X_0(v)\}\right].$$

With discrete data on $\mathcal{U}_J$, the eigenvalues of the covariance operator with kernel $c^\star$ can be estimated from the tensor

$$c_0^\star(t_j, t_k, t_{j'}, t_{k'}) = \left\{ \frac{1}{T} \sum_{i=1}^{T} X_i^c(t_j) X_i^c(t_{j'}) \right\} \left\{ \frac{1}{T} \sum_{i=1}^{T} X_i^c(t_k) X_i^c(t_{k'}) \right\}$$

from which a quantile $\hat{\Xi}_{\text{i.i.d.},1-\alpha}$ can be calculated as described in (13) and the sentences that follow. This gives a $1-\alpha$ confidence bound for ρ_h under the strong white noise assumption of

$$\widehat{B}_{\text{i.i.d.}}(1-\alpha) = \frac{\sqrt{\hat{\Xi}_{\text{i.i.d.},1-\alpha}}}{\sqrt{T} \int \hat{\gamma}_0(t,t)\mu(dt)}.$$

Plots of $\hat{\rho}_h$, $\widehat{B}_h(1-\alpha)$, and $\widehat{B}_{\text{i.i.d.}}(1-\alpha)$ are useful in model identification under potential conditional heteroscedasticity. Examples and applications are considered in Section 3.

3. Finite-sample performance and application to price data

We first show that the tests proposed above perform well in finite samples by reporting rejection rates for simulated data. We also demonstrate the usefulness of the confidence bands for the functional autocorrelation defined in Section 2.5. We then apply the above methods to intraday returns on assets in several different classes. In particular, our tests supply evidence that the functional GARCH model (3) is compatible with the autocorrelation observed in these assets.

3.1. Simulation study

In order to evaluate finite-sample properties of the proposed tests, we used the following data generating processes (DGP's):

(a) IID: $X_i(u) = W_i(u)$, where $\{W_i(u),\ u \in [0, 1]\}_{i=1}^{\infty}$ is a sequence of independent and identically distributed Brownian motions.
(b) fGARCH(1,1): $X_i(u)$ follows (3), where α and β are integral operators defined, for $x \in L^2(\mu)$ and $t \in [0, 1]$, by

$$(\alpha x)(t) = \int \alpha(t, s)x(s)\mu(ds), \quad (\beta x)(t) = \int \beta(t, s)x(s)\mu(ds),$$

where $\alpha(t, s) = \beta(t, s) = 12t(1 - t)s(1 - s)$. We set $\delta = 0.01$ (a constant function), and

$$\epsilon_i(u) = \frac{\sqrt{\ln(2)}}{2^{200u}} B_i\left(\frac{2^{400u}}{\ln 2}\right), \quad u \in [0, 1],$$

where $\{B_i(u), u \in [0, 1]\}_{i\in\mathbb{Z}}$ are independent and identically distributed Brownian bridges.
(c) FAR(1, S)-IID: $X_i(u) = \int \psi_c(u, s)X_{i-1}(s)\mu(ds) + \varepsilon_i(u)$, $u \in [0, 1]$, where ε_i follows the DGP IID, and $\psi_c(t, s) = c \exp\{-(t^2 + s^2)/2\}$. The constant c is then chosen so that $\|\psi_c\| = S \in (0, 1)$.
(d) FAR(1, S)-fGARCH(1,1): Same as above, but with ε_i following fGARCH(1,1).

If $S > 0$, the two FAR processes represent data for which the null hypotheses are violated. The specifications for the fGARCH(1,1) process are the same as those studied in [2], and satisfy Assumptions 1 and 2. For the fGARCH(1,1) and both FAR(1, S) processes, we used a burn-in sample of length 50 before producing a sample of length T. In all DGPs, each functional observation was simulated on an equally spaced grid of $J = 50$ points on the unit interval. We did not notice any significant difference in the results when we studied observations following these specifications generated on a more refined subset of the unit interval. The measure μ is taken to be μ_J.

We begin by reporting simulation results aimed at assessing the empirical size of the tests of $\mathcal{H}_{0,h}$ and $\mathcal{H}'_{0,K}$ described in Section 2. We did not notice any pronounced difference in terms of size when testing $\mathcal{H}_{0,h}$ for $h \in \{1, \ldots, 5\}$, so we report the results for $h = 1$ only. We further note that in this case the hypotheses $\mathcal{H}_{0,1}$ and $\mathcal{H}'_{0,1}$ are equivalent, as are the test statistics $Q_{T,1}$ and $V_{T,1}$. However, the tests differ because the rejection regions are computed in a different way. We use a direct estimation of eigenvalues to compute the critical values for $Q_{T,1}$ as compared to a two moment χ^2 approximation for $V_{T,1}$, and so comparing the tests in this case gives a comparison of these methods. The eigenvalue problem related to the asymptotic quantiles for the test statistic $Q_{T,1}$ was solved using (11), and the quantiles of $V_{T,K}$ were calculated using Monte Carlo integration method "rTrap" to evaluate $\hat{\eta}_{i,j}$ with $R = 25$ points as described in Section 2.4. We did not notice a significant difference when taking R larger in these examples. The empirical size with nominal levels of 10%, 5%, and 1% calculated from 1000 independent trials are reported in Table 1 for $T \in \{65, 125, 250, 500\}$, corresponding, roughly, to the number of trading days in a quarter, six months, a year, and two years.

The results can be summarized as follows:

(a) The test of $\mathcal{H}_{0,1}$ based on $Q_{T,1}$ has quite good size even when the data exhibit conditional heteroscedasticity. This statement remains true when considering lags 2, 3 and 4, which we studied in unreported simulations.
(b) When comparing these results to those for the same hypothesis test based on $V_{T,1}$, we see that the direct estimation of the eigenvalues is somewhat less conservative than the two-moment χ^2 approximation, and that this improvement is more marked in the presence of conditional heteroscedasticity.
(c) The tests for $\mathcal{H}'_{0,K}$ based on $V_{T,K}$ tended to be somewhat conservative, although the empirical size clearly approaches the nominal levels as T grows. This effect is more pronounced in the presence of conditional heteroscedasticity and large K.

To study the empirical power of the tests based on statistics $Q_{T,h}$ and $V_{T,K}$, we applied each of them to data generated according to FAR(1, S)-IID and FAR(1, S)-GARCH. The empirical rejection rates for increasing values of T are reported in Fig. 1 as power curves, where the y-axis denotes the empirical rejection rate out of 1000 independent trials with the level of each test set to $\alpha = 0.05$, and the x-axis denotes the size of the norms of the functional autoregression function, which we took to be $S = 0, 0.15, 0.30, 0.45, 0.60$ and 0.75. The empirical rejection rates were on the whole slightly higher when using $Q_{T,1}$ with the limiting quantiles estimated directly from the empirical eigenvalues in (11) compared to the χ^2 approximation approach, and so we just report the results of the χ^2 approximation below, i.e., we only tested $\mathcal{H}'_{0,1}$ using $V_{T,1}$.

Table 1
Empirical sizes based on 1000 independent simulations with nominal levels of 10%, 5%, and 1% for tests of $\mathcal{H}_{0,1}$ and $\mathcal{H}'_{0,K}$ based on $Q_{T,1}$ and $V_{T,K}$ with both IID and fGARCH(1,1) data.

Test, statistic:		$\mathcal{H}_{0,1}$, $Q_{T,1}$					
DGP:		IID			fGARCH(1,1)		
Nominal level:		10%	5%	1%	10%	5%	1%
T	65	0.100	0.050	0.006	0.095	0.035	0.004
	125	0.092	0.045	0.009	0.093	0.048	0.005
	250	0.113	0.048	0.007	0.109	0.052	0.007
	500	0.118	0.056	0.01	0.089	0.040	0.006
Test, statistic:		$\mathcal{H}'_{0,1}$, $V_{T,1}$					
		IID			fGARCH(1,1)		
		10%	5%	1%	10%	5%	1%
T	65	0.088	0.040	0.013	0.081	0.037	0.008
	125	0.096	0.066	0.017	0.097	0.040	0.014
	250	0.086	0.049	0.021	0.085	0.042	0.007
	500	0.102	0.050	0.018	0.090	0.044	0.008
Test, statistic:		$\mathcal{H}'_{0,5}$, $V_{T,5}$					
		IID			fGARCH(1,1)		
		10%	5%	1%	10%	5%	1%
T	65	0.090	0.44	0.010	0.067	0.035	0.008
	125	0.083	0.043	0.010	0.086	0.041	0.008
	250	0.078	0.038	0.008	0.077	0.049	0.014
	500	0.100	0.051	0.008	0.089	0.043	0.013
Test, statistic:		$\mathcal{H}'_{0,10}$, $V_{T,10}$					
		IID			fGARCH(1,1)		
		10%	5%	1%	10%	5%	1%
T	65	0.074	0.035	0.005	0.078	0.035	0.009
	125	0.075	0.036	0.009	0.075	0.039	0.016
	250	0.099	0.056	0.015	0.094	0.043	0.012
	500	0.094	0.044	0.012	0.102	0.051	0.009
Test, statistic:		$\mathcal{H}'_{0,20}$, $V_{T,20}$					
		IID			fGARCH(1,1)		
		10%	5%	1%	10%	5%	1%
T	65	0.103	0.066	0.022	0.091	0.051	0.019
	125	0.101	0.060	0.019	0.086	0.053	0.013
	250	0.084	0.050	0.014	0.092	0.049	0.013
	500	0.090	0.042	0.008	0.088	0.058	0.019

For small samples sizes ($T = 65$), it takes a fairly strong signal in terms of the strength of the autocorrelation in the sequence measured by the size of the autoregressive operator ($S \geq 0.5$) in order for the test to reliably reject the zero autocovariance hypothesis. However for large sample sizes, $T \geq 250$, even fairly light autocorrelation ($S \approx 0.15$) is reliably detected. Under this functional autoregressive alternative in which the autocovariance decays geometrically with the lag, the power is a decreasing function of the number of lags used in the test statistic $V_{T,K}$, as expected.

As a demonstration of the confidence bounds for ρ_h, $\widehat{B}_h(1-\alpha)$ and $\widehat{B}_{\text{i.i.d.}}(1-\alpha)$, developed in Section 2.5, we present in Fig. 2 plots of $\hat{\rho}_h$ versus $\widehat{B}_h(0.95)$ and $\widehat{B}_{\text{i.i.d.}}(0.95)$ computed from a simulated sample of length $T = 250$ following an fGARCH(1,1) process and a FAR(1, 0.75)-GARCH(1, 1) process. It can bee seen that for the functional GARCH process, the autocorrelation estimate $\hat{\rho}_h$ often exceeds the bound for a strong white noise sequence, but stays below the confidence bounds computed assuming functional conditional heteroscedasticity. This means that using the bound derived under the assumption of a strong white noise would incorrectly indicate serial autocorrelation, while using the lag dependent bounds we derived would lead to the correct conclusion that there is no serial autocorrelation. Under functional autoregression, $\hat{\rho}_h$ often goes well above both bounds, indicating that serial autocorrelation exists. In this case, both bounds lead to a correct conclusion.

3.2. Application to representative assets

We now consider an application of our methodology to intraday price data from a selection of assets representing various asset classes. Our objective in this section is to determine whether the autocovariance observed in intraday return functions constructed from these asset prices is consistent with a weak white noise. The assets that we considered are listed in Table 2.

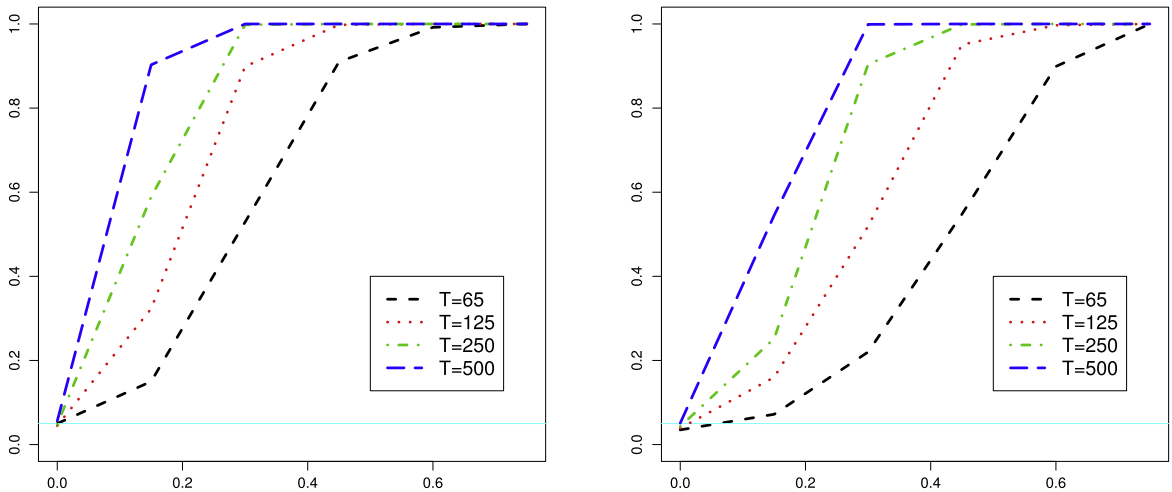

Fig. 1. Empirical rejection rates with the nominal significance level at $\alpha = 0.05$, as a function of the norm, S, of the functional autoregressive kernel; left panel: test of $\mathcal{H}'_{0,1}$, right panel: test of $\mathcal{H}'_{0,10}$. The DGP is FAR(1, S)-GARCH. The empirical rejection rates were somewhat higher for data following FAR(1, S)-IID.

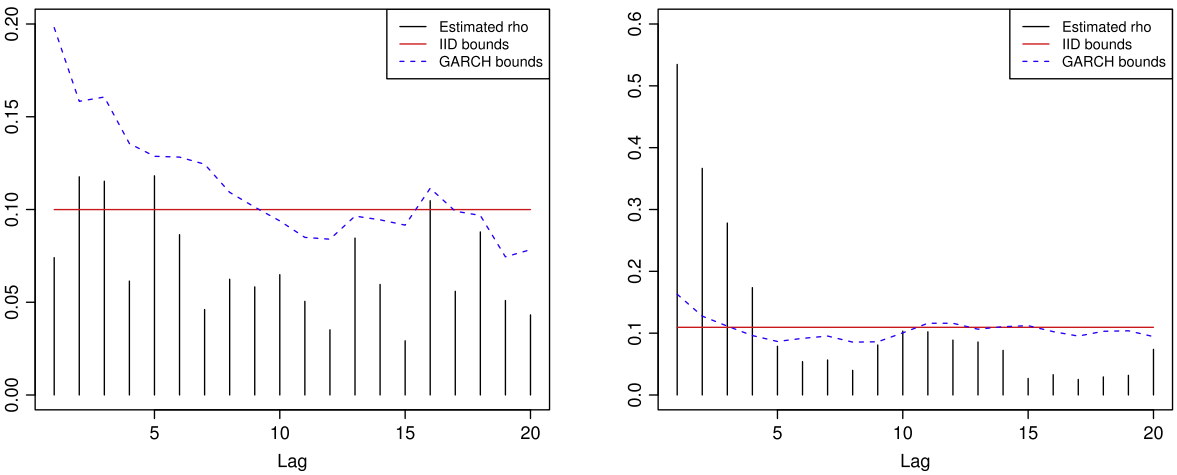

Fig. 2. Plots of $\hat{\rho}_h$ with bounds $\widehat{B}_h(0.95)$ and $\widehat{B}_{iid}(0.95)$ defined in Section 2.5 computed from a simulated sample of length $T = 250$ following a fGARCH(1,1) process (left panel) and a FAR(1, 0.75)-GARCH(1, 1) process (right panel).

The data for the S&P 500 index, Apple, Wells Fargo, and Exxon Mobile were obtained from `nasdaq.com`, and the commodity and currency exchange data were obtained from the Chicago Mercantile Exchange. We considered 5-min resolution data, $J = 78$, the time period from 2/January/2014 to 31/December/2014, which contains $T = 249$ trading days after removing a half trading day on December 24. For each asset, price data, $P_i(u)$, are available on each trading day during this time period in a 1-min resolution, which we used to construct two sequences of functions: the 5-min log-returns,

$$R_i(u) = \ln P_i(u) - \ln P_i(u-5),$$

and the cumulative intraday returns (CIDR's)

$$C_i(u) = \ln P_i(u) - \ln P_i(0).$$

While CIDR curves have not been postulated to follow a functional GARCH model, they could potentially form a functional white noise with a different dependence structure agreeing with Assumptions 1 and 2. CIDR curves are important to investors, as they show how a unit investment evolves throughout a trading day, see, e.g., [27]. Plots of the first five 5-min log return curves and cumulative intraday return curves computed from S&P 500 index are given in Fig. 3. For each asset, we plotted $\hat{\rho}_h$ and $\widehat{B}_h(0.95)$, $1 \le h \le 20$, as well as $\widehat{B}_{i.i.d.}(0.95)$ calculated from the 5-min returns and CIDR curves. These plots for 5-min log returns curves of the Apple stock price and the S&P 500 index are displayed in Fig. 4.

Table 2
Assets used in this study.

Class	Symbol	Description
Index	S&P 500	Standard & Poor 500 Index
Currency exchange	EC	Euro to Dollar
Commodity futures	CL	Crude Oil (WTI Sweet Light) Futures
Technology	AAPL	Apple Inc
Financials	WFC	Wells Fargo & Company
Energy	XOM	Exxon Mobile Corporation

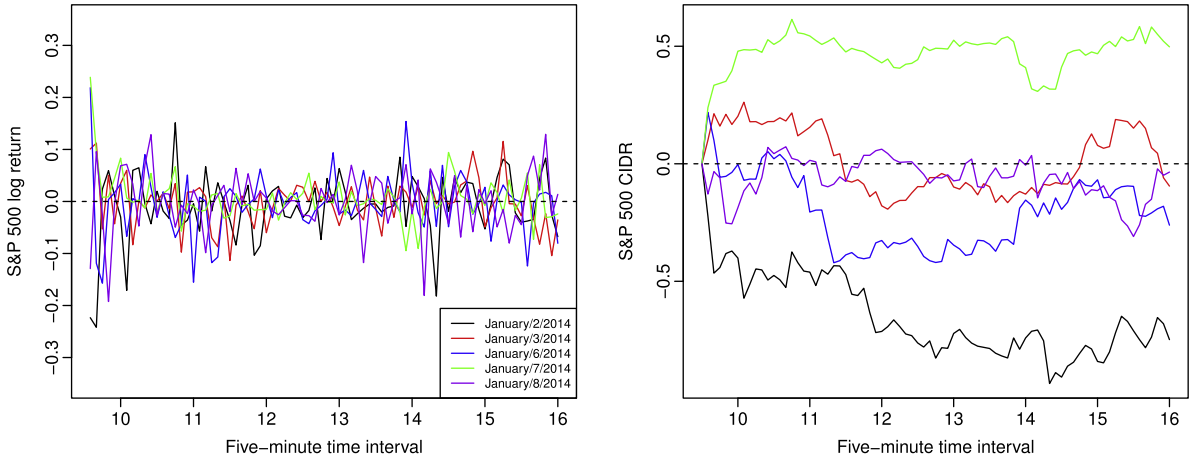

Fig. 3. Plots of the first five 5-min log return curves (left panel) and cumulative intraday return curves (right panel) constructed from the S&P 500 index.

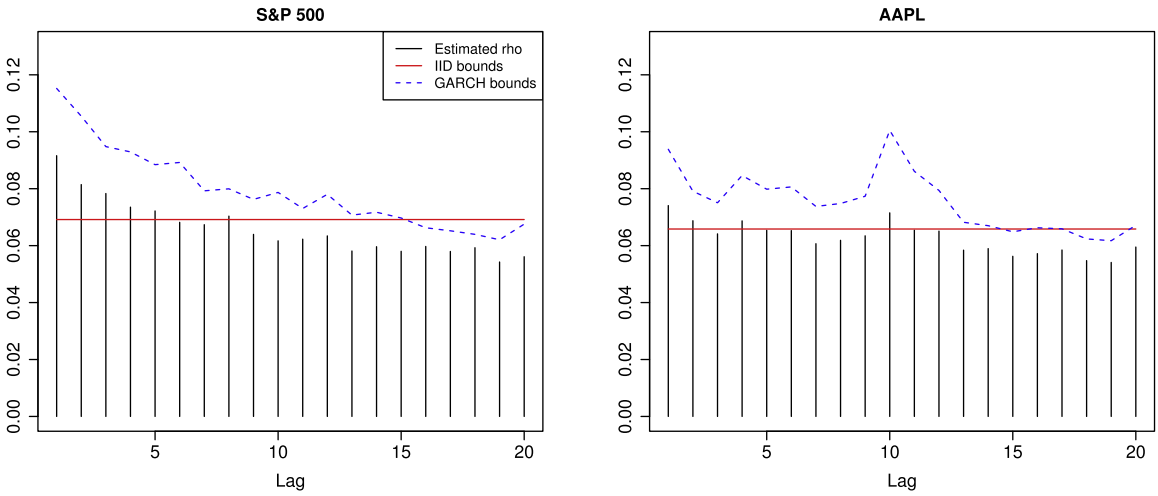

Fig. 4. Functional sample autocorrelations $\hat{\rho}_h$ defined by (17) for the 5-min log-return curves based on the S&P 500 index and Apple stock prices. Approximately 95% of the autocorrelations of a strong white noise should lie below the horizontal solid line. Approximately 95% of the autocorrelations a functional GARCH white noise should lie below the dashed curve.

One often notices in these plots that $\hat{\rho}_h$ regularly goes outside the 95% confidence interval for ρ_h assuming the sequence follows a strong white noise, but typically lies within the confidence intervals of the same level calculated under the assumption of conditional heteroscedasticity. We noticed this same pattern across all types of assets. This is in accordance with the adequacy of the functional GARCH model for the data.

Further evidence in support of the adequacy of a functional GARCH type model for these curves was supplied by testing for the cumulative significance of the first 10 functional autocovariances, a test $\mathcal{H}'_{0,10}$. We applied a test of this hypothesis based on $V_{249,10}$ to each series, and the results, in the form of p-values, are reported in Table 3. In each case, we could not reject $\mathcal{H}'_{0,10}$ with any significance. By comparison, when we applied the portmanteau test of [13] with maximum lag 10,

Table 3
p-values for portmanteau tests measuring the cumulative significance of the first 10 empirical autocovariance functions computed from the CIDR and 5-min return curves under conditional heteroscedasticity as well as under a strong white noise assumption as in [13].

p-values for portmanteau tests				
Symbol	CIDR		5-min returns	
	$\mathcal{H}'_{0,10}$	SWN test	$\mathcal{H}'_{0,10}$	SWN test
S&P 500	0.9721	0.4061	0.5170	0.0000
EC	0.5820	0.0798	0.4934	0.0001
CL	0.0788	0.0000	0.4462	0.0000
AAPL	0.5045	0.0315	0.4737	0.0989
WFC	0.1713	0.0604	0.4483	0.2193
XOM	0.3759	0.6513	0.5035	0.0332

which tests the null hypothesis that the series follows a strong white noise (abbreviated SWN Test below), the null was strongly rejected in 6 out of 12 total tests: in 4 out of 6 assets for the 5-min return curves, and 2 out of 6 assets for the CIDR curves. The SWN p-values are also given in Table 3.

Our data analysis shows that both types of return curves generally follow a weak white noise, but not a sequence of i.i.d. curves. Insights of this type have been a cornerstone of modeling of scalar returns, but to the best of our knowledge, have not been rigorously established in the context of daily return curves. We hope that our work will spur further research into the dependence structure and prediction of functional objects derived from price data, in particular towards modeling functional conditional heteroscedasticity.

Appendix. Proofs of the asymptotic results of Section 2

A.1. Proof of Theorem 1 and Corollary 1

The first lemma shows that the estimation of the mean function has an asymptotically negligible effect. Recall that we assume that $\mathrm{E}(X_i) = 0$ in the weak sense in Assumption 2. A key element of the proof is the bound $\mathrm{E}\|\sum_{i=1}^{T} X_i\|^4 = O(T^2)$ under Assumptions 1 and 2 obtained by [4,35].

Lemma 1. *Suppose Assumptions 1 and 2 hold, and let*

$$\tilde{\gamma}_h(t,s) = \frac{1}{T}\sum_{i=1}^{T-h} X_i(t)X_{i+h}(s). \tag{A.1}$$

Then

$$\left\{\iint (\hat{\gamma}_h(t,s) - \tilde{\gamma}(t,s))^2 \mu(dt)\mu(ds)\right\}^{1/2} = O_P\left(1/T\right).$$

Proof. According to the definitions of $\hat{\gamma}_h$ and $\tilde{\gamma}_h$,

$$\hat{\gamma}_h(t,s) = \tilde{\gamma}_h(t,s) - \bar{X}(s)\frac{1}{T}\sum_{i=1}^{T-h} X_i(t) - \bar{X}(t)\frac{1}{T}\sum_{i=1}^{T-h} X_i(s) + \bar{X}(t)\bar{X}(s).$$

It follows then from the triangle inequality in $L^2(\mu \otimes \mu)$ that

$$\left\{\iint (\hat{\gamma}_h(t,s) - \tilde{\gamma}(t,s))^2 \mu(dt)\mu(ds)\right\}^{1/2} \le G_1 + G_2 + G_3,$$

where

$$G_1 = \left[\iint \left\{\bar{X}(s)\frac{1}{T}\sum_{i=1}^{T-h} X_i(t)\right\}^2 \mu(dt)\mu(ds)\right]^{1/2}, \quad G_2 = \left[\iint \left\{\bar{X}(t)\frac{1}{T}\sum_{i=1}^{T-h} X_i(s)\right\}^2 \mu(dt)\mu(ds)\right]^{1/2},$$

and

$$G_3 = \left[\iint \left\{\bar{X}(t)\bar{X}(s)\right\}^2 \mu(dt)\mu(ds)\right]^{1/2}.$$

It follows from the Cauchy–Schwarz inequality that

$$\mathrm{E}(G_1^2) = \mathrm{E}\int \bar{X}^2(s)\mu(ds)\int\Big\{\frac{1}{T}\sum_{i=1}^{T-h}X_i(t)\Big\}^2\mu(dt) \le (\mathrm{E}\|\bar{X}\|^4)^{1/2}\Big(\mathrm{E}\Big\|\frac{1}{T}\sum_{i=1}^{T-h}X_i\Big\|^4\Big)^{1/2}.$$

Assumptions 1 and 2(i) imply that (X_i) is a mean zero and L^4-m-approximable series, and hence by Proposition 4 of [4], which can be adapted to the separable Hilbert space setting, see Proposition 2.61 of [35], we have that

$$\mathrm{E}\Big\|\sum_{i=1}^{T}X_i\Big\|^4 = O(T^2).$$

We therefore conclude that $\mathrm{E}\|\bar{X}\|^4 = O(1/T^2)$. It follows similarly that

$$\mathrm{E}\Big\|\frac{1}{T}\sum_{i=1}^{T-h}X_i\Big\|^4 = O\left(1/T^2\right),$$

and so $\mathrm{E}(G_1^2) = O(1/T^2)$. This along with Chebyshev's inequality implies that $G_1 = O_P\left(1/T\right)$. Similar arguments show that $G_i = O_P(1/T)$ for $i \in \{2, 3\}$, which implies the result. □

Our objective is to establish limit distribution of the sample autocovariance function for a fixed lag h. The first step is to show that the products $X_{i-h}(t)X_i(s)$ form an L^2-m-approximable sequence. Under Assumption 1, $\tilde{\gamma}_h(t,s)$ has the same distribution as

$$\bar{\gamma}_h(t,s) = \frac{1}{T}\sum_{j=1}^{T-h}Z_{i,h}(t,s), \tag{A.2}$$

where $Z_{i,h}(t,s) = X_{i-h}(t)X_i(s)$. This time shift is introduced to obtain a more convenient Bernoulli shift representation. The sequence $(Z_{i,h})$ is evidently strictly stationary, and has mean zero under Assumption 2(ii). It follows from the Cauchy–Schwarz inequality for the expectation and Assumption 1 that

$$\mathrm{E}\Big\{\iint Z_{i,h}^2(t,s)\mu(dt)\mu(ds)\Big\}^{1/2} < \infty,$$

and so $Z_{i,h}$ is a.s. an element of $L^2(\mu\otimes\mu)$. Moreover, according to Assumption 1,

$$Z_{i,h} = f_h(\varepsilon_i, \varepsilon_{i-1}, \ldots),$$

where $f_h : S^\infty \to L^2(\mu\otimes\mu)$. Let $Z_{i,h}^{(m)}$ be defined as in (6). We note that for $m > h$, $Z_{i,h}^{(m)} = X_{i-h}^{(m-h)} \otimes X_i^{(m)}$.

Lemma 2. *Under Assumption 1, the sequence $(Z_{i,h})$ is L^2-m-approximable, i.e.,*

$$\sum_{m=1}^{\infty}\Big[\mathrm{E}\iint\{Z_{i,h}(t,s) - Z_{i,h}^{(m)}(t,s)\}^2\mu(dt)\mu(ds)\Big]^{1/2} < \infty.$$

Proof. By adding and subtracting the same term under the square in the integrand, we get that

$$\begin{aligned}&\mathrm{E}\iint\{Z_{i,h}(t,s) - Z_{i,h}^{(m)}(t,s)\}^2\mu(dt)\mu(ds)\\ &\quad= \mathrm{E}\iint\{X_{i-h}(t)X_i(s) - X_{i-h}(t)X_i^{(m)}(s) + X_{i-h}(t)X_i^{(m)}(s) - X_{i-h}^{(m-h)}(t)X_i^{(m)}(s)\}^2\mu(dt)\mu(ds)\end{aligned}$$

and hence using the triangle inequalities in $L^2(\mu\otimes\mu)$ and for $\{\mathrm{E}(\cdot)^2\}^{1/2}$ it follows that

$$\begin{aligned}\Big[\mathrm{E}\iint\{Z_{i,h}(t,s) - Z_{i,h}^{(m)}(t,s)\}^2\mu(dt)\mu(ds)\Big]^{1/2} &\le \Bigg\{\mathrm{E}\Bigg(\Big[\iint\{X_{i-h}(t)X_i(s) - X_{i-h}(t)X_i^{(m)}(s)\}^2\mu(dt)\mu(ds)\Big]^{1/2}\\ &\qquad+\Big[\iint\{X_{i-h}(t)X_i^{(m)}(s) - X_{i-h}^{(m-h)}(t)X_i^{(m)}(s)\}^2\mu(dt)\mu(ds)\Big]^{1/2}\Bigg)^2\Bigg\}^{1/2}\\ &\le \Big[\mathrm{E}\iint\{X_{i-h}(t)X_i(s) - X_{i-h}(t)X_i^{(m)}(s)\}^2\mu(dt)\mu(ds)\Big]^{1/2}\\ &\qquad+\Big[\mathrm{E}\iint\{X_{i-h}(t)X_i^{(m)}(s) - X_{i-h}^{(m-h)}(t)X_i^{(m)}(s)\}^2\mu(dt)\mu(ds)\Big]^{1/2}.\end{aligned} \tag{A.3}$$

We obtain by some simple algebra and the Cauchy–Schwarz inequality that the second term in the last line of (A.3) satisfies

$$\left[\mathrm{E}\iint\{X_{i-h}(t)X_i^{(m)}(s)-X_{i-h}^{(m-h)}(t)X_i^{(m)}(s)\}^2\mu(dt)\mu(ds)\right]^{1/2} = (\mathrm{E}\|X_i^{(m)}\|^2\|X_{i-h}-X_{i-h}^{(m-h)}\|^2)^{1/2}$$
$$\leq (\mathrm{E}\|X_i\|^4)^{1/4}(\mathrm{E}\|X_{i-h}-X_{i-h}^{(m-h)}\|^4)^{1/4}. \tag{A.4}$$

We recall that by Assumption 1, $\mathrm{E}(\|X_i\|^4)<\infty$. Similar arguments show that the first term on the last line of (A.3) satisfies

$$\left[\mathrm{E}\iint\{X_{i-h}(t)X_i(s)-X_{i-h}(t)X_i^{(m)}(s)\}^2\mu(dt)\mu(ds)\right]^{1/2} \leq (\mathrm{E}\|X_{i-h}\|^4)^{1/4}(\mathrm{E}\|X_i-X_i^{(m)}\|^4)^{1/4},$$

from which we obtain, when combined with (A.4), that

$$\sum_{m=h}^{\infty}\left[\mathrm{E}\iint\{Z_{i,h}(t,s)-Z_{i,h}^{(m)}(t,s)\}^2\mu(dt)\mu(ds)\right]^{1/2}$$
$$\leq (\mathrm{E}\|X_{i-h}\|^4)^{1/4}\left[\sum_{m=h}^{\infty}(\mathrm{E}\|X_{i-h}-X_{i-h}^{(m-h)}\|^4)^{1/4}+\sum_{m=h}^{\infty}(\mathrm{E}\|X_i-X_i^{(m)}\|^4)^{1/4}\right]<\infty.$$

This completes the proof. □

Lemma 3. *Suppose Assumptions 1 and 2 hold. Then one may define a sequence of Gaussian processes $\Gamma_{h,T}$, $T\geq 1$, in $L^2(\mu\otimes\mu)$ on the same, perhaps enlarged, probability space as (X_i) such that*

$$\iint\{\sqrt{T}\bar{\gamma}_h(t,s)-\Gamma_{h,T}(t,s)\}^2\mu(ds)\mu(dt)=o_P(1),$$

where for t,s,t',s' in the support of μ, $\mathrm{E}\Gamma_{h,T}(t,s)=0$, and

$$c_h(t,s,t',s')=\mathrm{E}\left\{\Gamma_{h,T}(t,s)\Gamma_{h,T}(t',s')\right\}=\mathrm{E}\left\{X_0(t)X_0(t')X_h(s)X_h(s')\right\}.$$

Proof. It follows from similar arguments as those used to establish Lemma 1 that if

$$\gamma_h^\star(t,s)=\frac{1}{T}\sum_{j=1}^{T}Z_{i,h}(t,s),$$

then

$$\left[\iint\{\sqrt{T}\bar{\gamma}_h(t,s)-\sqrt{T}\gamma_h^\star(t,s)\}^2\mu(dt)\mu(ds)\right]^{1/2}=O_P(h/\sqrt{T})=o_P(1).$$

It follows from Lemma 2 that the sequence $(Z_{i,h})$ is L^2-m-approximable, and hence we obtain from Theorem 1.2 of [19] that a sequence of Gaussian processes $\Gamma_{h,T}(t,s)$ as defined in the statement of the lemma exists and satisfies

$$\iint\{\sqrt{T}\gamma_h^\star(t,s)-\Gamma_{h,T}(t,s)\}^2\mu(ds)\mu(dt)=o_P(1),$$

with

$$\mathrm{E}\Gamma_{h,T}(t,s)\Gamma_{h,T}(t',s')=\mathrm{E}Z_{0,h}(t,s)Z_{0,h}(t',s')+\sum_{j=1}^{\infty}\mathrm{E}Z_{0,h}(t,s)Z_{j,h}(t',s')+\sum_{j=1}^{\infty}\mathrm{E}Z_{0,h}(t',s')Z_{j,h}(t,s).$$

According to Assumption 2(iii), $\mathrm{E}Z_{0,h}(t',s')Z_{j,h}(t,s)=0$ and $\mathrm{E}Z_{0,h}(t,s)Z_{j,h}(t',s')=0$ for all $j\geq 1$, which implies that

$$\mathrm{E}\Gamma_{h,T}(t,s)\Gamma_{h,T}(t',s')=\mathrm{E}Z_{0,h}(t,s)Z_{0,h}(t',s')=c_h(t,s,t',s').$$

This completes the argument. □

Proof of Theorem 1. Theorem 1 follows directly from Lemmas 1 and 3. □

Proof of Corollary 1. According the triangle inequality in $L^2(\mu \otimes \mu)$ and Theorem 1,

$$\left| \left\{ \iint T\hat{\gamma}_h^2(t,s)\mu(dt)\mu(ds) \right\}^{1/2} - \left\{ \iint \Gamma_{h,T}^2(t,s)\mu(dt)\mu(ds) \right\}^{1/2} \right|$$
$$\leq \left[\iint \{\sqrt{T}\hat{\gamma}_h(t,s) - \Gamma_{h,T}(t,s)\}^2 \mu(dt)\mu(ds) \right]^{1/2} = o_P(1).$$

This implies

$$\iint T\hat{\gamma}_h^2(t,s)\mu(dt)\mu(ds) = \iint \Gamma_{h,T}^2(t,s)\mu(dt)\mu(ds) + o_P(1).$$

Finally, since the elements of the approximating Gaussian sequence each have the same distribution, we have that

$$\iint \Gamma_{h,T}^2(t,s)\mu(dt)\mu(ds) \stackrel{D}{=} \iint \Gamma_{h,0}^2(t,s)\mu(dt)\mu(ds)$$

for every integer $T \geq 0$. This completes the proof. □

A.2. Proof of Theorem 2 and (15)

To begin, let K be a positive integer as in the definition of $\mathcal{H}'_{0,K}$, and define

$$\tilde{V}_{T,K} = T \sum_{h=1}^{K} \|\tilde{\gamma}_h\|^2$$

where $\tilde{\gamma}_h$ is defined in (A.1).

Lemma 4. *Under Assumptions 1 and 2, $|V_{T,K} - \tilde{V}_{T,K}| = o_P(1)$.*

Proof. The proof follows along similar lines as Lemma 1, and so the details are omitted. □

Analogous to the proof of Theorem 1, we also define

$$\bar{V}_{T,K} = T \sum_{h=1}^{K} \|\bar{\gamma}_h\|^2,$$

where $\bar{\gamma}$ is defined in (A.2). Under Assumption 1, $\tilde{V}_{T,K}$ and $\bar{V}_{T,K}$ have the same distribution.

Let K be a positive integer as in the definition of $\mathcal{H}'_{0,K}$. In order to prove Theorem 2, we must introduce some notation related to two Hilbert spaces of square integrable, finite-dimensional, functions. Consider the space $\mathcal{H}_1$ of functions $f : [0,1]^2 \to \mathbb{R}^K$, mapping the unit square to the space of K-dimensional column vectors with real entries, satisfying

$$\iint \{f(t,s)\}^\top f(t,s)\mu(dt)\mu(ds) < \infty.$$

This space is a separable Hilbert space when equipped with the inner product

$$\langle f, g\rangle_{\mathcal{H},1} = \iint \{f(t,s)\}^\top g(t,s)\mu(dt)\mu(ds).$$

Let $\|\cdot\|_{\mathcal{H},1}$ denote the norm induced by this innerproduct. Let $\langle\cdot,\cdot\rangle_{\mathbb{F}}$ denote the matrix Frobenius inner product, and let $\|\cdot\|_{\mathbb{F}}$ denote the corresponding norm; see Chapter 5 of [29]. Further let $\mathcal{H}_2$ denote the space of functions $f : [0,1]^4 \to \mathbb{R}^{K\times K}$, equipped with the inner product

$$\langle f, g\rangle_{\mathcal{H},2} = \iiiint \langle f(t,s,u,v), g(t,s,u,v)\rangle_{\mathbb{F}}\mu(dt)\mu(ds)\mu(du)\mu(dv).$$

for which $\langle f, f\rangle_{\mathcal{H},2} < \infty$. $\mathcal{H}_2$ is also a separable Hilbert space when equipped with this inner product.

Now, proceeding with this notation towards establishing Theorem 2, let $\psi_K : [0,1]^4 \to \mathbb{R}^{K\times K}$ be defined by

$$\psi_K(t,s,u,v) = \{\mathrm{E}\{X_{-i}(t)X_0(s)X_{-j}(u)X_0(v)\}, 1 \leq i,j \leq K\}.$$

The kernel ψ_K defines a linear operator $\Psi_K : \mathcal{H}_1 \to \mathcal{H}_1$ by

$$\Psi_K(f)(t,s) = \iint \psi_K(t,s,u,v)f(u,v)\mu(du)\mu(dv), \tag{A.5}$$

where the integration is carried out coordinate-wise. The operator Ψ_K has the following three properties:

1. Ψ_K is Hilbert–Schmidt, and therefore is compact. This follows if we show that the kernel ψ_K is an element of $\mathcal{H}_2$, see Chapter 6 of [31]. We have by Assumption 1 and two applications of the Cauchy–Schwarz inequality that

$$\begin{aligned}\langle \psi_K, \psi_K\rangle_{\mathcal{H},2} &= \iiiint \|\psi_K(t,s,u,v)\|_{\mathbb{F}}^2 \mu(dt)\mu(ds)\mu(du)\mu(dv)\\ &= \iiiint \sum_{i,j=1}^{K}[\mathrm{E}\{X_{-i}(t)X_0(s)X_{-j}(u)X_0(v)\}]^2\mu(dt)\mu(ds)\mu(du)\mu(dv)\\ &\le \sum_{i,j=1}^{K}\mathrm{E}\left\{\int X_{-i}^2(t)dt\int X_0^2(s)ds\right\}\mathrm{E}\left\{\int X_{-j}^2(u)du\int X_0^2(v)dv\right\}\\ &= \sum_{i,j=1}^{K}\mathrm{E}(\|X_{-i}\|^2\|X_0\|^2)\mathrm{E}(\|X_{-j}\|^2\|X_0\|^2)\\ &\le \sum_{i,j=1}^{K}(\mathrm{E}(\|X_0\|^4)^2) = K^2\{\mathrm{E}(\|X_0\|^4)^2\} < \infty.\end{aligned}$$

2. Ψ_K is symmetric. We have, by the definition of ψ_K, that $\psi_K^\top(t,s,u,v) = \psi_K(u,v,t,s)$, and therefore for all $f, g \in \mathcal{H}_1$,

$$\begin{aligned}\langle \Psi_K(f), g\rangle &= \iiiint f^\top(u,v)\psi_K^\top(t,s,u,v)g(t,s)\mu(dt)\mu(ds)\mu(du)\mu(dv)\\ &= \iiiint f^\top(u,v)\psi_K(u,v,t,s)g(t,s)\mu(dt)\mu(ds)\mu(du)\mu(dv) = \langle f, \Psi_K(g)\rangle.\end{aligned}$$

3. Ψ_K is positive definite. To see this, let $\zeta_K(t,s) = (X_0(t)X_1(s), \ldots, X_0(t)X_K(s))^\top$. Then it follows by Fubini's theorem that for any $f \in \mathcal{H}_1$,

$$\langle f, \Psi_K(f)\rangle = \mathrm{E}\left\{\iint f^\top(t,s)\zeta_K(t,s)\mu(dt)\mu(ds)\right\}^2 \ge 0.$$

Due to these three properties, we have by the spectral theorem for positive definite, self-adjoint, compact operators, see Chapter 6.2 of [31], that Ψ_K defines a nonnegative and decreasing sequence of eigenvalues, $\xi_{1,K} \ge \xi_{2,K} \ge \cdots$ and corresponding orthonormal basis of eigenfunctions $\varphi_{i,K}(t,s)$, $1 \le i < \infty$, satisfying

$$\Psi_K(\varphi_i)(t,s) = \xi_{i,K}\varphi_{i,K}(t,s), \quad \text{with } \sum_{i=1}^{\infty}\xi_{i,K} < \infty. \tag{A.6}$$

Let

$$\hat{\Gamma}_{T,K}(t,s) = \left\{\frac{1}{\sqrt{T}}\sum_{i=1}^{T}X_{i-1}(t)X_i(s), \ldots, \frac{1}{\sqrt{T}}\sum_{i=1}^{T}X_{i-K}(t)X_i(s)\right\}^\top.$$

Lemma 5. *Under the conditions of Theorem 2, $\hat{\Gamma}_{T,K}(t,s) \xrightarrow{D(\mathcal{H}_1)} \Gamma_K(t,s)$, where $\Gamma_K(t,s)$ is a Gaussian process in $\mathcal{H}_1$ with mean zero and covariance operator Ψ_K defined in (A.5).*

Proof. Let $v(t,s) = (v_1(t,s), \ldots, v_K(t,s))^\top \in \mathcal{H}_1$. Then

$$\langle \hat{\Gamma}_{T,K}, v\rangle_{\mathcal{H},1} = \iint \Gamma_{K,T}^\top(t,s)v(t,s)\mu(dt)\mu(ds) = \frac{1}{\sqrt{T}}\sum_{i=1}^{T}\sum_{j=1}^{K}\iint X_{i-j}(t)X_i(s)v_j(t,s)\mu(dt)\mu(ds) = \frac{1}{\sqrt{T}}\sum_{i=1}^{T}\theta_{i,K}(v).$$

Along similar lines as the proof of Lemma 2, one can show that the stationary real-valued sequence $\{\theta_{i,K}(v)\}_{i\in\mathbb{Z}}$ is L^2-m-approximable. Moreover, Fubini's theorem and Assumption 2 imply that $\mathrm{E}\{\theta_{i,K}(v)\} = 0$. Also, for integers $\ell \neq 0$ we have according to Fubini's theorem and Assumption 2(iii) that

$$\mathrm{E}\{\theta_{0,K}(v)\theta_{\ell,K}(v)\} = \sum_{i=1}^{K}\sum_{j=1}^{K}\iiiint \mathrm{E}\{X_{-j}(t)X_0(s)X_{\ell-j}(u)X_\ell(r)\}v_i(t,s)v_j(u,r)\mu(dt)\mu(ds)\mu(du)\mu(dr) = 0,$$

and for $\ell = 0$,

$$\mathrm{E}\{\theta_{0,K}(v)\theta_{0,K}(v)\} = \sum_{i=1}^{K}\sum_{j=1}^{K}\iiiint \mathrm{E}\{X_{-j}(t)X_0(s)X_{-j}(u)X_0(r)\}v_i(t,s)v_j(u,r)\mu(dt)\mu(ds)\mu(du)\mu(dr) = \langle \Psi_K(v), v\rangle_{\mathcal{H},1},$$

from which it follows that

$$\sum_{i=-\infty}^{\infty} \mathrm{E}\{\theta_{0,K}(v)\theta_{i,K}(v)\} = \langle \Psi_K(v), v\rangle_{\mathcal{H},1}.$$

Therefore, by Theorem 3 of [38],

$$\langle \hat{\Gamma}_{T,K}, v\rangle_{\mathcal{H},1} \rightsquigarrow \mathcal{N}(0, \langle \Psi_K(v), v\rangle_{\mathcal{H},1}) \stackrel{D}{=} \langle \Gamma_K, v\rangle_{\mathcal{H},1},$$

where $\mathcal{N}(m, \sigma^2)$ denotes a normal random variable with mean m and variance σ^2. This shows that the finite-dimensional distributions of $\hat{\Gamma}_{T,K}$ converge to those of Γ_K. We now aim to show that the sequence $\hat{\Gamma}_{T,K}$ is tight in $\mathcal{H}_1$. We begin by showing that the sequence $\mathbf{X}_{i,K}(t,s) = (X_{i-1}(t)X_i(s), \ldots, X_{i-K}(t)X_i(s))^\top$ is L^2-m-approximable in $\mathcal{H}_1$. Evidently, by Assumption 1, $\mathbf{X}_{i,K}$ is of the form $\mathbf{X}_{i,K} = f_K(\varepsilon_i, \varepsilon_{i-1}, \ldots)$, where $f : S^\infty \to \mathcal{H}_1$. Let $\mathbf{X}_{i,K}^{(m)}$ be defined as in (6) so that when $m > K$, $\mathbf{X}_{i,K}^{(m)}(t,s) = (X_{i-1}^{(m-1)}(t)X_i^{(m)}(s), \ldots, X_{i-K}^{(m-K)}(t)X_i^{(m)}(s))^\top$. We then have that for $m > K$ by (A.3) and (A.4), that

$$\begin{aligned}
(\mathrm{E}\|\mathbf{X}_{i,K} - \mathbf{X}_{i,K}^{(m)}\|_{\mathcal{H},1}^2)^{1/2} &= \left[\sum_{j=1}^{K} \mathrm{E}\iint \{X_{i-j}(t)X_i(s) - X_{i-j}^{(m-j)}(t)X_i^{(m)}(s)\}^2 \mu(dt)\mu(ds)\right]^{1/2} \\
&\le \sum_{j=1}^{K} [\mathrm{E}\iint \{X_{i-j}(t)X_i(s) - X_{i-j}^{(m-j)}(t)X_i^{(m)}(s)\}^2 \mu(dt)\mu(ds)]^{1/2} \\
&\le \sum_{j=1}^{K} (\mathrm{E}\|X_0\|^4)^{1/4}\{(\mathrm{E}\|X_0 - X_0^{(m-j)}\|^4)^{1/4} + (\mathrm{E}\|X_0 - X_0^{(m-j)}\|^4)^{1/4}\}.
\end{aligned}$$

Therefore by Assumption 1 and since K is fixed,

$$\sum_{m=K+1}^{\infty} (\mathrm{E}\|\mathbf{X}_{i,K} - \mathbf{X}_{i,K}^{(m)}\|_{\mathcal{H},1}^2)^{1/2} \le \sum_{j=1}^{K} (\mathrm{E}\|X_0\|^4)^{1/4} \sum_{m=K+1}^{\infty} \{(\mathrm{E}\|X_0 - X_0^{(m-j)}\|^4)^{1/4} + (\mathrm{E}\|X_0 - X_0^{(m-j)}\|^4)^{1/4}\} < \infty,$$

showing that $\mathbf{X}_{i,K}(t,s)$ is L^2-m-approximable in $\mathcal{H}_1$. Since $\hat{\Gamma}_{T,K} = \sum_{i=1}^{T} \mathbf{X}_{i,K}/\sqrt{T}$, it follows from Lemmas 5, 6, and 7 in [8] that the sequence $\hat{\Gamma}_{T,K}$ is tight in $\mathcal{H}_1$. The result now follows from Theorem 2.2 of [7]. □

Proof of Theorem 2. This is now a simple consequence of Lemma 5. It follows from the latter, the continuous mapping theorem, and the Karhunen–Loève theorem that

$$\bar{V}_{T,K} = \|\hat{\Gamma}_{T,K}\|_{\mathcal{H},1}^2 \rightsquigarrow \|\Gamma_K\|_{\mathcal{H},1}^2 \stackrel{D}{=} \sum_{\ell=1}^{\infty} \xi_{\ell,K}\mathcal{N}_\ell^2,$$

where $(\xi_{\ell,K})_{\ell=1}^{\infty}$ are the eigenvalues of Ψ_K. Now the result for $V_{T,K}$ follows from Lemma 4. □

Justification of (15). According to Proposition 5.10.16 of [5] (see Theorems 1.5 and 1.6 in [33] in the finite-dimensional case),

$$\mathrm{E}(\|\Gamma_K\|_{\mathcal{H},1}^2) = \mathrm{tr}(\Psi_K) = \iint \mathrm{E}X_0^2(t)X_i^2(s)\mu(dt)\mu(ds),$$

and

$$\mathrm{var}(\|\Gamma_K\|\|_{\mathcal{H},1}^2) = 2\,\mathrm{tr}(\Psi_K^2) = 2\sum_{1\le i,j\le K} \iiiint [\mathrm{E}\{X_{-i}(t)X_0(s)X_{-j}(u)X_0(v)\}]^2 \mu(dt)\mu(ds)\mu(du)\mu(dv).$$

A.3. Proof of Theorem 3

Proof of Theorem 3. Since $V_{T,K} \ge Q_{T,h}$ for all $h \in \{1, \ldots, K\}$, it is enough to just prove the first part of the theorem. By repeating the arguments in Lemma 1, it follows that it is enough to show that under the conditions of the theorem, $T\|\tilde{\gamma}_h\|^2 \stackrel{P}{\to} \infty$, as $T \to \infty$. By Assumption 3, it follows that there exists a nonzero function $a_h \in L^2(\mu \otimes \mu)$ such that

$\mathrm{E}\{X_0(t)X_h(s)\} = a_h(t,s)\ \mu\otimes\mu$ almost everywhere. From this we obtain that

$$T\|\tilde{\gamma}_h\|^2 = \frac{1}{T}\iint\left[\sum_{j=1}^{T-h}\{X_j(t)X_{j+h}(s) - a_h(t,s) + a_h(t,s)\}\right]^2 dtds \tag{A.7}$$

$$= \iint\left[\frac{1}{\sqrt{T}}\sum_{j=1}^{T-h}\{X_j(t)X_{j+h}(s) - a_h(t,s)\}\right]^2 dtds$$

$$+\ 2\iint\left[\frac{T-h}{T}\,a_h(t,s)\sum_{j=1}^{T-h}\{X_j(t)X_{j+h}(s) - a_h(t,s)\}\right]dtds + \frac{(T-h)^2}{T}\,\|a_h\|^2.$$

It follows again from similar arguments as those used to establish Lemma 2 that the stationary sequence $X_j(t)X_{j+h}(s) - a_h(t,s) \in L^2(\mu\otimes\mu), j\in\mathbb{Z}$ is mean zero and L^2-m-approximable, from which we obtain that

$$\iint\left[\frac{1}{\sqrt{T}}\sum_{j=1}^{T-h}\{X_j(t)X_{j+h}(s) - a_h(t,s)\}\right]^2 dtds = O_P(1),$$

and

$$\iint\left[\frac{T-h}{T}\,a_h(t,s)\sum_{j=1}^{T-h}\{X_j(t)X_{j+h}(s) - a_h(t,s)\}\right]dtds = O_P(\sqrt{T}).$$

The result then follows in light of (A.7) since $(T-h)^2/T\|a_h\|^2$ diverges to infinity at rate T. □

References

[1] A. Aue, S. Hörmann, L. Horváth, M. Reimherr, Break detection in the covariance structure of multivariate time series models, Ann. Statist. 37 (2009) 4046–4087.
[2] A. Aue, L. Horváth, D.F. Pellatt, Functional generalized autoregressive conditional heteroskedasticity, J. Time Ser. Anal. 38 (2017) 3–21.
[3] L. Bauwens, S. Laurent, J. Rombouts, Multivariate GARCH models: A survey, J. Appl. Econometrics 21 (2006) 79–109.
[4] I. Berkes, S. Hörmann, J. Schauer, Split invariance principles for stationary processes, Ann. Probab. 39 (2011) 2441–2473.
[5] V.I. Bogachev, Gaussian Measures, American Mathematical Society, 1998.
[6] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, J. Econometrics 31 (1986) 307–327.
[7] D. Bosq, Linear Processes in Function Spaces, Springer, New York, 2000.
[8] C. Cerovecki, S. Hörmann, On the CLT for discrete Fourier transforms of functional time series, J. Multivariate Anal. 154 (2017) 282–295.
[9] P. Duchesne, Testing for multivariate autoregressive conditional heteroskedasticity using wavelets, Comput. Statist. Data Anal. 51 (2006) 2142–2163.
[10] R.F. Engle, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50 (1982) 987–1007.
[11] C. Francq, R. Roy, J.-M. Zakoian, Diagnostic checking in ARMA models with uncorrelated errors, J. Amer. Statist. Assoc. 100 (2005) 532–544.
[12] C. Francq, J.-M. Zakoian, GARCH Models, Wiley, New York, 2010.
[13] R. Gabrys, P. Kokoszka, Portmanteau test of independence for functional observations, J. Amer. Statist. Assoc. 102 (2007) 1338–1348.
[14] C. Gouriéroux, ARCH Models and Financial Applications, Springer, New York, 1997.
[15] S. Hörmann, L. Horváth, R. Reeder, A functional version of the ARCH model, Econom. Theory 29 (2013) 267–288.
[16] S. Hörmann, P. Kokoszka, Weakly dependent functional data, Ann. Statist. 38 (2010) 1845–1884.
[17] L. Horváth, P. Kokoszka, Inference for Functional Data with Applications, Springer, 2012.
[18] L. Horváth, G. Rice, S. Whipple, Adaptive bandwidth selection in the estimation of the long run covariance of functional time series, Comput. Statist. Data Anal. 100 (2016) 676–693.
[19] M. Jirak, On weak invariance principles for sums of dependent random functionals, Statist. Probab. Lett. 83 (2013) 2291–2296.
[20] V. Kargin, A. Onatski, Curve forecasting by functional autoregression, J. Multivariate Anal. 99 (2008) 2508–2526.
[21] J. Klepsch, C. Klüppelberg, An innovations algorithm for the prediction of functional linear processes, J. Multivariate Anal. 155 (2017) 252–271.
[22] P. Kokoszka, H. Miao, X. Zhang, Functional dynamic factor model for intraday price curves, J. Financ. Econ. 13 (2015) 456–477.
[23] P. Kokoszka, M. Reimherr, Predictability of shapes of intra day price curves, Econom. J. 16 (2013) 285–308.
[24] P. Kokoszka, M. Reimherr, Introduction To Functional Data Analysis, CRC Press, Boca Raton, FL, 2017.
[25] K. Krishnamoorthy, Modified normal-based approximation to the percentiles of linear combination of independent random variables with applications, Comm. Statist. Simulation Comput. 45 (2016) 2428–2444.
[26] W.K. Li, Diagnostic Checks in Time Series, Chapman & Hall, London, 2004.
[27] D.O. Lucca, E. Moench, The pre-FOMC announcement drift, J. Financ. 70 (2015) 329–371.
[28] E. Mahdi, I. McLeod, Improved multivariate portmanteau test, J. Time Ser. Anal. 33 (2012) 211–222.
[29] C.D. Meyer (Ed.), Matrix Analysis and Applied Linear Algebra, SIAM, 2000.
[30] V.M. Panaretos, S. Tavakoli, Fourier analysis of stationary time series in function space, Ann. Statist. 41 (2013) 568–603.
[31] F. Riesz, B.Sz. -Nagy, Functional Analysis, Dover, 1990.
[32] J. Romano, L. Thombs, Inference for autocorrelations under weak assumptions, J. Amer. Statist. Assoc. 91 (1996) 590–600.
[33] G.A.F. Seber, A.J. Lee, Linear Regression Analysis, Wiley, New York, 2003.
[34] X. Shao, W.B. Wu, Asymptotic spectral theory for nonlinear time series, Ann. Statist. 35 (2007) 1773–1801.
[35] L. Torgovitski, Hilbert space valued signal plus noise models: Analysis of structural breaks under high dimensionality and temporal dependence, Ph.D. thesis, Universität zu Köln, Germany, 2016.
[36] Y.K. Tse, Residual-based diagnostics for conditional heteroscedasticity models, Econom. J. 5 (2002) 358–373.

[37] K.G. van den Boogaart, tensorA: Advanced tensors arithmetic with named indices, 2010. URL https://CRAN.R-project.org/package=tensorA. R package version 0.36.
[38] W. Wu, Nonlinear system theory: another look at dependence, in: Proceedings of the National Academy of Sciences of the United States, Vol. 102, National Academy of Sciences, 2005.
[39] J.T. Zhang, Analysis of Variance for Functional Data, Chapman & Hall, London, 2013.
[40] X. Zhang, White noise testing and model diagnostic checking for functional time series, J. Econometrics 194 (2016) 76–95.
[41] X. Zhang, X. Shao, Two sample inference for the second-order property of temporally dependent functional data, Bernoulli 21 (2015) 909–929.