Теорія ймовір. та матем. статист.
Вип. 95, 2016, стор. 55–74

Teor. Ĭmovir. ta Matem. Statyst.
No. 95, 2016, pp. 55–74

SPECTRAL ESTIMATION IN THE PRESENCE OF MISSING DATA

UDC 519.21

NATALIA BAHAMONDE AND PAUL DOUKHAN

Abstract. In this article we propose a quasi-Whittle estimator for parametric families of time series models in the presence of missing data. This estimator extends results to the incompletely observed case. This extension is valid to non-Gaussian and non-linear models. It also allows to bound the variance of an associated quasi-periodogramm. A simulation study validates empirically the proposed estimate for mixing and non-mixing models.

1. Introduction

A common problem in parametric time series analysis is the estimation of a finite variance stationary ARMA model in the presence of missing observations. Missing data appear either by nature, or because aberrant data were detected and thus eliminated. A large number of such data appear in the literature, including hydrology and environmental sciences.

In most of the literature, the asymptotic properties of estimators in time series with missing observations is concerned with linear processes with normal innovations. Some extensions to linear processes also exclude heteroscedasticity which is a standard feature of financial time series.

We first quote that stationarity does essentially not hold for real data: however the spectrum takes its signification for 2nd order stationary processes and we shall restrict to the stationary frame. For locally stationary processes [4] introduced an analogue of the Whittle estimate which will be addressed in a further paper in presence of missing observations; this was processed in [18] for the case of independently distributed missing data who accurately achieved by a localization through a wavelet transform. Anyway the latter reference is more methodological in § 5.1 this is indeed mentioned that results are not rigorously stated: asymptotic limit theory remains thus to be processed.

In order to omit independence many proposals have been suggested as strong mixing properties in [22]. Weak dependence in [9] includes wider classes of models used in econometrics and in statistics. Financial time series often exhibit heteroscedasticity e.g. ARCH type processes introduced in [16] are commonly used in the financial econometrics literature, because of their stylistic properties.

The Whittle approximation of the likelihood yields suitable estimates for parametric models such as ARMA, LARCH, GARCH or Bilinear models. The Whittle contrast does not depend on marginal distributions but only on spectral densities. It has quicker to calculate that other parametric estimates such as MLE; moreover the latter based on the likelihood is only accessible for Markovian time series. The seminal paper [25] proved asymptotic normality of Whittle estimates for Gaussian and linear causal sequences, in

2000 *Mathematics Subject Classification.* Primary: 60; Secondary: 60F, 62M10, 62M15.

Key words and phrases. Limit theorems, Time Series, Auto Correlation, Whittle estimator, Weakly dependent.

This work has been developed within the MME-DII center of excellence (ANR-11-LABEX-0023-01) and Mathamsud 16-MATH-03 SIDRE project.

the Gaussian case again [15] provides oracle bounds for its convergence. Moreover [23] considered strong mixing processes. The case of LARCH(1) processes is in [17] and more general GARCH models are considered in [24] for completely observed ARCH type models.

Here a stationary sequence with Bernoulli marginal distributions is designed for setting the fact that data are missing or not.

We obtain the asymptotic normality of a quasi-periodogram and the asymptotic properties of the quasi-Whittle estimator under stationary missing data. A very rich class of weakly dependent time series, including those previously quoted (with independent missing data procedures) is considered. Our technique also includes never studied non-causal, non-linear and non-mixing processes such as the simple AR(1)-model with Bernoulli innovations.

Besides proving the CLT for quasi-Whittle estimates the present work yields $\mathbb{L}^2$-bounds for its MSE (Mean Squares Error) which is a quite challenging question since the proposed quasi-periodogram essentially writes as linear combinations of ratios. Covariance estimation results into a ratio, due to the fact that data are incompletely observed, thus the considered quasi-periodogram take now a complicated form. This is widely known that ratios moments are not easy to compute but [7] provides us with an essential tool for deriving bounds for ratios moments. We thus bound the MSE of the involved quasi-periodogram.

The paper is organized as follows. In Section 2 we first introduce the estimates of the autocovariances under missing data. Then we recall features and results concerning Whittle estimation for completely observed data.

Section 3 is devoted to the the asymptotic properties of the quasi-periodogram built from the estimation of the covariances. This quasi-periodogram is integrated over a Sobolev class, we obtain an uniform $\mathbb{L}^2$-convergence with $\sqrt{N}-$rate which implies an uniform SLLN. An uniform CLT is also proved for the normalized process.

In section 4, a.s. consistency and asymptotic normality are proved for the quasi-Whittle estimate by applying results for section 3.

A simulation example for an ARCH(1)-model with discrete innovations (which does not exhibit mixing properties) demonstrates the performances of this quasi-Whittle estimate.

A last section 6 gives proofs and technical results of a proper interest are given in the last section. Among the technical results we prove SLLN, CLT and $\mathbb{L}^2$-convergence of the autocovariances estimates as well as a Donsker type CLT which allows to consider change point procedures.

2. Preliminaries

2.1. **Time series process with missing observations.** Let $(X_k)_{k\in\mathbb{Z}}$ be a stationary time series with zero mean and autocovariance function $\gamma_X(\ell)$. The aim of this article is to study autocovariance and spectral estimation when in the time series $(X_k)_{k\in\mathbb{Z}}$ some observations may be missed. In this context, the data X_k are not observed for each $k = 1, 2, \dots, N$ but rather only for $k = k_1, k_2, \dots, k_m$, where $k_j - k_{j-1} \geq 1$. Following [20], we introduce the amplitude modulated observations

$$Y_k = C_k X_k, \qquad k = 1, \dots, N \tag{1}$$

where

$$C_k = \begin{cases} 1, & \text{if } X_k \text{ is measured,} \\ 0, & \text{if } X_k \text{ is missing.} \end{cases}$$

We assume throughout the paper that $(X_k)_{k\in\mathbb{Z}}$ and $(C_k)_{k\in\mathbb{Z}}$ are independent sequences.

In order that the following convergence holds

$$\overline{C}_N \quad = \quad N^{-1}\sum_{k=1}^{N} C_k \xrightarrow{a.s.} \mu_C.$$

we shall assume that the sequence $(C_k)_{k\in\mathbb{Z}}$ is ergodic.

We set $\mu_C = \mathbb{E}C_0$ and $\gamma_C(\ell) = \mathrm{Cov}(C_0, C_\ell)$ and for notational convenience, we set

$$\nu(\ell) = \mathbb{E}C_0C_\ell = \gamma_C(\ell) + \mu_C^2.$$

In order to be able to estimate covariances we shall need the following assumption throughout the paper

Assumption (A). $\mathbb{E}C_0 \neq 0$ and $\mathbb{E}C_0C_\ell \neq 0$ for each $\ell = 0, 1, \ldots$

Remark 1. *The Assumption* **(A)** *essentially means that we indeed to observe a reasonable amount of data. The condition* **(A)** *holds if* $\mathbb{E}C_0 = \mathbb{P}(C_0 = 1) \neq 0$.

This follows from ergodicity which holds eg. for the case of independently distributed C.

We denote by $\overline{X}_N$, $\overline{Y}_N$ the sample means. Usual estimates of the autocovariance coefficients $\gamma_Y(\ell) = \mathrm{Cov}(Y_k, Y_{k+\ell})$ are empirical estimates $\widehat{\gamma}_{Y,N}(\ell)$ and $\widehat{\nu}_N(\ell)$ as follow:

$$\widehat{\gamma}_{Y,N}(\ell) = \begin{cases} \frac{1}{N-\ell}\sum_{k=1}^{N-\ell} Y_kY_{k+\ell}, & \text{if} \quad 0 \le \ell < N. \\ \frac{1}{N-\ell}\sum_{k=1-\ell}^{N} Y_kY_{k+\ell}, & \text{if} \quad -N < \ell < 0, \end{cases}$$

$$\widehat{\nu}_N(\ell) \; = \begin{cases} \frac{1}{N-\ell}\sum_{k=1}^{N-\ell} C_kC_{k+\ell}, & \text{if} \quad 0 \le \ell < N. \\ \frac{1}{N-\ell}\sum_{k=1-\ell}^{N} C_kC_{k+\ell}, & \text{if} \quad -N < \ell < 0. \end{cases}$$

Both estimates are obtained from the observations $\{Y_1, \ldots, Y_N\} = \{X_1C_1, \ldots, X_NC_N\}$ according to Equation (1) and they are unbiased. The following estimator of the theoretical ACF of interest is defined under the Equation (1) and was introduced by [20],

$$\widetilde{\gamma}_{X,N}(\ell) = \frac{\widehat{\gamma}_{Y,N}(\ell)}{\widehat{\nu}_N(\ell)}, \qquad \text{if} \qquad \widehat{\nu}_N(\ell) \neq 0. \tag{2}$$

Quote that Assumption **(A)** entails together with any type law of large number that

$$\lim_{N\to\infty} \mathbb{P}(\widehat{\nu}_N(\ell) = 0) \neq 0$$

thus this empirical covariances are ultimately well defined.

The correlation function $\rho_X(\ell)$ is estimated by $\widehat{\rho}_{X,N}(\ell) = \widetilde{\gamma}_{X,N}(\ell)/\widetilde{\gamma}_{X,N}(0)$.

The asymptotic properties of the estimator (2) were investigated under various assumptions on the noise of the linear representation $(\epsilon_k)_{k\in\mathbb{Z}}$ and assuming that $(C_k)_{k\in\mathbb{Z}}$ is asymptotically stationary in [13] for stationary processes of the form

$$X_k = \sum_{j=0}^{\infty} \beta_j \xi_{k-j}, \qquad \sum_{j=0}^{\infty} \beta_j^2 < \infty,$$

where the white noise process $(\xi_k)_{k\in\mathbb{Z}}$ consists of uncorrelated random variables with mean zero and variance 1.

More recently [26] compare three estimators of the autocorrelation function for a stationary process with missing observations. The first estimator is (2) and the other estimators are extensions of this one. The authors derive the asymptotic distribution for both short memory and long memory models for those three estimators of the ACF with missing observations. They impose the same assumptions on the innovations $(\xi_k)_{k\in\mathbb{Z}}$ as those in [13].

The construction of the estimator $\widetilde{\gamma}_{X,N}(\ell)$ defined by (2) assumes that the observed process $(Y_k)_{k\in\mathbb{Z}}$ is centered. This assumption is convenient for deriving the asymptotic behavior, but it is not necessarily the best formulation from the theoretic viewpoint and it is unrealistic. An alternative estimate representation for the sample ACF of time series in presence of missing observations can be proposed

$$\breve{\gamma}_{X,N}(\ell) = \widetilde{\gamma}_{X,N}(\ell) - \left\{\overline{Y}_N\right\}^2 \tag{3}$$

Using the fact that C_k is a non-negative and bounded random variables we derive that $\breve{\gamma}_{X,N}(\ell) - \widetilde{\gamma}_{X,N}(\ell)$ is arbitrarily small. Therefore $\breve{\gamma}_{X,N}(\ell) - \widetilde{\gamma}_{X,N}(\ell) = o_{\mathbb{P}}(1)$ so that limiting distributional results proved for $\widetilde{\gamma}_{X,N}(\ell)$ will hold for $\breve{\gamma}_{X,N}(\ell)$.

We thus may assume that X_k has zero-mean in order to simplify proofs: this problem is directly addressed in [13], Chapter 5; those authors prove that the same results hold true also in case a mean correction has been performed.
On the other hand,

$$Z_k^{(\ell)} = C_k C_{k+\ell}(X_k X_{k+\ell} - \gamma_X(\ell)), \tag{4}$$

then $\delta(\ell) \equiv \mathbb{E}Z_k^{(\ell)} = \gamma_X(\ell)\nu(\ell)$ and $\widehat{\delta}_N(\ell) = \widehat{\gamma}_{Y,N}(\ell) - \gamma_X(\ell)\widehat{\nu}_N(\ell)$.

2.2. **Whittle estimator for totally observed time series.** After [25] working with Gaussian data, [8] and [2] introduced Whittle's parametric estimation in a framework general dependence. We consider here situations with non totally observed data. Let $X = (X_k)_{k\in\mathbb{Z}}$ be a zero mean fourth-order stationary time series with real values. Denote $(\gamma_X(s))_s$ the covariogram of X, and $\kappa_4(a,b,c)$ the fourth order cumulants of X, such that for all $a,b,c \in \mathbb{Z}$:

$$\begin{aligned}
\gamma_X(a) &= \mathrm{Cov}(X_0, X_a) = \mathbb{E}\left(X_0 X_a\right), \\
\kappa_4(a,b,c) &= \mathbb{E}X_0X_aX_bX_c - \mathbb{E}X_0X_a\mathbb{E}X_bX_c - \mathbb{E}X_0X_b\mathbb{E}X_aX_c - \mathbb{E}X_0X_c\mathbb{E}X_aX_b.
\end{aligned}$$

The following assumptions on the process X are used to define the limit variance of the periodogram:

Assumption (M): X *is such that:*

$$\gamma = \sum_{\ell\in\mathbb{Z}} \gamma_X^2(\ell) < \infty, \quad \textit{and}, \quad \kappa_4 = \sum_{a,b,c} |\kappa_4(a,b,c)| < \infty.$$

The periodogram of X writes as:

$$I_{X,N}(\lambda) = \frac{1}{2\pi \cdot N}\left|\sum_{k=1}^N X_k e^{-ik\lambda}\right|^2, \qquad \text{for } \lambda \in [-\pi, \pi[.$$

Now, let $g : [-\pi,\pi[\to \mathbb{R}$ a 2π-periodic function such that $g \in \mathbb{L}^2([-\pi,\pi[)$ and define the integrated periodogram and spectrum of X:

$$J_{X,N}(g) = \int_{-\pi}^{\pi} g(\lambda) I_{X,N}(\lambda)\, d\lambda, \qquad J(g) = \int_{-\pi}^{\pi} g(\lambda) f_X(\lambda)\, d\lambda \tag{5}$$

We denote here by f_X spectral density of X, that exists in $\mathbb{L}^2([-\pi,\pi[)$ from Assumption **(M)**:

$$f(\lambda) = \frac{1}{2\pi}\sum_{\ell\in\mathbb{Z}} \gamma_X(\ell)\, e^{i\ell\lambda}, \qquad \text{for } \lambda \in [-\pi,\pi[.$$

Recall that

$$I_{X,N}(\lambda) = \frac{1}{2\pi}\sum_{|\ell|<N} \gamma_{X,N}(\ell) e^{-i\ell\lambda)}, \quad \gamma_{X,N}(\ell) = \frac{1}{N}\sum_{j=1\vee(1-\ell)}^{(N-\ell)\wedge N} X_j X_{j+\ell}$$

which is a biased estimate of $\gamma_X(\ell)$, and

$$\mathbb{E}\gamma_{X,N}(\ell) = \frac{N-\ell}{N} \cdot \gamma_X(\ell) \,\Big(= \frac{N-\ell}{N} \mathbb{E}\widehat{\gamma}_{X,N}(\ell)\Big).$$

A special case of the integrated periodogram is the Whittle contrast

$$\beta \mapsto J_n(h_\beta)$$

where h_β belongs to a class of regular functions: the Whittle estimate minimizes this contrast.

Hence uniform limit theorems (over a class of function $\mathcal{G}$) for the integrated periodogram $(J_n(g))_{g\in\mathcal{G}}$ are an appropriate tool to derive uniform limit theorems of the Whittle contrast.

Additional conditions with respect to the regularity of the spectral density are needed to obtain a reasonable limit theory for this Whittle estimate.

Let $X = (X_k)_{k\in\mathbb{Z}}$ be a time series satisfying Assumption **(M)**. We denote by f the spectral density of X (see eqn. (12) if $C_k \equiv 1$). Define

$$\sigma^2 = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(\lambda)\, d\lambda\right). \tag{6}$$

Following Rosenblatt (1985) we shall assume that the one-step prediction error variance satisfies $2\pi\sigma^2 > 0$. Assume that f belongs to the family of functions defined in the form

$$f(\lambda) = f_{(\beta,\sigma^2)}(\lambda) = \sigma^2 \cdot g_\beta(\lambda) \text{ for all } \lambda \in [-\pi,\pi[, \tag{7}$$

the function f thus depends on a finite number of unknown parameters, a variance term σ^2 and a $\mathbb{R}^p$-vector β, where $\beta = (\beta^{(1)},\dots,\beta^{(p)})$. The normalization condition (6) implies:

$$\int_{-\pi}^{\pi} \log g_\beta(\lambda)\, d\lambda = 0. \tag{8}$$

Denote also σ^* and $\beta^* = (\beta^{(1)*},\dots,\beta^{(p)*})$ the true value of σ and β. As a consequence, for all $\lambda \in [-\pi,\pi[$, we will now denote $\sigma^{*2} g_{\beta^*}(\lambda)$ the spectral density of X.

In order to prove the asymptotic properties of the Whittle estimator we consider a space $\mathcal{G}$ of functions regular enough to contain h_β for each acceptable value of the parameter β.

The Sobolev space $\mathcal{G} = \mathcal{H}_s$ for $s > \frac{1}{2}$ is such a nice space because of its simple Hilbert structure:

$$\mathcal{H}_s = \left\{g \in \mathbb{L}^2[-\pi,\pi[\,;\, g(-x) = g(x), \|g\|^2_{\mathcal{H}_s} < \infty\right\}, \quad \text{with} \quad \|g\|^2_{\mathcal{H}_s} = \sum_{\ell\in\mathbb{Z}} (1+|\ell|)^{2s}|g_\ell|^2.$$

for g a 2π–periodic function such that $g \in \mathbb{L}^2([-2\pi,2\pi[)$ and $g(\lambda) = \sum_{\ell\in\mathbb{Z}} g_\ell e^{i\ell\lambda}$.

This space $\mathcal{H}_s$ is included in the space $C^\star$ of continuous and 2π–periodic functions and $\|g\|_\infty = \sup_{[-\pi,\pi[} |g| \le \sqrt{c_s} \cdot \|g\|_{\mathcal{H}_s}$ with

$$c_s = \sum_{\ell\in\mathbb{Z}} \frac{1}{(1+|\ell|)^{2s}}. \tag{9}$$

We thus introduce the following conditions **(C)**:

- *Condition C1: the true values σ^* and β^* are such that $\sigma^* > 0$ and β^* lies in a region $\mathcal{K} \subset \mathbb{R}^p$ where $\mathcal{K}$ is an open and relatively compact set.*
- *Condition C2: if β_1, β_2 are distinct elements of $\mathcal{K}$, the set $\{\lambda \in [-\pi,\pi[,\ g_{\beta_1}(\lambda) \neq g_{\beta_2}(\lambda)\}$ has positive Lebesgue measure.*
- *Condition C3: for all $\beta \in \mathcal{K}$, $g_\beta^{-1} \equiv \frac{1}{g_\beta} \in \mathcal{H}_s$ and $\sup_{\beta\in\mathcal{K}} \|g_\beta^{-1}\|_{\mathcal{H}_s} < \infty$ with $s > 1/2$.*
- *Condition C4: for all $\lambda \in [-\pi,\pi[$, the function $\beta \mapsto g_\beta^{-1}(\lambda)$ is continuous on $\mathcal{K}$.*

- *Condition C5: for all $\lambda \in [-\pi, \pi[$, the function $\beta \mapsto g_\beta^{-1}(\lambda)$ is twice continuously differentiable on $\mathcal{K}$ and $\beta \mapsto \int_{-\pi}^{\pi} \log(g_\beta(\lambda))\, d\lambda$ can be differentiated twice under the integral sign.*
- *Condition C6: there exists $s > 1/2$ such that for all $\beta \in \mathcal{K}$ and $(i, j) \in \{1, \dots, p\}$, $\left(\frac{\partial g_\beta^{-1}}{\partial \beta^{(i)}}\right)_\beta, \left(\frac{\partial^2 g_\beta^{-1}}{\partial \beta^{(i)} \partial \beta^{(j)}}\right)_\beta \in \mathcal{H}_s$ and $\sup_{\beta \in \mathcal{K}} \left\| \frac{\partial^2 g_\beta^{-1}}{\partial \beta^{(i)} \partial \beta^{(j)}} \right\|_{\mathcal{H}_s} < \infty$.*
- *Condition C7: for all $\beta \in \mathcal{K}$, the function $\lambda \mapsto g_\beta(\lambda)$ is continuously differentiable on $[-\pi, \pi[$.*

Let $(X_1, \dots, X_N)$ be a sample from the process X. As usual define Whittle estimator of β^* and σ^{*2} as:

$$\beta_N = \mathrm{Argmin}_{\beta \in \mathcal{K}} J_{X,N}(g_\beta^{-1}) = \mathrm{Argmin}_{\beta \in \mathcal{K}} \int_{-\pi}^{\pi} \frac{I_{X,N}(\lambda)}{g_\beta(\lambda)}\, d\lambda \text{ and } \sigma_N^2 = \frac{1}{2\pi} J_{X,N}(g_{\beta_N}^{-1}) \tag{10}$$

Condition (C2) implies that β_N is uniquely defined.

For a practical use those definitions β_N and σ_N^2 are modified to answer the previous question of centering and the integral are replaced by its approximation by Riemann sums:

$$\beta_N' = \mathrm{Argmin}_{\beta \in \mathcal{K}} \frac{2\pi}{N} \sum_{k=1}^{N} \frac{I_{X,N}'(\pi k/N)}{g_\beta(\pi k/N)}, \ \sigma_N'^2 = \frac{1}{N} \sum_{k=1}^{N} \frac{I_{X,N}'(\pi k/N)}{g_{\beta_N'}(\pi k/N)},$$

$$I_{X,N}'(\lambda) = \frac{1}{2\pi \cdot N} \Big| \sum_{k=1}^{N} (X_k - \overline{X}_N) e^{-ik\lambda} \Big|^2 \text{ for } \lambda \in [-\pi, \pi[.$$

Let us write $f(\lambda) = f(\lambda, \beta)$ where $\beta = (\beta^{(0)}, \beta^{(1)}, \dots, \beta^{(p)})$, we define the full Whittle's contrast as

$$U_N(\beta) = \int_{-\pi}^{\pi} \log f(\lambda, \beta) d\lambda + \int_{-\pi}^{\pi} \frac{I_{X,N}(\lambda)}{f(\lambda, \beta)} d\lambda.$$

This function is related with the following contrast function

$$U(\beta, \beta^*) = \int_{-\pi}^{\pi} \log f(\lambda, \beta) d\lambda + \int_{-\pi}^{\pi} \frac{f(\lambda, \beta^*)}{f(\lambda, \beta)} d\lambda,$$

assuming β^* the true parameter value. Now:

$$\frac{\partial U(\cdot, \beta^*)}{\partial \beta^{(j)}} \Big|_{\beta = \beta^*} = 0, \qquad j = 1, \dots, p$$

and also

$$\frac{\partial U_N}{\partial \beta^{(j)}}(\beta) = -\int_{-\pi}^{\pi} \frac{\partial \log f(\lambda, \beta)}{\partial \beta^{(j)}} \frac{[I_{X,N}(\lambda) - f(\lambda, \beta)]}{f(\lambda, \beta)}\, d\lambda.$$

Whittle's estimator β_N is the only solution of the equation $\frac{\partial U_N}{\partial \beta^{(j)}}(\beta_N) = 0$ for $j = 1, \dots, p$. Now we must modified conditions C1-6 conveniently, to be applied to function f instead of g. In first place by using the general theory of contrast functions, conditions C1-4 and in theorem 1 they shows that $\beta_N \to \beta^*$ almost surely.

Moreover under the conditions C1-6 and by using a CLT [2] we are able to derive for some suitable $\Sigma(\beta)$:

$$\begin{aligned} \sqrt{N}\Big(\frac{\partial U_N}{\partial \beta^{(j)}}(\beta)\Big)_{1 \le j \le p} &= -\sqrt{N} \int_{-\pi}^{\pi} \Big(\frac{\partial \log f(\lambda, \beta)}{\partial \beta^{(j)}}\Big)_{1 \le j \le p} \frac{(I_{X,N}(\lambda) - f(\lambda, \beta))}{f(\lambda, \beta)}\, d\lambda \\ &\xrightarrow[N \to \infty]{\mathcal{D}} \mathcal{N}_p(0, \Sigma(\beta)). \end{aligned}$$

The central limit theorem for Whittle estimates is then based upon a careful use of the $\Delta-$method.

3. Uniform limit theorems for the quasi-periodogram under missing observations

In this section we consider the estimation of functionals of the spectral density function from time series in presence of missing observations. Using the estimator of the covariance of $(X_k)_{k\in\mathbb{Z}}$ given by $\widetilde{\gamma}_{X,N}(\ell)$, we introduce estimates of the spectral density.

More precisely, we introduce a quasi-periodogram defined with the empirical covariance $\widetilde{\gamma}_{X,N}(\ell)$ of the process in presence of missing observations

$$\widetilde{I}_{X,N}(\lambda) = \sum_{\ell\in\mathbb{Z}} \widetilde{\gamma}_{X,N}(\ell)e^{-i\ell\lambda} = \sum_{|\ell|<N} \frac{\widehat{\gamma}_{Y,N}(\ell)}{\widehat{\nu}_N(\ell)} e^{-i\ell\lambda}. \tag{11}$$

A new feature of this quasi-periodogram is that it includes ratios. This is a main point of the paper that [7]'s results allow us to bound moments of those expressions. For example the mean square error is evaluated contrary to what was proved in all the previous papers. This provides us with a general proof including mean squares convergences. The proofs are then simple adaptations of those in [2].

Assume now that we wish to estimate $J_X(g)$ (see (5), quote here that the spectral density f_X of X coincides with f in eqn. (12) in case of totally observed data, $C_k \equiv 1$), we use the estimates of the covariance of $(X_k)_{k\in\mathbb{Z}}$, $\widetilde{\gamma}_{X,N}(\ell)$, to build an estimate of the integrated periodogram of $(X_k)_{k\in\mathbb{Z}}$:

$$\widetilde{J}_{X,N}(g) = \int_{-\pi}^{+\pi} g(\lambda)\widetilde{I}_{X,N}(\lambda)d\lambda.$$

Under general conditions the integrated periodogram is a consistent estimator of the $J_X(g)$ provided the spectral density $f_X(\lambda)$ is well defined.

We consider the asymptotic behavior of $\mathbb{E}|\widetilde{J}_{X,N}(g) - J_X(g)|^2$ for $g \in \mathcal{H}_s$ if $s > \frac{1}{2}$.

The dual space $\mathcal{H}'_s$ of $\mathcal{H}_s$ is equipped with the norm

$$\|T\|^2_{\mathcal{H}'_s} = \sup_{\|g\|_{\mathcal{H}_s}\le 1} |T(g)|^2 = \sum_{\ell\in\mathbb{Z}} \frac{1}{(1+|\ell|)^{2s}} |T(e^{i\ell\lambda})|^2.$$

Note that $J_X \in \mathcal{H}'_s$ and it will be proved in Theorem 1 that *(a.s.)* $\widetilde{J}_{X,N} \in \mathcal{H}'_s$ under additional assumptions.

We shall use dependence conditions as follows. Let $(Z_k)_{k\in\mathbb{Z}}$ be a random process. Rosenblatt (1956) and Dedecker and Doukhan (2003) respectively define strong mixing and θ-weak dependence through sequences of coefficients from the respective relations

$$\alpha_Z(i) \to_{i\to\infty} 0, \qquad \theta_Z(i) \to_{i\to\infty} 0$$

(below we will set those conditions for *resp.* $Z_k = X_k, C_k, Y_k$ or $Z_k = (X_k, C_k)$.)

Theorem 1 (Uniform Spectral SLLN). *Assume that conditions* **(M)-(A)** *hold, and that* $\mathbb{E}|X_0|^r < \infty$, $s > 1 + 2/r$. *Moreover assume*

- *either that the sequence C is strong mixing and:*

$$\sum_{i=1}^{\infty} i^{\frac{8}{r-4}} \cdot \alpha_C(i) < \infty,$$

- *or that the sequence C is θ-weakly dependent and:*

$$\sum_{i=1}^{\infty} i^{\frac{8}{r-4}} \cdot \theta_C(i) < \infty.$$

Then

$$\|\widetilde{J}_{X,N} - J_X\|_{\mathcal{H}'_s} \xrightarrow[N\to\infty]{a.s.} 0, \qquad \textit{and in } \mathbb{L}^2.$$

In order to prove this main result we need the following essential Lemma based on [7]. Here indeed division is a tricky argument.

Lemma 1 (Spectral moment inequality). *Assume that conditions of Theorem 1 hold, then there exists some $a > 0$ such that*

$$\mathbb{E}\|\widetilde{J}_{X,N} - J_X\|^2_{\mathcal{H}'_s} \leq \frac{a}{\sqrt{N}}.$$

The proof of those laws of large numbers are reported in § 6.1.

The previous assumptions are needed for those laws of large numbers for $\widetilde{J}_{X,N}$. We shall need additional dependence conditions on the stationary processes (X_k), and (C_k) to prove central limit theorems. They follow

Assumptions (D): $\mathbb{E}|X_0|^r < \infty$ *for some $r > 4$, that condition* **(A)** *holds.*
Moreover the independent processes X and C are either

- *strongly mixing with*

$$\sum_{i=0}^{\infty} i^{\frac{4}{r-4}} \cdot \alpha_X(i) < \infty, \text{ and } \quad \sum_{i=0}^{\infty} i^{\frac{8}{r-4}} \cdot \alpha_C(i) < \infty,$$

or they are

- *θ-weakly dependent with*

$$\sum_{i=0}^{\infty} i^{\frac{2}{r-4}} \cdot \theta_X^{\frac{r-2}{r-1}}(i) < \infty, \text{ and } \quad \sum_{i=1}^{\infty} i^{\frac{8}{r-4}} \cdot \theta_C(i) < \infty.$$

Quote that the process C admits marginals bounded by 1.

In order to describe the limiting behavior of $W_N = \sqrt{N}(\widetilde{J}_{X,N} - J)$ in distribution we need some further notations. Under Assumption **(M)**, we may define for any $\lambda, \mu, \zeta \in \mathbb{R}$, the quasi-spectral and quasi-bispectral densities:

$$f(\lambda) = \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \mathbb{E}C_0 C_\ell \cdot \gamma_X(\ell) e^{i\ell\lambda} \tag{12}$$

$$f_4(\lambda, \mu, \zeta) = \frac{1}{(2\pi)^3} \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} \sum_{c=-\infty}^{\infty} \mathbb{E}C_0 C_a C_b C_c \cdot \kappa_4(a, b, c) e^{i(a\lambda + b\mu + c\zeta)} \tag{13}$$

and for g_1 and g_2 in $\mathcal{H}_s$, the limiting covariance Γ of the process W_N writes:

$$\Gamma(g_1, g_2) = 4\pi \int_{-\pi}^{\pi} g_1(\lambda) g_2(\lambda) f^2(\lambda)\, d\lambda + 2\pi \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g_1(\lambda) g_2(\mu) f_4(\lambda, -\mu, \mu)\, d\lambda\, d\mu. \tag{14}$$

Theorem 2 (Uniform Spectral CLT). *Assume that the dependence assumptions* **(M)**, *and* **(D)** *hold true with $s > 1 + 2/r$, then the following functional central limit theorem holds*

$$W_N \equiv \sqrt{N}\big(\widetilde{J}_{X,N} - J_X\big) \xrightarrow[N\to\infty]{\mathcal{D}} W, \quad \text{in the space } \mathcal{H}'_s.$$

Here $(W(g))_{g\in\mathcal{H}_s}$ denotes a centered Gaussian process such that

$$\mathbb{E}W(g_1)W(g_2) = \Gamma(g_1, g_2)$$

with Γ defined in eqn. (14).

Remark 2. *Since this theorem makes use of* [7]*'s division trick (see Lemma 1) we observe that assumptions* **(D)** *are assymmetrical with respect to the random processes* X *and* C. *Let us assume Riemaniann decays* $\alpha_X(i) \sim i^{-\alpha_X}$, $\alpha_C(i) \sim i^{-\alpha_C}$ *and* $\theta_X(i) \sim i^{-\theta_X}$, $\theta_C(i) \sim i^{-\theta_C}$ *for either the strong mixing coefficients or the weak dependence coefficients, then those assumptions need:*

$$\alpha_X > \frac{r}{r-4}, \qquad \alpha_C > \frac{r+4}{r-4},$$

and for the θ*-weak coefficients they we ask:*

$$\theta_X > \frac{r-1}{r-4}, \qquad \theta_C > \frac{r+4}{r-4}.$$

Now consider corollary 2 in [2] *then it is clear that the same assumptions are essentially needed for this uniform CLT in case of power rate decays if* $C_k \equiv 1$.

A careful attention proves in fact that we slightly improve on this previous work; see Remarks 5 and 6 second items, for the needed tools related to quantile functions.

The proofs of those central limit theorems are reported in § 6.2.

4. Quasi-Whittle estimation of irregularly sampled data

Define now quasi-Whittle estimates by mimicking (10) as:

$$\widehat{\beta}_N = \text{Argmin}_{\beta\in\mathcal{K}} \widetilde{J}_{X,N}(g_\beta^{-1}) = \text{Argmin}_{\beta\in\mathcal{K}} \int_{-\pi}^{\pi} \frac{\widetilde{I}_{X,N}(\lambda)}{g_\beta(\lambda)}\, d\lambda, \text{ and } \widehat{\sigma}_N^2 = \frac{1}{2\pi}\widetilde{J}_{X,N}(g_{\widehat{\beta}_N}^{-1}) \tag{15}$$

Condition C2 implies that $\widehat{\beta}_N$ *a.s.* uniquely defined.

Theorem 3. *Let* X *satisfy the assumptions of Theorem 1 and also that* $\mathbb{E}X_0 = 0$.

Under Conditions (A) and C1-4:

$$\widehat{\beta}_N \xrightarrow[N\to\infty]{a.s.} \beta^* \quad \text{and} \quad \widehat{\sigma}_N^2 \xrightarrow[N\to\infty]{a.s.} \sigma^{*2}.$$

Proof. From Theorem 1 and Condition C3 with probability 1,

$$\lim_{N\to\infty} J_N(g_\beta^{-1}) = J(g_\beta^{-1}),$$

uniformly on $\beta \in \mathcal{K}$. From Conditions C2 and normalization condition (8),

$$J(g_\beta^{-1}) > 2\pi\sigma^{*2} = J(g_{\beta^*}^{-1}) \text{ for all } \beta \neq \beta^*$$

The proof follows the lines in theorem 4 [2]. □

Now again the discretized and centered version of the estimates write:

$$\breve{\beta}_N = \text{Argmin}_{\beta\in\mathcal{K}} \frac{2\pi}{N}\sum_{k=1}^{N} \frac{\breve{I}_{X,N}(\pi k/N)}{g_\beta(\pi k/N)}, \quad \breve{\sigma}_N^2 = \frac{1}{N}\sum_{k=1}^{N} \frac{\breve{I}_{X,N}(\pi k/N)}{g_{\breve{\beta}_N}(\pi k/N)},$$

Now the estimate of X's covariances makes use of eqn. (3), and the quasi-periodogram (11) is replaced by its centered version:

$$\breve{I}_{X,N}(\lambda) = \sum_{|\ell|<N} \frac{\breve{\gamma}_{Y,N}(\ell)}{\widehat{\nu}_N(\ell)} e^{-i\ell\lambda}.$$

Corollary 1 (Corrected estimates)**.** *Let* X *satisfy the assumptions of Theorem 1, then*

$$\breve{\beta}_N \xrightarrow[N\to\infty]{a.s.} \beta^* \quad \text{and } \breve{\sigma}_N^2 \xrightarrow[N\to\infty]{a.s.} \sigma^{*2}.$$

Proof. Consider the process $(X_k - \mathbb{E}X_0)_{k\in\mathbb{Z}}$ instead of $X = (X_k)_{k\in\mathbb{Z}}$ and the proof is that of corollary 4 in [2]. □

Theorem 4. *Let X satisfy either the assumptions of Theorem 2. Under Conditions C1-6 and if the matrix $W^* = (w^*_{ij})_{1\le i,j\le p}$, with*

$$w^*_{ij} = \int_{-\pi}^{\pi} g^2_{\beta^*}(\lambda)\Big(\frac{\partial g_\beta^{-1}}{\partial \beta^{(i)}}\Big)_{\beta^*}(\lambda)\Big(\frac{\partial g_\beta^{-1}}{\partial \beta^{(j)}}\Big)_{\beta^*}(\lambda)\,d\lambda$$

is nonsingular, then

$$\sqrt{N}(\widehat{\beta}_N - \beta^*) \quad \xrightarrow[N\to\infty]{\mathcal{D}} \quad \mathcal{N}_p\Big(0\,,\ \frac{1}{\sigma^{*4}}W^{*-1}Q^*W^{*-1}\Big), \tag{16}$$

with a matrix $Q^* = (q^*_{ij})_{1\le i,j\le p}$ whose entries are defined by

$$q^*_{ij} = 2\pi\left(2\sigma^{*4}w^*_{ij} + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_4(\lambda,\mu,-\mu)\Big(\frac{\partial g_\beta^{-1}}{\partial \beta^{(i)}}\Big)_{\beta^*}(\lambda)\Big(\frac{\partial g_\beta^{-1}}{\partial \beta^{(j)}}\Big)_{\beta^*}(\mu)\,d\lambda\,d\mu\right).$$

The proof is that of theorem 5 [2] and the following result is similar to their theorem 6.

Theorem 5. *Let X satisfy either the assumptions of Theorem 2. Under Conditions C1-7, then*

$$\sqrt{N}(\widehat{\sigma}^2_N - \sigma^{*2}) \ \xrightarrow[N\to\infty]{\mathcal{D}} \ \mathcal{N}\Big(0\,,\,2\sigma^{*4} + 2\pi\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_4(\lambda,\mu,-\mu)g^{-1}_{\beta^*}(\lambda)g^{-1}_{\beta^*}(\mu)\,d\lambda\,d\mu\Big). \tag{17}$$

Moreover, $\sqrt{N}(\widehat{\sigma}^2_N - \sigma^{*2})$ and $\sqrt{N}(\widehat{\beta}_N - \beta^*)$ are jointly asymptotically normal with covariance given by

$$\lim_{n\to\infty}\sqrt{N}\big(Cov(\widehat{\sigma}^2_N,\,\widehat{\beta}^{(i)}_N)\big)_{1\le i\le p}$$
$$= \frac{1}{\sigma^{*2}W^*}\cdot\left(2\pi\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_4(\lambda,\mu,-\mu)g^{-1}_{\beta^*}(\lambda)\left(\frac{\partial}{\partial\beta^{(j)}}g^{-1}_{\beta^*}(\mu)\right)\,d\lambda\,d\mu\right)_{1\le j\le p}.$$

Remark 3.
- *We defer the reader to* [2] *for more precise statements of the examples.*
- *As in* [2] *we are not in position to exhibit a non-centered and discretized version of Theorem 5. Indeed such results needs a $\sqrt{N}$-acurate uniform approximation of the quasi-periodogram $\widetilde{I}_{X,N}(\lambda)$.*

5. Simulation experiments

In this section we analyse the finite sample behavior of the quasi-Whittle estimate by Monte Carlo study. 5000 repetitions of samples with length 1000 were processed for each case.

We consider here an independent sequence C Bernoulli trials with $q = \{C_k = 0\}$, probabilities $q = 0$ (the completely observed case), 0.05 and 0.1 of missing observation are considered. The sample mean and Standard error (Std.) are compared for different values of the parameters and different estimates.

We consider both a linear and a non-linear process. An alternative moment-based technique is derived in each case. Moreover the 2 first models considered admit one parameter while the last one has 2 parameter. For all the forthcoming examples which are Markov processes the QMLE is a more efficient method too anyway as stated in [3] this technique does not apply simply to the case of missing observations. We compare estimates of their parameters. The comparison is thus not at the advantage of quasi-Whittle method but our point is that this is a systematic technique which really work for simulated data. Under missing data the asymptotic for such empirical estimates is already considered in [7] and there is no doubt that it will be more efficient but this relies on the simplicity of the considered models: the quasi-Whittle method systematically applies.

5.1. **AR(1)–model.** Set

$$X_k = \beta(X_{k-1} + \xi_k)$$

for $|\beta| < 1$ and $(\xi_k)_{k\in\mathbb{Z}}$ an iid $\mathbb{L}^2$-sequence. Then [1] proves that for $\beta = \frac{1}{2}$ and for Bernoulli $b(\frac{1}{2})$-distributed marginal distribution this sequence is non-mixing; quote that for $\beta = .2$ the model is not known to be non-mixing but it is still θ-weakly dependent. A simple way to estimate the parameter is to check that here $\mathbb{E}X_0 = \beta/2(1-\beta)$ thus an empirical estimate $\widetilde{\beta} = 2\widetilde{m}/(1+2\widetilde{m})$ is $\sqrt{N}$-consistent. An alternative is again to consider the quasi-Whittle estimate. Both estimates behave extremely well (both in the

Parameter	Estimate Q-Whittle	Estimates SD	Estimate Q-Moment	Estimates SD
β	$\widehat{\beta}$	SD($\widehat{\beta}$)	$\widetilde{\beta}$	SD($\widetilde{\beta}$))
q=0				
0.2	0.217	0.017	0.199	0.003
0.5	0.507	0.024	0.499	0.005
q=0.05				
0.2	0.219	0.021	0.199	0.003
0.5	0.514	0.030	0.499	0.005
q=0.10				
0.2	0.220	0.027	0.199	0.003
0.5	0.512	0.316	0.499	0.005

Table 1. AR(1) estimation under missing data, quasi-Whittle and empirical estimates.

non mixing case, if $\beta = .2$). One spectacular behavior of the Q-moment is that up 10^{-3} the quality of the estimate does not depend on the number of missing data.

5.2. **LARCH–models.** In [12], LARCH(∞)–models are a vector valued equivalent of the solution of a recursion given from an iid sequence $(\xi_k)_{k\in\mathbb{Z}}$ and a sequence of real parameters $\beta_0, \beta_1, \beta_2, \ldots$

$$Z_k = \xi_k \left(\beta_0 + \sum_{j=1}^{\infty} \beta_j Z_{k-j}\right).$$

If $\|\xi_0\|_\infty \sum_{j\ge1} |\beta_j| < 1$ for some $p \ge 1$, then there exists a stationary solution of the previous equation which is in $\mathbb{L}^p$. Anyway the parameter β is here infinite dimensional and does not fit the present frame so we restrict to the simplest case.

Eg. for LARCH(1)-models $Z_k = \xi_k(\beta_0 + \beta_1 Z_{k-1})$ this simply means that $|\beta_1| < 1$.

5.2.1. *LARCH(1)-symmetric model.* Assume first that $\beta_0 = 1$. The current model is a weak white noise with a symmetric marginal distribution in case ξ_0 admits such a symmetric distribution. Note that in [11] it is proved that this model is non-mixing in the special case of Rademacher innovations $\mathbb{P}(\xi_0 = \pm1) = \frac{1}{2}$ and $\beta \equiv \beta_1 \in \left]\frac{3-\sqrt{5}}{2}, \frac{1}{2}\right]$.

Here the stationary distribution satisfies $\mathbb{E}Z_0^2 = \beta^2\mathbb{E}Z_0^2 + \mathbb{E}\xi_0^2$ but Rademacher rvs satisfy $\mathbb{E}\xi_0^2 = 1$. Hence a consistent estimate of β writes $\widetilde{\beta} = \sqrt{1 - 1/\widehat{M}}$ for $\widehat{M}$ the empirical counterpart of $\mathbb{E}Z_0^2$. We shall process simulations for true values $\beta = .45$ and .3, the first model is thus non-mixing again and the both are θ-weakly dependent. In the simulation a stripking point is that even with only 50 receptions the Q-moment behaves the same as for 5000 repetitions, contrary to the quasi-Whittle which needs a large amount of such replications. Here again the empirical technique is better that quasi-Whittle without much surprise. Here again both estimates are quite reasonable too.

Parameter	Estimate Q-Whittle	Estimates SD	Estimate Q-Moment	Estimates SD
β	$\widehat{\beta}$	SD($\widehat{\beta}$)	$\widetilde{\beta_0}$	SD($\widetilde{\beta}$))
q=0				
0.3 0.45	0.307 0.453	0.021 0.035	0.299 0.449	0.017 0.016
q=0.05				
0.3 0.45	0.314 0.462	0.034 0.042	0.298 0.449	0.018 0.017
q=0.10				
0.3 0.45	0.319 0.467	0.039 0.049	0.298 0.449	0.018 0.017

TABLE 2. LARCH(1)-symmetric estimation under missing data, quasi-Whittle and empirical estimates.

5.2.2. *ARCH(1)-model.* We now consider a class of parametric ARCH models, in this class of time series, Whittle estimation is based on squared observations because this class of models has no linear correlation but exhibit dependence in their squares. All the experiments are based on 1000 replications of the ARCH(1) model defined by

$$X_k = \varepsilon_k \sqrt{\beta_0 + \beta_1 X_{k-j}^2}, \tag{18}$$

where (ε_k) is a sequence of iid zero-mean Gaussian random variables with unit variance.

Conditions for existence and uniqueness in $\mathbb{L}^p$ for the ARCH(1) model have been derived before since $Z_k = X_k^2$ is again a LARCH(1)-model with innovations $\xi_k = \varepsilon_k^2$. Quote the such χ_1^2−distributed innovations admit now a density and hence [6] entails that the model is geometrically but here innovations are no more centered and a moment method will now be based upon the empirical estimation of $m = \mathbb{E}X_0^2$ and $M = \mathbb{E}X_0^4$ (again under missing data) since this is standard to derive

$$\beta_0 = \left(1 - \sqrt{\frac{M - 3m^2}{3(M - m^2)}}\right) m, \qquad \beta_1 = \sqrt{\frac{M - 3m^2}{3(M - m^2)}}.$$

The quasi-Whittle estimation for β_0 and β_1 is efficient for all situations studied here. Based on this simulation study, we can conclude that the proposed method is very precise to estimate the parameters of an ARCH model with missing observations.

Both estimates are reasonable for the estimation of 'main' coefficient β_0 but the quasi-moment method (Q-Moment) is much more efficient that the quasi-Whittle estimate. This sounds reasonable the adapted moment techniques are more efficient that a general Whittle type method. In order to check the quality of those Q-moment estimates we thus consider in Table 5.2.2 the case when $\beta_0 = 1$ is known and the moment technique only relies on the only coefficient $\beta = \beta_1 \equiv 1 - 1/M$ yields a simple moment based estimation of $\beta \in [0, 1[$. Now both estimates seem to run analogously. This means maybe that estimation of parameters in high dimension for this quasi-Whittle estimate may cause some trouble. Quasi-Whittle estimate is thus a performing and general technique under missing data which is much more natural that estimation of missing values as processed as this was proved [3] for least square estimates under mixing assumptions.

Parameters		Estimates Q-Whittle		Estimates SD		Estimates Q-Moment		Estimates SD	
β_0	β_1	$\widehat{\beta_0}$	$\widehat{\beta_1}$	SD($\widehat{\beta_0}$)	SD($\widehat{\beta_0}$)	$\widetilde{\beta_0}$	$\widetilde{\beta_1}$	SD($\widetilde{\beta_0}$)	SD($\widetilde{\beta_0}$)
q=0									
0.15	0.25	0.151	0.251	0.0073	0.0987	0.155	0.225	0.016	0.079
0.15	0.30	0.153	0.263	0.0097	0.0.893	0.158	0.262	0.017	0.083
0.20	0.25	0.200	0.255	0.0084	0.991	0.207	0.226	0.021	0.077
q=0.05									
0.15	0.25	0.141	0.251	0.093	0.153	0,155	0.226	0.016	0.079
0.15	0.30	0.157	0.283	0.089	0.130	0.157	0.264	0.017	0.081
0.20	0.25	0.206	0.235	0.094	0.137	0.206	0.226	0.016	0.079
q=0.10									
0.15	0.25	0.156	0.234	0.103	0.175	0.155	0.226	0.016	0.080
0.15	0.30	0.173	0.263	0.093	0.146	0.158	0.261	0.017	0.081
0.20	0.25	0.222	0.214	0.100	0.155	0.207	0.226	0.021	0.078

TABLE 3. ARCH(1) estimation ($d = 2$) under missing data, quasi-Whittle and empirical estimates.

Parameter	Estimate Q-Whittle	Estimates SD	Estimate Q-Moment	Estimates SD
β	$\widehat{\beta}$	SD($\widehat{\beta}$)	$\widetilde{\beta_0}$	SD($\widetilde{\beta}$))
q=0				
0.25	0.257	0.087	0.247	0.049
0.30	0.311	0.079	0.247	0.049
0.50	0.499	0.089	0.496	0.063
0.75	0.741	0.085	0.713	0.073
q=0.05				
0.25	0.261	0.997	0.249	0.048
0.30	0.314	0.998	0.294	0.049
0.50	0.511	0.106	0.490	0.062
0.75	0.678	0.124	0.713	0.074
q=0.10				
0.25	0.267	0.123	0.248	0.051
0.30	0.319	0.132	0.296	0.051
0.50	0.521	0.015	0.494	0.062
0.75	0.659	0.015	0.715	0.073

TABLE 4. ARCH(1) estimation ($d = 1$) under missing data, quasi-Whittle and empirical estimates.

6. Technical Results

6.1. **Proofs of laws of large numbers.**
Proof of Theorem 1. The proof is that of theorem 1 in [2] by using Lemma 1. □
Proof of Lemma 1. We follow the lines in the proof of lemma 3 in [2]. Lemma 3 provides the asymptotic for the finite dimensional repartitions of the process W_N. Here again $\widetilde{J}_{X,N} - J_X = -T_1 - T_2 + T_3$ where the two first terms write the same way here

and only T_3 is altered by the missing data, then:

$$\mathbb{E}\|T_3\|^2_{\mathcal{H}'_s} \quad = \quad \sum_{\ell\in\mathbb{Z}} \frac{1}{(1+|\ell|)^{2s}} \cdot \mathbb{E}(\widetilde{\gamma}_{X,N}(\ell) - \gamma_X(\ell))^2 \tag{19}$$

with c_s defined in (9).

Now from [7] we derive that if

$$\|\widehat{\nu}(\ell) - \nu(\ell)\|_q \le \frac{c}{\sqrt{N-\ell}}, \tag{20}$$

for $q = \frac{2(r+4)}{r-4}$ and assumptions **(M)-(A)** hold (which implies that $\|\widehat{\delta}_N\|_2 \le c/\sqrt{N-\ell}$), then $\|\widetilde{\gamma}_{X,N}\|_2 \le c'/\sqrt{N-\ell}$.

For large values of $|\ell| \ge N/2$ the idea of Pisier lemma (already used in [7], eqn. (15)) implies with Cauchy-Schwartz inequality

$$\mathbb{E}(\widetilde{\gamma}_{X,N}(\ell) - \gamma_X(\ell))^2 \le 2(\mathbb{E}(\widetilde{\gamma}^2_{X,N}(\ell) + \gamma^2_X(\ell)) \le 2((N-\ell)^{\frac{4}{r}}\|X_0\|^4_r + \gamma^2_X(0))$$

Indeed

$$\mathbb{E}\widetilde{\gamma}^2_{X,N}(\ell) \le \mathbb{E}\max_{1\le i\le N-\ell}(X_i^2 X_{i+\ell}^2) \le (\mathbb{E}\max_{1\le i\le N-\ell}|X_i X_{i+\ell}|^{\frac{r}{2}})^{\frac{4}{r}} \le ((N-\ell)\mathbb{E}|X_0|^r)^{\frac{4}{r}}.$$

Thus in eqn. (19) we derive the result in case $s > 1 + 2/r$.

Finally [21], théorème 2.5. (b) proves that inequality (20) holds if

$$\sum_{i=1}^{\infty} i^{\frac{8}{r-4}} \alpha_C(i) < \infty.$$

Denote by G_C the generalized inverse of $x \mapsto \int_0^x Q_C(u)du$, then corollary 2 in [5] implies (20) in case:

$$\int_0^{\|C_0\|_1} (\theta_C^{-1}(u))^{\frac{q}{2}} Q_C(u)du < \infty.$$

Now C_0 being bounded so does Q_C, and from $q/2 - 1 = 8/(r-4)$ this relation now holds if

$$\sum_{i=1}^{\infty} i^{\frac{8}{r-4}} \theta_C(i) < \infty. \qquad \square$$

6.2. **Proof of the central limit theorems.** We first present a strong law of large numbers (SLLN). We will then provide proofs of the limit theorems.

Lemma 2 (SLLN)**.** *Let us assume that* $\mathbb{E}\|X_0\|^r < \infty$ *for some* $r > 2$.

If the independent processes X *and* C *are either strongly mixing with*

$$\sum_{i=0}^{\infty} \frac{1}{i+1} \cdot \alpha_X^{1-\frac{2}{r}}(i) < \infty, \text{ and } \quad \sum_{i=0}^{\infty} \frac{1}{i+1} \cdot \alpha_C^{1-\frac{2}{r}}(i) < \infty$$

or they are θ*-weakly dependent with*

$$\sum_{i=0}^{\infty} \frac{1}{i+1} \cdot \theta_X^{\frac{r-2}{r-1}}(i) < \infty, \text{ and } \quad \sum_{i=0}^{\infty} \frac{1}{i+1} \cdot \theta_C^{\frac{r-2}{r-1}}(i) < \infty$$

Then

$$\widehat{\gamma}_{Y,N}(\ell) \xrightarrow[N\to\infty]{a.s.} \mathbb{E}C_0C_\ell \cdot \gamma_X(\ell), \qquad \widehat{\nu}(\ell) \xrightarrow[N\to\infty]{a.s.} \mathbb{E}C_0C_\ell,$$

thus in case $\mathbb{E}C_0C_\ell \ne 0$*:*

$$\widetilde{\gamma}_{X,N}(\ell) \xrightarrow[N\to\infty]{a.s.} \gamma_X(\ell).$$

Remark 4. *In case those coefficients admit Riemannian decay rates $\mathcal{O}(i^{-a})$ for some $a > 0$ then the previous conditions automatically holds.*

For the more general case of spectral estimation assumption **(A)** *implies that* $\mathbb{E}C_0C_\ell \neq 0$ *for each value of* ℓ.

Proof of Lemma 2. Set either $Z_k = C_kC_{k+\ell}(X_kX_{k+\ell} - \gamma_X(\ell))$ or $Z_k = C_kC_{k+\ell}$. This is clear that $\|Z\|_{\frac{r}{2}} \leq 2\|C_0\|_\infty^2\|X_0\|_r^2$ (resp. $\|Z\|_{\frac{r}{2}} \leq \|C_0\|_\infty^2$).

Then [21]'s eqn. (3.11) for $p = 1$ derives the SLLN in the strong mixing case if

$$\sum_{i=0}^{\infty} \frac{1}{i+1} \int_0^{\alpha_Z(i)} Q_Z(u)du < \infty, \text{with } Q_Z \text{ the inverse function of } t \mapsto \mathbb{P}(|Z| > t). \quad (21)$$

Now Hölder inequality entails

$$\int_0^{\alpha_Z(i)} Q_Z(u)du \leq \left(\int_0^1 Q_Z^{\frac{r}{2}}(u)du\right)^{\frac{2}{r}} \left(\int_0^{\alpha_Z(i)} du\right)^{1-\frac{2}{r}} = \|Z\|_{\frac{r}{2}}\, \alpha_Z^{1-\frac{2}{r}}(i).$$

Under weak dependence, Theorem 3 in [5] with $p = 1$ implies the a.s. convergence of the same sums

$$\frac{1}{N-\ell} \sum_{k=1}^{N-\ell} Z_k \underset{N\to\infty}{\overset{a.s.}{\longrightarrow}} \mathbb{E}Z_0,$$

if

$$\sum_{i=0}^{\infty} \frac{1}{i+1} \alpha_Z^{1-\frac{2}{r}}(i) < \infty, \qquad \text{or} \qquad \sum_{i=0}^{\infty} \frac{1}{i+1} \theta_{(X,C)}^{\frac{r-2}{r-1}}(i) < \infty$$

For this one also needs the heredity result lemma 6, in [2] entails that there exists a constant $a > 0$ such that

$$\theta_Z(i) \leq a\, \theta_{(X,C)}^{\frac{r-2}{r-1}}(i)$$

Since X and C are independent processes this is simple to derive in both cases that

$$\alpha_Z(i) \leq \alpha_{(X,C)} \leq \alpha_X(i) + \alpha_C(i), \qquad \theta_{(X,C)}(i) \leq \theta_X(i) + \theta_C(i).$$

The subadditivity of $t \mapsto t^a$ for $a \in [0,1]$ yields the desired result. □

Remark 5. *Alternative simplified assumptions.*

- *The convergence of $\widehat{\nu}(\ell)$ holds in case*

$$\sum_{i=0}^{\infty} \frac{\alpha_C(i)}{i+1} < \infty, \qquad \textit{or} \qquad \sum_{i=0}^{\infty} \frac{\theta_C(i)}{i+1} < \infty$$

because of the boundedness of C_k which allows to set $r = \infty$, but the first convergence cannot be improved directly.

- *Under strong mixing, the expression (21) may be handled somehow differently and setting $\alpha_Z(0)=1$ as in Rio, we rewrite:*

$$\begin{aligned}
\sum_{i=0}^{\infty}\frac{1}{i+1}\int_0^{\alpha_Z(i)} Q_Z(u)du &= \int_0^1\left(\sum_{i=1}\frac{1}{i+1}1\!\!I_{[0,\alpha_Z(i)]}\right)Q_Z(u)du\\
&= \int_0^1 \sum_{\{i/\alpha_Z(i)>u\}}\frac{1}{i+1}\cdot Q_Z(u)du\\
&= \int_0^1 \sum_{i\le\alpha_Z^{-1}(u)}\frac{1}{i+1}\cdot Q_Z(u)du\\
&\le \int_0^1 \ln\left(1+\alpha_Z^{-1}(u)\right)Q_Z(u)du\\
&\le \|Z_0\|_{\frac{r}{2}}\left(\int_0^1 \ln^{\frac{r}{r-2}}\left(1+\alpha_Z^{-1}(u)\right)du\right)^{1-\frac{2}{r}}
\end{aligned}$$

To this end quote that

$$\frac{1}{i+1}\le\int_i^{i+1}\frac{du}{u}=\ln(i+1)-\ln i.$$

By using Abel transformation (discrete integration by parts) the right-hand side of the last inequality is seen to be finite in case

$$\sum_{i=1}^{\infty}\frac{\ln^{\frac{2}{r-2}} i}{i}\cdot\alpha_Z(i)<\infty$$

which is a bit better assumption than the previously used condition

$$\sum_{i=0}^{\infty}\frac{1}{i+1}\cdot\alpha_Z^{1-\frac{2}{r}}(i)<\infty$$

The previous remarks are rejected after the proof since the previous Remark 4 proves that in case of power decaying coefficients all those assumptions always hold!

Central limit theorems are now based upon weak invariance principles under the 2 weak dependence conditions used here.

Lemma 3 (Multidimensional CLT)**.** *Assume that the dependence assumptions* **(M), (D)** *hold true, then for arbitrary finite repartitions $(\ell_1,\dots,\ell_k)$:*

$$\sqrt{N}\Big(\nu(\ell_i)\big(\widetilde{\gamma}_{X,N}(\ell_i)-\gamma_X(\ell_i)\big)\Big)_{1\le i\le k}\xrightarrow[N\to\infty]{\mathcal{D}}\mathcal{N}_k(0,\Sigma), \tag{22}$$

where $\Sigma=(\sigma^2_{\ell_u,\ell_v})_{1\le u,v\le k}$ is given by

$$\sigma^2_{s,t}=\sum_{k\in\mathbb{Z}}\mathbb{E}(C_0C_sC_kC_{k+t})[\kappa_4(s,k,k+t)+\gamma_X(k+t)\gamma_X(k-s)-\gamma_X(k)\gamma_X(k+t-s)] \tag{23}$$

where $\kappa_4(a,b,c)$ denotes the fourth order cumulant of (X_0,X_a,X_b,X_c).

Remark 6. *Comparison of the assumptions.*

- *We note that if all the $(X_k)_{k\in\mathbb{Z}}$ are observed then $\mathbb{E}(C_0C_\ell C_kC_{k+\ell})=1$ and σ^2_ℓ is the same as in* [23]*, Theorem 3, p. 58.*

- *The assumptions* **(D)** *on the process in presence of missing observations may be relaxed in this Lemma and we respectively only need*

$$\sum_i i^{\frac{4}{r-4}} \cdot \alpha_C(i) < \infty, \qquad \sum_i i^{\frac{2}{r-4}} \cdot \theta_C^{\frac{r-2}{r-1}}(i) < \infty$$

instead of the assumed assumptions

$$\sum_i i^{\frac{8}{r-4}} \cdot \alpha_C(i) < \infty, \qquad \sum_i i^{\frac{8}{r-4}} \cdot \theta_C(i) < \infty$$

The last assumption implies that the sequence $i^{\frac{8}{r-4}}\theta_C(i)$ *is bounded thus* $\theta_C(i) = \mathcal{O}(i^{-\frac{8}{r-4}})$ *which implies* $\sum_i i^{\frac{2}{r-4}}\theta_C^{\frac{r-2}{r-1}}(i) < \infty$. *Assumptions for Lemmas 3 and 1 are thus highly related.*

Proof of Lemma 3. To study the asymptotic behavior of $\widetilde{\gamma}_{X,N}(\ell) = \widehat{\gamma}_{Y,N}(\ell)/\widehat{\nu}(\ell)$ we quote that:

$$W_N^{(\ell)} \equiv \sqrt{N}(\widetilde{\gamma}_{X,N}(\ell) - \gamma_X(\ell)) \quad = \quad \sqrt{N}\frac{\widehat{\delta}_N(\ell)}{\widehat{\nu}_N(\ell)}$$

The denominator term converges in probability to $\nu(\ell)$ from Lemma 2. In order to prove this fidi convergence we consider vectors $(W_N^{(\ell_1)}, \ldots, W_N^{(\ell_k)})$ and arbitrary real numbers $a_1, \ldots, a_m$ and a linear combination of the initial rvs

$$Z_k = \sum_{j=1}^m a_j Z_k^{(\ell_j)}, \qquad Z_k^{(\ell)} = C_k C_{k+\ell}(X_k X_{k+\ell} - \gamma_X(\ell)).$$

We proceed as in Lemma 2 and use the convergence in probability of numerators asserted before and the CLT for the numerators from [10] which results from the convergence

$$\sum_i i^{\frac{4}{r-2}} \alpha_Z(i) < \infty$$

We then apply the heredity arguments provided in the proof of Lemma 2 to conclude.

It remains to show that the limiting covariances are given by

$$\sigma_\ell^2 = \sum_{k \in \mathbb{Z}} \mathbb{E} Z_0^{(\ell)} Z_k^{(\ell)} \tag{24}$$

Check first that

$$\sigma_\ell^2 = \sum_{k \in \mathbb{Z}} \mathbb{E}(C_0 C_\ell C_k C_{k+\ell})[\mathbb{E}(X_0 X_\ell X_k X_{k+\ell}) - 3\gamma_X^2(\ell)] \tag{25}$$

Cumulants are defined for all $(i, j, k) \in \mathbb{Z}^3$, they appear in the previous expression (25) to derive the limit covariances

$$\kappa_4(i, j, k) = \mathbb{E}X_0 X_i X_j X_k - \mathbb{E}X_0 X_i \mathbb{E}X_j X_k - \mathbb{E}X_0 X_j \mathbb{E}X_i X_k - \mathbb{E}X_0 X_k \mathbb{E}X_i X_j$$

Then the following expression, which exists for all finite ℓ, is equivalent to the asymptotic variance in equation (25)

$$\sigma_\ell^2 = \sum_{k \in \mathbb{Z}} \mathbb{E}(C_0 C_\ell C_k C_{k+\ell})[\kappa_4(\ell, k, k+\ell) + \gamma_X(k+\ell)\gamma_X(k-\ell) - \gamma_X^2(\ell)].$$

The convergence of finite dimensional repartitions (22) follows under the same lines as before. $\square$

Remark 7. *From both* [10] *and* [5], *we know that conditions of Lemma 3 entail the weak invariance principle for both dependence frames*

$$\frac{1}{\sqrt{N}}\frac{\sum_{k+1}^{[Nt]}(Y_k-\mathbb{E}Y_k)}{\sum_{k+1}^{[Nt]}(C_k-\mathbb{E}C_k)} \underset{N\to\infty}{\longrightarrow} \sigma\cdot W_t, \quad \text{in the Skorohod space } D([0,1])$$

$$\frac{1}{\sqrt{N}}\frac{\sum_{k+1}^{[Nt]}(Y_kY_{k+\ell}-\mathbb{E}Y_kY_{k+\ell})}{\sum_{k+1}^{[Nt]}(C_kC_{k+\ell}-\mathbb{E}C_kC_{k+\ell})} \underset{N\to\infty}{\longrightarrow} \sigma_\ell\cdot W_t, \quad \text{in the Skorohod space } D([0,1]).$$

Such results yield change-point detection for the mean or for the covariance of X (see [19]*).*

A natural and interesting example that extends central limit theorem above under weak dependence conditions and missing observations is the following:

Corollary 2 (Empirical correlations)**.** *Let $(X_k)_{k\in\mathbb{Z}}$ and $(C_k)_{k\in\mathbb{Z}}$ satisfy the dependence assumptions* **(M), (D)** *hold true. Then,*

$$\sqrt{N}(\widehat{\rho}_{X,N}(\ell)-\rho_X(\ell))_{1\le\ell\le m} \underset{N\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \mathcal{N}_m(0,\Omega)$$

where $\Omega=(\omega^2_{st})$ is given by

$$\omega^2_{s,t}=\sigma^2_{s,t}-\rho_X(s)\sigma^2_{0,t}-\rho_X(t)\sigma^2_{s,0}+\rho_X(s)\rho_X(t)\sigma^2_{0,0}. \tag{26}$$

Proof of Corollary 2. Define $f:\mathbb{R}^*\times\mathbb{R}^\ell\to\mathbb{R}^\ell$ by

$$f(x_0,x_1,\ldots,x_\ell)=(x_1/x_0,\ldots x_\ell/x_0) \tag{27}$$

If $\gamma_X(\cdot)$ is the auto-covariance function of $(X_k)_{k\in\mathbb{Z}}$, then by the multivariate continuous mapping theorem

$$\widehat{\rho}_{X,N}(\ell)=f(\widehat{\gamma}_{X,N}(0),\ldots,\widehat{\gamma}_{X,N}(\ell)) \underset{N\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \mathcal{N}(f(\gamma_X(0),\ldots,\gamma_X(\ell)),N^{-1}D\Sigma D) \tag{28}$$

i.e. $\widehat{\rho}_{X,N}(\ell)$ is asymptotic joint normality, where Σ is defined by (23) and D is the matrix of partial derivatives,

$$D=\gamma_X(0)^{-1}\begin{pmatrix} -\rho(1) & 1 & 0 & \ldots & 0\\ -\rho(2) & 0 & 1 & \ldots & 0\\ \vdots & & & \ddots & \vdots\\ -\rho(h) & 0 & 0 & \ldots & 1\end{pmatrix}.$$

Denoting $\sigma^2_{s,t}$ and $\omega^2_{s,t}$ the current element of Σ and $\Omega=D\Sigma D'$ respectively, we find that

$$\omega^2_{s,t}=\sigma^2_{s,t}-\rho(s)\sigma^2_{0,t}-\rho(t)\sigma^2_{s,0}+\rho(s)\rho(t)\sigma^2_{0,0}.$$

Alternatively, making the decomposition

$$\sqrt{N}(\widehat{\rho}_{X,N}(\ell)-\rho_X(\ell))=\sqrt{N}\left(\frac{\widetilde{\gamma}_{X,N}(\ell)-\rho_X(\ell)\widetilde{\gamma}_{X,N}(0)}{\widetilde{\gamma}_{X,N}(0)}\right)$$

the denominator behaves like $\gamma_X(0)$ and we write

$$N\mathrm{Cov}\big(\widetilde{\gamma}_{X,N}(\ell)-\rho_X(\ell)\widetilde{\gamma}_{X,N}(0),\widetilde{\gamma}_{X,N}(h)-\rho_X(h)\widetilde{\gamma}_{X,N}(0)\big)=T_1-T_2-T_3+T_4$$

with

$$T_1 = N\mathrm{Cov}(\widetilde{\gamma}_{X,N}(\ell),\widetilde{\gamma}_{X,N}(h)), \qquad T_2 = \rho_X(\ell)\cdot N\mathrm{Cov}(\widetilde{\gamma}_{X,N}(0),\widetilde{\gamma}_{X,N}(h)),$$
$$T_3 = \rho_X(h)\cdot N\mathrm{Cov}(\widetilde{\gamma}_{X,N}(0),\widetilde{\gamma}_{X,N}(\ell)), \qquad T_4 = \rho_X(h)\rho_X(\ell)\cdot N\mathrm{Var}(\widetilde{\gamma}_{X,N}(0)).$$

and we apply Lemmas 2 and 3 to get the result.

Therefore we obtain

$$\omega^2_{s,t}=\sigma^2_{s,t}-\rho(s)\sigma^2_{0,t}-\rho(t)\sigma^2_{s,0}+\rho(s)\rho(t)\sigma^2_{0,0}. \qquad \square$$

Proof Theorem 2. Lemma 3 provides us with finite dimensional convergence we thus only need to derive the tightness of the sequence of processes W_N. Lemma 1 proves that Γ, the covariance function of the Gaussian limit process W exists and thus that W is a tight process in $\mathcal{H}_s'$.

We mimic here the tightness argument of lemma 5 in [2]. The flat concentration De Acosta tightness argument is used again. We thus need to derive that

$$\mathbb{E} \sup_{g \in B_L} W_N^2(g) \to 0 \quad \text{as } L \to \infty$$

where $B_L = \{g \in \mathcal{H}_s / \, \|g\|_{\mathcal{H}_s} \le 1, g_\ell = 0, \ell \ge L\}$.

But $W_N^2(g) \le 3(|T_1(g)|^2 + |T_2(g)|^2 + |T_3(g)|^2)$, where the first two (deterministic) terms are bounded above in terms of the Sobolev space $\mathcal{H}_s'$ and the third one is such that

$$\mathbb{E} \sup_{g \in B_L} |T_3(g)|^2 \le N \sum_{|\ell| \ge L} \frac{1}{(1+|\ell|)^{2s}} \mathbb{E}(\widetilde{\gamma}_{X,N}(\ell) - \gamma_X(\ell))^2.$$

The same argument as in Lemma 1 allows to conclude. The limit is here with respect to $L \to \infty$ thus we may assume $L > N/2$, there exists constants $a, b > 0$ such that:

$$\mathbb{E} \sup_{g \in B_L} |T_3(g)|^2 \le aN^{1+\frac{4}{r}} \sum_{\ell \ge L \vee N/2} \frac{1}{(1+|\ell|)^{2s}} \le bN^{1+\frac{4}{r}} (L \vee N/2)^{1-2s} \to_{L \to \infty} 0. \quad \square$$

Acknowledgements. We acknowledge the suggestions, and the efficient and kind help of Eric Moulines (Ecole Polytechnique, Paris).

References

1. Andrews, D. Non strong mixing autoregressive processes. *J. Appl. Probab.*, 21:930–934 (1984).
2. Bardet, J., Doukhan, P., and León, J. Uniform limit theorems for the periodogram of weakly dependent time series and their applications to whittle's estimate. *J. of Time Series Anal.*, 29(1):906–945 (2008).
3. Bondon, P. and Bahamonde, N. Least squares estimation of arch models with missing observations. *Journal of Time Series Analysis*, 33 (6):880–891 (2012).
4. Dahlhaus, R. A likelihood approximation for locally stationary processes. *Annals of Statistics*, 28:1762–1794. (2000).
5. Dedecker, J. and Doukhan, P. A new covariance inequality and applications. *Stochastic Process. Appl.*, 106(1):63–80 (2003).
6. Doukhan, P. *Mixing: Properties and examples*, volume 85 of *Lecture Notes in Statistics*. Springer-Verlag, New York. (1994).
7. Doukhan, P. and Lang, G. Evaluation for moments of a ratio with application to regression estimation. *Bernoulli*, 15(4):1259–1286 (2009).
8. Doukhan, P. and León, J. R. Cumulants for stationary mixing random sequences and applications to empirical spectral density. *Probab. Math. Statist.*, 10(1):11–26 (1989).
9. Doukhan, P. and Louhichi, S. A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.*, 84(2):313–342 (1999).
10. Doukhan, P., Massart, P., and Rio, E. The functional central limit theorem for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 30(1):63–82 (1994).
11. Doukhan, P., Mayo, N., and Truquet, T. Weak dependence, models and some applications. *Metrika*, 69, 199–225 (2009).
12. Doukhan, P., Teyssière, G., and Winant, P. A LARCH(∞) vector valued process. In *Dependence in probability and statistics*, volume 187 of Lecture Notes in Statist., pages 245–258. Springer, New York (2006).
13. Dunsmuir, W. and Robinson, P. M. Asymptotic theory for time series containing missing and amplitude modulated observations. *Sankhyā Ser. A*, 43(3):260–281 (1981a).
14. Dunsmuir, W. and Robinson, P. M. Estimation of time series models in the presence of missing data. *J. Amer. Statist. Assoc.*, 76(375):560–568 (1981b).
15. Efromovich, S. Efficient non-parametric estimation of the spectral density in presence of missing observations. *Journal of Time Series Analysis*, 35:407–427 (2014).
16. Engle, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007.
17. Giraitis, L. and Robinson, P. M. (2001). Whittle estimation of arch models. *Econometric Theory*, 17(3):608–631 (1982).

18. Knight, M. I., Nunes, M. A., and Nason, G. P. Spectral estimation for locally stationary time series with missing observations. *Stat. Comput.*, 22(4):877–895 (2012).
19. Lavielle, M. Detection of multiple changes in a sequence of dependent variables. *Stochastic Processes and Appl.*, 83(1):79–102.
20. Parzen, E. (1963). On spectral analysis with missing observations and amplitude modulation. *Sankhyā Ser. A*, 25:383–392 (1999).
21. Rio, E. *Théorie asymptotique des processus aléatoires faiblement dépendants*. Collection Mathématiques & Applications 31. Springer (2000).
22. Rosenblatt, M. A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA*, 42:43–47.
23. Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Birkhäuser, New York (1956).
24. Straumann, D. *Estimation in conditionally heteroscedastic time series models*, volume 181 of *Lecture Notes in Statistics*. Springer-Verlag, Berlin (2005).
25. Whittle, P. Gaussian estimation in stationary time series. *Bull. Inst. Internat. Statist.*, 39(livraison 2):105–129 (1962).
26. Yajima, Y. and Nishino, H. Estimation of the autocorrelation function of a stationary time series with missing observations. *Sankhyā Ser. A*, 61(2):189–207 (1999).

Instituto de Estadística, Pontificia Universidad Católica de Valparaíso
Current address: Avenida Errazuriz 2734, Valparaíso, Chile
E-mail address: natalia.bahamonde@pucv.cl

UMR 8088 AGM, University Cergy-Pontoise, France
Current address: Mathematics, office E 5.28, UCP site Saint-Martin, 2 Bd. Adolphe Chauvin, 95000 Cergy-Pontoise, France
E-mail address: doukhan@u-cergy.fr

Received 06/09/2016