

## STATIONARITY AGAINST INTEGRATION IN THE AUTOREGRESSIVE PROCESS WITH POLYNOMIAL TREND

BY

FRÉDÉRIC PROÏA (ANGERS)

*Abstract.* We tackle the stationarity issue of an autoregressive path with a polynomial trend, and generalize some aspects of the LMC test, the testing procedure of Leybourne and McCabe. First, we show that it is possible to get the asymptotic distribution of the test statistic under the null hypothesis of trend-stationarity as well as under the alternative of nonstationarity for any polynomial trend of order  $r$ . Then, we explain the reason why the LMC test, and by extension the KPSS test, does not reject the null hypothesis of trend-stationarity, mistakenly, when the random walk is generated by a unit root located at  $-1$ . We also observe it on simulated data and correct the procedure. Finally, we describe some useful stochastic processes that appear in our limiting distributions.

**2010 AMS Mathematics Subject Classification:** Primary: 62M10, 60G10, 62M02; Secondary: 91G70, 91B84.

**Key words and phrases:** LMC test, KPSS test, unit root, stationarity testing procedure, polynomial trend, stochastic nonstationarity, random walk, integrated process, ARIMA process, Donsker's invariance principle, continuous mapping theorem.

### 1. A CONSISTENT TEST FOR A UNIT ROOT

**Notation.** In all the paper, we define  $L$  as the lag operator with the convention that  $L^0 = I$ . In addition,  $k_T = k/T$  is the renormalization of any  $k \in \mathbb{N}$ , and  $\mathbb{I}$  means the indicator function. We will always consider that  $0 < \tau \leq 1$  and that  $[T\tau]$  denotes the integer part of  $T\tau$ . To simplify the notation, we will usually refer to the corresponding vector by removing the implicit subscript on the variable. For example,  $\varepsilon' = (\varepsilon_1 \dots \varepsilon_T)$ , where  $\varepsilon'$  is the transpose of  $\varepsilon$ .

We consider the autoregressive process of order  $p$  on  $\mathbb{Z}$  with a polynomial trend of order  $r$ , driven by a random walk and an additive error. For an observed path of size  $T$ , we investigate the model given, for all  $1 \leq t \leq T$ , by

$$(1.1) \quad \Theta(L)X_t = (\alpha_0 + \alpha_1 t_T + \dots + \alpha_r t_T^r) \mathbb{I}_{\{\kappa \neq 0\}} + S_t^\eta + \varepsilon_t,$$

where, for all  $z \in \mathbb{C}$ ,  $\Theta(z) = 1 - \theta_1 z - \dots - \theta_p z^p$  is an autoregressive polynomial

having all its zeroes outside the unit circle, where, for any  $|\rho| = 1$ ,

$$(1.2) \quad S_t^\eta = \rho S_{t-1}^\eta + \eta_t$$

is a random walk starting from  $S_0^\eta = 0$ , and where  $(\varepsilon_t)$  and  $(\eta_t)$  are uncorrelated white noises of variance  $\sigma_\varepsilon^2 > 0$  and  $\sigma_\eta^2 \geq 0$ , respectively. From now on, white noises are to be interpreted in the strong sense, that is, as sequences of independent and identically distributed random variables. For the sake of simplicity, we consider that  $X_{-p+1} = \dots = X_{-1} = 0$ . We also normalize the known part of the trend, by selecting  $t_T = t/T$ , to simplify the treatment of the projections, as we will see in the technical proofs. The order of the polynomial trend is  $r$ , but we will also take account of the case where no trend is introduced in (1.1). We switch from one situation to another by selecting  $\kappa \neq 0$  or  $\kappa = 0$ . Our objective is to establish a testing procedure for

$$\mathcal{H}_0 : \sigma_\eta^2 = 0 \quad \text{against} \quad \mathcal{H}_1 : \sigma_\eta^2 > 0.$$

One can observe that (1.1) is a trend-stationary process under the null  $\mathcal{H}_0$ , since the process  $(S_t^\eta)$  is almost surely zero, and an integrated process of order one under the alternative  $\mathcal{H}_1$ . Hence, evaluating  $\mathcal{H}_0$  against  $\mathcal{H}_1$  is equivalent to testing stationarity against integration in the stochastic part of the process. In this context, our work is a generalization of the procedure of Leybourne and McCabe [12], shortened the *LMC test* in all the sequel. In their original paper, they propose to make use of the maximum likelihood estimator of  $\theta$  on a given path of size  $T$  and to estimate the trend parameters using a least squares methodology on the residual process. Then, they build a test statistic and establish its behavior under the null hypothesis of stationarity for specific trends (none, constant or linear). Under  $\mathcal{H}_1$ , they show that the test statistic diverges with rate  $T$ , and that it is possible to get its correctly renormalized asymptotic distribution. In the simple case where  $p = 0$ , Nabeya and Tanaka [16] had already investigated the founding principles of this strategy. This restriction seems nevertheless far from the reality of time series since all correlation phenomenon has disappeared. Earlier, the authors of [19], [18] and [11] had already taken an interest in such test statistics for closely related models. The procedure of Kwiatkowski et al. [9], shortened from now on the *KPSS test*, translates any correlation in the residual process, to avoid any preliminary estimation of  $p$  and  $\theta$ . Their test statistic (described later in Remark 1.2) is shown to reach the same asymptotic distribution but, as a long-run variance has to be estimated instead, there is a truncation at a lag  $\ell$  such that  $\ell = \ell(T) \rightarrow \infty$  to ensure consistency, and the divergence under  $\mathcal{H}_1$  occurs with rate  $T/\ell = o(T)$ . One can accordingly expect that the LMC procedure will be more powerful to discriminate  $\mathcal{H}_1$ , and such observations are made in [12]. However, the true value of  $p$  is needed and all flexibility is sacrificed, contrasting with the KPSS procedure. The stationarity of time series being a contemporary issue, it is not surprising to find an abundant literature on empirical studies, anomalies detection or improvements brought to these

strategies: let us mention the papers [24], [13], [17], or [15], [7], [4], [20] and all associated references, without completeness. First we will show that in the context of the LMC test, it is possible to get the asymptotic distribution of the test statistic under  $\mathcal{H}_0$  as well as under  $\mathcal{H}_1$  for any polynomial trend of order  $r$ . Then, we will explain, and observe it on some straightforward simulated data, the reason why the LMC test – and by extension the KPSS test – does not reject the null hypothesis of trend-stationarity, mistakenly, when the random walk is generated by a unit root located at  $-1$ . We have widely been inspired by the calculation methods of [21], [9] and [12], themselves relying on the Donsker's invariance principle and the Mann–Wald's theorem, that we will also recall. Finally, we will describe some useful stochastic processes that appear in our limiting distributions, and we will prove our results.

The case  $|\rho| < 1$  corresponds to a trend-stationary process both under  $\mathcal{H}_0$  and under  $\mathcal{H}_1$ , it is consequently not of interest as part of this paper. By combining (1.1) and (1.2), the model under  $\mathcal{H}_1$  is

$$(1.3) \quad \Theta(L)X_t = (\alpha_0 + \alpha_1 t_T + \dots + \alpha_r t_T^r) \mathbb{I}_{\{\kappa \neq 0\}} + \sum_{k=1}^t \rho^{t-k} \eta_k + \varepsilon_t,$$

where the source of the stochastic nonstationarity of  $(X_t)$  is

$$(1.4) \quad S_t^\eta = \sum_{k=1}^t \rho^{t-k} \eta_k,$$

which is the partial sum process of  $(\eta_t)$  when  $\rho = 1$ . First,

$$\begin{aligned} \Theta(L)(I - \rho L)X_t &= (I - \rho L)(\alpha_0 + \alpha_1 t_T + \dots + \alpha_r t_T^r) \mathbb{I}_{\{\kappa \neq 0\}} \\ &\quad + (I - \rho L)(S_t^\eta + \varepsilon_t) \\ &= (\alpha_0^* + \alpha_1^* t_T + \dots + \alpha_r^* t_T^r) \mathbb{I}_{\{\kappa \neq 0\}} + \eta_t + (I - \rho L)\varepsilon_t, \end{aligned}$$

where  $\alpha_0^*, \alpha_1^*, \dots, \alpha_r^*$  are easily identifiable (e.g.,  $\alpha_r^* = 0$  when  $\rho = 1$ ) and the process  $(\eta_t + (I - \rho L)\varepsilon_t)$  is second-order equivalent in moments to an MA(1) residual, as it is explained in [9]. We obtain the integrated model given, for all  $1 \leq t \leq T$ , by

$$(1.5) \quad \Theta(L)(I - \rho L)X_t = (\alpha_0^* + \alpha_1^* t_T + \dots + \alpha_r^* t_T^r) \mathbb{I}_{\{\kappa \neq 0\}} + (I + \beta L)\xi_t,$$

where  $(\xi_t)$  is a white noise of variance  $\sigma_\xi^2$  depending on the so-called *signal-to-noise ratio*  $\sigma_\eta^2/\sigma_\varepsilon^2$ . For the generating process (1.5), we build a consistent estimator of  $\theta$  (see Remark 1.1 below), and consider the residual process

$$(1.6) \quad \tilde{X}_t = X_t - \tilde{\theta}_1 X_{t-1} - \dots - \tilde{\theta}_p X_{t-p}.$$

Note that under  $\mathcal{H}_1$ ,  $|\beta| < 1$ , implying that the differentiated process is causal and invertible. On the other hand,  $|\beta| = 1$  under  $\mathcal{H}_0$ , and the process is not invertible.

REMARK 1.1. *The consistency of  $\check{\theta}_T$  is a crucial issue of the study. Let  $\Lambda_{r,\rho}(L)$  be the operator defined as*

$$\Lambda_{r,\rho}(L) = (I - \rho L)(I - L)^r.$$

*It follows that*

$$\begin{aligned} \Lambda_{r,\rho}(L) \Theta(L) X_t &= (I - L)^r (\alpha_0^* + \alpha_1^* t_T + \dots + \alpha_r^* t_T^r) \mathbb{I}_{\{\kappa \neq 0\}} \\ &\quad + (I - L)^r (I + \beta L) \xi_t \\ &= \mu^* + \Phi(L) \xi_t, \end{aligned}$$

*where  $\mu^*$  is easily identifiable ( $\mu^* = 0$  when  $\rho = 1$  or  $\kappa = 0$ ) and  $\Phi$  is a moving average polynomial of order  $r + 1$ . Now, let*

$$Y_t = \Lambda_{r,\rho}(L) X_t.$$

*Clearly,  $\Theta(L)Y_t = \mu^* + \Phi(L) \xi_t$ , implying that  $(Y_t)$  is a causal ARMA( $p, r + 1$ ) process having a potentially nonzero intercept. Under  $\mathcal{H}_1$ ,  $\Phi$  has  $r$  unit roots, and its last zero is outside the unit circle (since  $|\beta| < 1$ ). Under  $\mathcal{H}_0$ ,  $\Phi$  has  $r + 1$  unit roots. In both situations, Theorem 2.1 of [23] ensures that a pseudo-MLE is consistent for  $(\theta, \beta)$  treating  $(\xi_t)$  as a Gaussian noise, whereas  $\mu^*$  is easily estimated as an intercept of a stationary ARMA process. Nevertheless, only the AR part of the process is of interest for us, and a faster method is worth considering. The causality of  $\Theta$  implies that there exists a causal representation*

$$Y_t = \Theta^{-1}(L)(\mu^* + \Phi(L) \xi_t) = \nu^* + \sum_{k=0}^{\infty} \psi_k \xi_{t-k}$$

*such that, according to Chapter 7 of [2], the sample autocovariance function  $\hat{\gamma}_T$  of  $(Y_t - \nu^*)$  is a consistent estimator of its autocovariance function  $\gamma_Y$ . Using a Yule–Walker approach, for all  $h \in \{r + 2, \dots, p + r + 1\}$ , we obtain*

$$\gamma_Y(h) = \sum_{k=1}^p \theta_k \gamma_Y(h - k).$$

*Hence, a consistent estimator of  $\theta$  may be obtained via  $\hat{\gamma}_T$ . The selection of  $\rho$  will be widely discussed in Section 3.*

As a result of the previous remark, it makes sense to estimate  $\alpha$  under  $\mathcal{H}_0$  using a least squares methodology in the model given by

$$(1.7) \quad \check{X}_t = (\alpha_0 + \alpha_1 t_T + \dots + \alpha_r t_T^r) \mathbb{I}_{\{\kappa \neq 0\}} + \check{\varepsilon}_t,$$

where  $(\check{\varepsilon}_t)$  is the residual process coming from the estimation of  $\theta$ . A second-order residual set  $(\hat{\varepsilon}_t)$  is then built via

$$(1.8) \quad \hat{\varepsilon}_t = \check{X}_t - (\hat{\alpha}_0 + \hat{\alpha}_1 t_T + \dots + \hat{\alpha}_r t_T^r) \mathbb{I}_{\{\kappa \neq 0\}},$$

where  $\hat{\alpha}_T$  is the least squares estimator of  $\alpha$  in the model (1.7). Let the partial sum processes of  $(\hat{\varepsilon}_t)$  and  $(\hat{\varepsilon}_t^2)$  be defined as

$$(1.9) \quad S_t = \sum_{k=1}^t \hat{\varepsilon}_k \quad \text{and} \quad Q_t = \sum_{k=1}^t \hat{\varepsilon}_k^2.$$

Finally, consider the test statistic

$$(1.10) \quad \hat{K}_T = \frac{1}{TQ_T} \sum_{t=1}^T S_t^2.$$

REMARK 1.2. *The test statistic of the KPSS procedure is very close to  $\hat{K}_T$ . The main difference is that  $(\varepsilon_t)$  satisfies some weaker assumptions including correlation (see [9]), leading to  $p = 0$  and no parameter  $\theta$  to estimate. In return, a long-run variance defined as*

$$\sigma^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}[S_T^2]$$

*has to be estimated using a truncation method. The test statistic is*

$$\hat{K}_T = \frac{1}{T^2 \hat{\sigma}_T^2} \sum_{t=1}^T S_t^2$$

*and corresponds to (1.10) when  $\hat{\sigma}_T^2 = Q_T/T$ , that is, when the long-run variance is estimated as a white noise variance.*

We now establish the asymptotic behavior of  $\hat{K}_T$  under  $\mathcal{H}_0$ . The stochastic processes appearing in our limiting distributions are described in the next section.

THEOREM 1.1. *Assume that  $\sigma_\eta^2 = 0$ . Then, for  $\kappa \neq 0$ , we have the weak convergence*

$$\hat{K}_T \xrightarrow{\mathcal{D}} \int_0^1 B_r^2(s) \, ds,$$

*where  $(B_r(t), t \in [0, 1])$  is the generalized Brownian bridge of order  $r$ . In addition, for  $\kappa = 0$ , we have the weak convergence*

$$\hat{K}_T \xrightarrow{\mathcal{D}} \int_0^1 W^2(s) \, ds,$$

*where  $(W(t), t \in [0, 1])$  is the standard Wiener process.*

In the following theorem, we show that  $\hat{K}_T$  diverges under  $\mathcal{H}_1$  for  $\rho = 1$  with rate  $T$  and we study the asymptotic behavior of the test statistic correctly renormalized. We also show that it decreases to zero under  $\mathcal{H}_1$  for  $\rho = -1$ .

THEOREM 1.2. Assume that  $\sigma_\eta^2 > 0$ . Then, for  $\kappa \neq 0$  and  $\rho = 1$ , we have the weak convergence

$$\frac{\widehat{K}_T}{T} \xrightarrow{\mathcal{D}} \frac{\int_0^1 C_{r,1}^2(s) \, ds}{\int_0^1 W_{r,0}^2(s) \, ds},$$

where  $(C_{r,1}(t), t \in [0, 1])$  is the integrated Brownian bridge of order  $r \times 1$ , and  $(W_{r,0}(t), t \in [0, 1])$  is the detrended Wiener process of order  $r \times 0$ . In addition, for  $\kappa = 0$ , we have the weak convergence

$$\frac{\widehat{K}_T}{T} \xrightarrow{\mathcal{D}} \frac{\int_0^1 (W^{(1)})^2(s) \, ds}{\int_0^1 W^2(s) \, ds},$$

where  $(W^{(1)}(t), t \in [0, 1])$  is the integrated Wiener process of order 1, and  $(W(t), t \in [0, 1])$  is the standard Wiener process. Finally, for  $\rho = -1$ ,

$$\widehat{K}_T \xrightarrow{\mathbb{P}} 0.$$

The situation where  $\rho = -1$  is the cause of a number of complications as we will see in the associated proofs, that is the reason why we limit ourselves to stipulate the convergence of  $\widehat{K}_T$  to zero in the general case. However, in the particular case where  $\kappa = 0$ , we reach the following result.

PROPOSITION 1.1. Assume that  $\sigma_\eta^2 > 0$ . Then, for  $\kappa = 0$  and  $\rho = -1$ , we have the weak convergence

$$T \widehat{K}_T \xrightarrow{\mathcal{D}} \frac{2\sigma_\varepsilon^2 \int_0^1 W_\varepsilon^2(s) \, ds + \sigma_\eta^2 \int_0^1 W_\eta^2(s) \, ds}{2\sigma_\eta^2 \int_0^1 W_\eta^2(s) \, ds},$$

where  $(W_\varepsilon(t), t \in [0, 1])$  and  $(W_\eta(t), t \in [0, 1])$  are independent standard Wiener processes.

One can notice that this is the only situation in which  $(\varepsilon_t)$  and  $(\eta_t)$  simultaneously play a role in the asymptotic behavior; this explains why we had to make such a decomposition into  $W_\varepsilon(t)$  and  $W_\eta(t)$ . As a matter of fact, under  $\mathcal{H}_0$ ,  $(\varepsilon_t)$  is the only perturbing process whereas under  $\mathcal{H}_1$  with  $\rho = 1$ ,  $(\varepsilon_t)$  is dominated by  $(\eta_t)$ . We are pretty convinced, on the basis of a simulation study, that it is possible to find an identifiable limiting distribution for  $T \widehat{K}_T$  when  $\kappa \neq 0$  and  $\rho = -1$ .

However, we have not reached the explicit expression in this work because of complications due to the phenomenon of compensation in the invariance principles, and calculations very hard to conduct. This could form an objective for a future study.

The proofs of Theorems 1.1 and 1.2 and Proposition 1.1 are given in Section 4.

**REMARK 1.3.** *It is also possible to extend the whole results to the multi-integrated processes under the alternative, such as ARI processes having more than one unit root. In the model (1.1), the random walk  $(S_t^\eta)$  is now itself generated by a random walk, and so on up to  $d \geq 0$  positive unit roots. Then, weak convergences in Theorem 1.2 become*

$$\frac{\widehat{K}_T}{T} \xrightarrow{\mathcal{D}} \frac{\int_0^1 C_{r,d}^2(s) ds}{\int_0^1 W_{r,d-1}^2(s) ds} \quad \text{and} \quad \frac{\widehat{K}_T}{T} \xrightarrow{\mathcal{D}} \frac{\int_0^1 (W^{(d)})^2(s) ds}{\int_0^1 (W^{(d-1)})^2(s) ds},$$

respectively for  $\kappa \neq 0$  and  $\kappa = 0$ . For  $d \geq 1$  negative unit roots, we still reach the convergence

$$\widehat{K}_T \xrightarrow{\mathbb{P}} 0.$$

Such results may be useful to produce a statistical testing procedure concerning the integration order  $d$  of the generating process of an observed path and/or to check the true value of  $r$ .

In Figure 1, we have represented the asymptotic distribution of  $\widehat{K}_T$  under  $\mathcal{H}_0$  for  $\kappa = 0$ , then for  $\kappa \neq 0$  and  $r \in \{0, \dots, 4\}$ , using Monte Carlo experiments.

## 2. SOME USEFUL STOCHASTIC PROCESSES

Throughout the study, we deal with some stochastic processes, built from the standard Wiener process  $(W(t), t \in [0, 1])$  that we are now going to introduce. In all definitions, we consider that  $d, r \in \mathbb{N}$ .

**DEFINITION 2.1 (Integrated Wiener Process).** The process given, for  $t \in [0, 1]$ , by the formula

$$W^{(d)}(t) = \int_0^t \int_0^{s_1} \dots \int_0^{s_{d-1}} W(s_d) ds_d \dots ds_1$$

is called an *integrated Wiener process of order  $d$*  in the whole paper. By convention,  $W^{(0)}(t) \equiv W(t)$ .

For example,

$$W^{(1)}(t) = \int_0^t W(s) ds \quad \text{and} \quad W^{(2)}(t) = \int_0^t \int_0^s W(u) du ds.$$

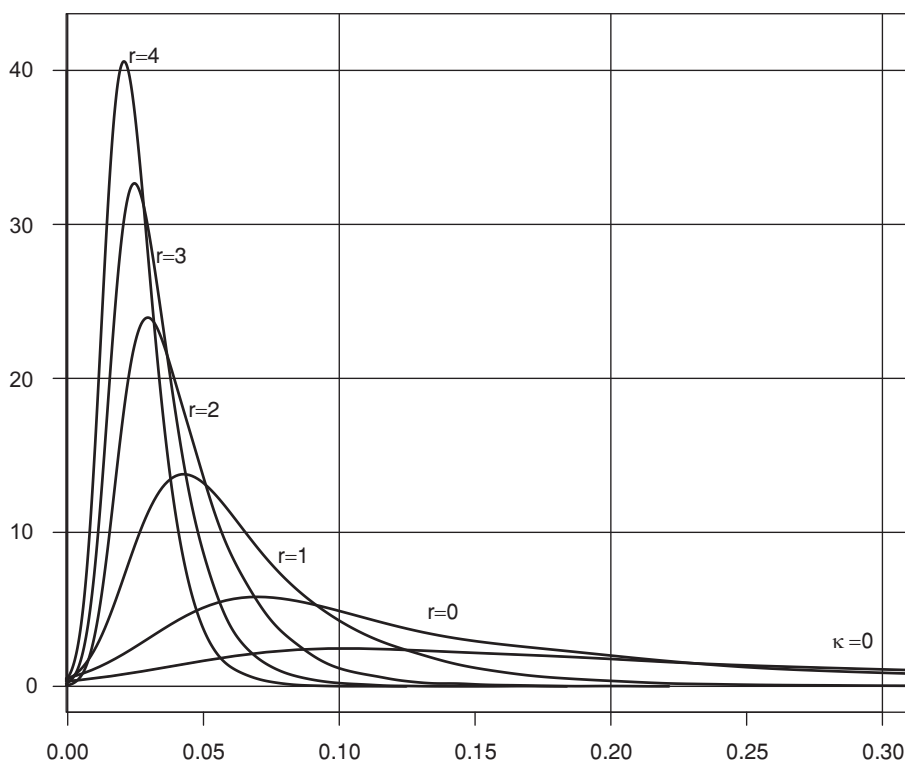


FIGURE 1. Asymptotic distribution of  $\widehat{K}_T$  under  $\mathcal{H}_0$  for  $\kappa = 0$ , then for  $\kappa \neq 0$  and  $r \in \{0, \dots, 4\}$ , using Monte Carlo experiments.

**DEFINITION 2.2 (Generalized Brownian Bridge).** The process given, for  $t \in [0, 1]$ , by

$$B_r(t) = h_r(W)(t),$$

where  $h_r$  is an application from  $C([0, 1])$  into  $C([0, 1])$  given by formula (8) in [14], is called a *generalized Brownian bridge of order  $r$*  in the whole paper.

**DEFINITION 2.3 (Integrated Brownian Bridge).** The process given, for  $t \in [0, 1]$ , by

$$C_{r,d}(t) = h_r(W^{(d)})(t)$$

is called an *integrated Brownian bridge of order  $r \times d$*  in the whole paper. By convention,  $C_{r,0}(t) \equiv B_r(t)$ .

**DEFINITION 2.4 (Detrended Wiener Process).** The process given, for  $t \in [0, 1]$ , by

$$W_{r,d}(t) = \frac{dC_{r,d+1}(t)}{dt}$$

is called a *detrended Wiener process of order  $r \times d$*  in the whole paper. It is explic-



itly defined as

$$W_{r,d}(t) = W^{(d)}(t) - P_d'(1)M^{-1}\Lambda(t),$$

where the nonsingular matrix  $M$  satisfies  $M_{ij} = 1/(i + j - 1)$  for all  $1 \leq i, j \leq r + 1$ ,  $\Lambda(t) = (1 \ t \ \dots \ t^r)'$ , and where

$$(2.1) \quad P_d'(t) = (W^{(d)}(t) \int_0^t s W^{(d-1)}(s) ds \ \dots \ \int_0^t s^r W^{(d-1)}(s) ds).$$

Let us illustrate these definitions on the standard cases  $r = \{0, 1\}$  and  $d = 0$ . According to Definition 2.2 and formula (8) in [14], for  $t \in [0, 1]$ ,

$$B_0(t) = h_0(W)(t) = W(t) - tW(1),$$

which is the usual “Brownian bridge”. It follows from Definitions 2.3 and 2.4 that

$$C_{0,1}(t) = h_0(W^{(1)})(t) = \int_0^t W(s) ds - t \int_0^1 W(s) ds$$

and that

$$W_{0,0}(t) = \frac{dC_{0,1}(t)}{dt} = W(t) - \int_0^1 W(s) ds,$$

which is the usual “demeaned Wiener process”. Similarly, for  $r = 1$ ,

$$B_1(t) = h_1(W)(t) = W(t) + t(2 - 3t)W(1) - 6t(1 - t) \int_0^1 W(s) ds$$

is the “second-level Brownian bridge”, leading to

$$C_{1,1}(t) = \int_0^t W(s) ds + t(3t - 4) \int_0^1 W(s) ds + 6t(1 - t) \int_0^1 s W(s) ds.$$

Finally,

$$W_{1,0}(t) = \frac{dC_{1,1}(t)}{dt} = W(t) + (6t - 4) \int_0^1 W(s) ds + (6 - 12t) \int_0^1 s W(s) ds$$

is the standard “detrended Wiener process”.

### 3. A CORRECTED TEST ADAPTED TO THE NEGATIVE UNIT ROOT

The empirical power of the KPSS and LMC procedures has been widely studied in the literature (see Section 1 for references). For  $\rho = 1$ , the improvements that we described in this paper (for any  $r$  and  $d$ ) are mainly theoretical. On the other hand, we thought useful to conduct an empirical study for  $\rho = -1$ , because in this case it is not only a matter of generalization but also a matter of *correction* of the existing procedures. To motivate the study, we have represented in Figures 2–5 below some examples of simulations according to (1.3) under  $\mathcal{H}_0 : “\sigma_\eta^2 = 0”$ ,

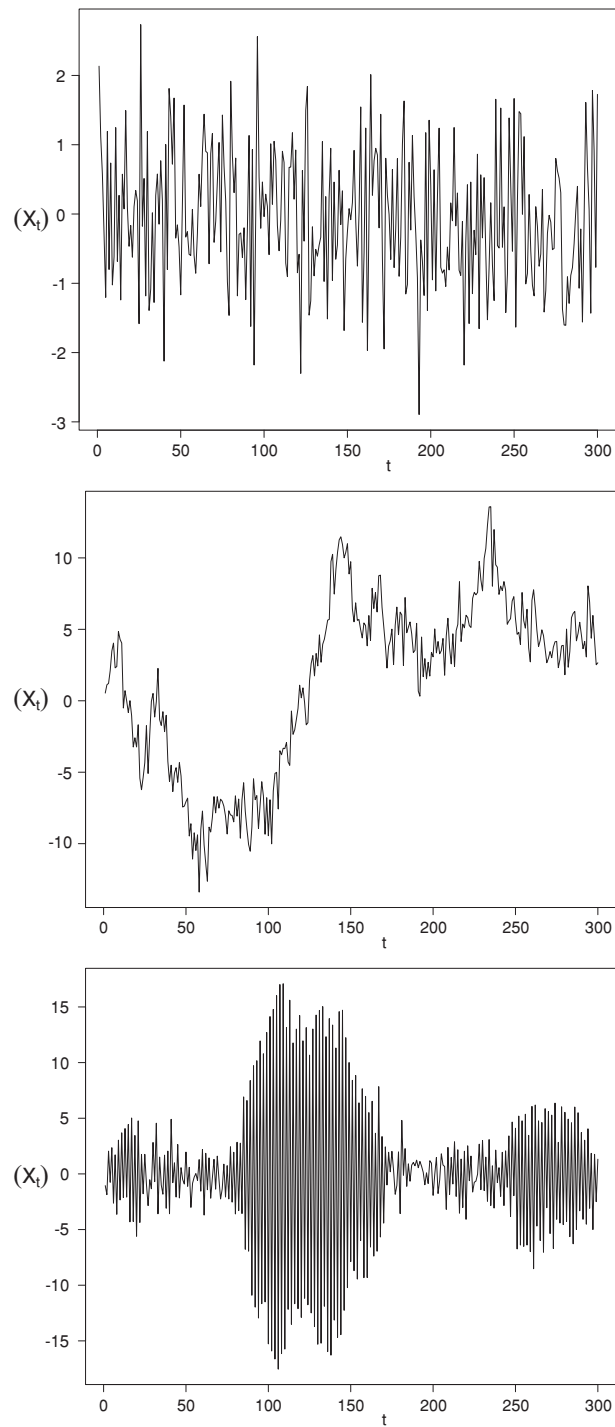


FIGURE 2. Example of simulations under  $\mathcal{H}_0$  (upper), under  $\mathcal{H}_1^+$  (middle) and under  $\mathcal{H}_1^-$  (lower) for  $T = 300, p = 0, \kappa = 0$  and standard Gaussian white noises.

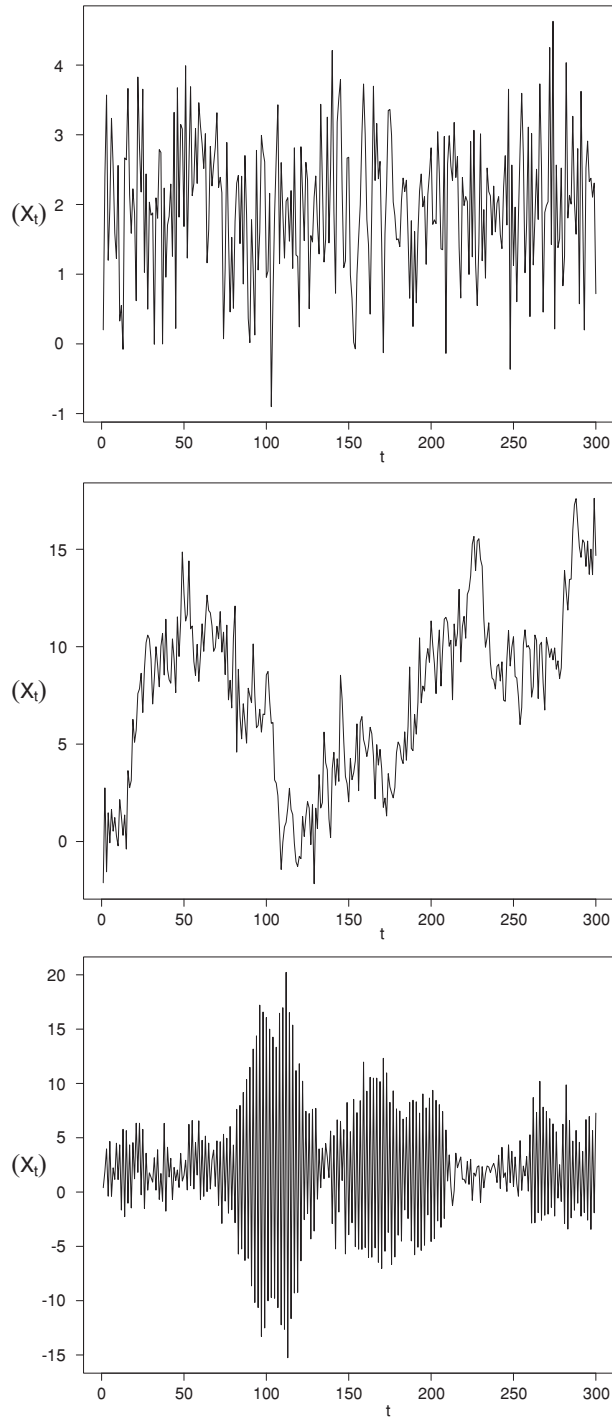


FIGURE 3. Example of simulations under  $\mathcal{H}_0$  (upper), under  $\mathcal{H}_1^+$  (middle) and under  $\mathcal{H}_1^-$  (lower) for  $T = 300$ ,  $p = 0$ ,  $\kappa \neq 0$ ,  $r = 0$  (with  $a_0 = 2$ ) and standard Gaussian white noises.

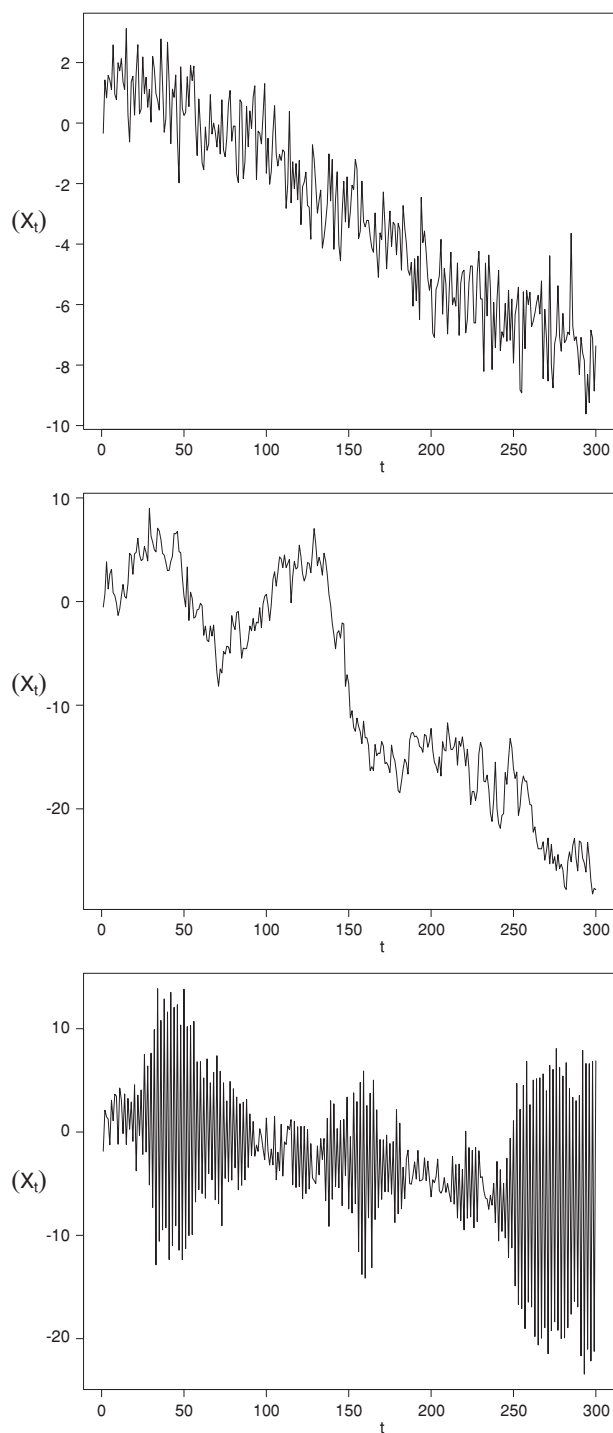


FIGURE 4. Example of simulations under  $\mathcal{H}_0$  (upper), under  $\mathcal{H}_1^+$  (middle) and under  $\mathcal{H}_1^-$  (lower) for  $T = 300$ ,  $p = 0$ ,  $\kappa \neq 0$ ,  $r = 1$  (with  $a_0 = 2$  and  $a_1 = -10$ ) and standard Gaussian white noises.

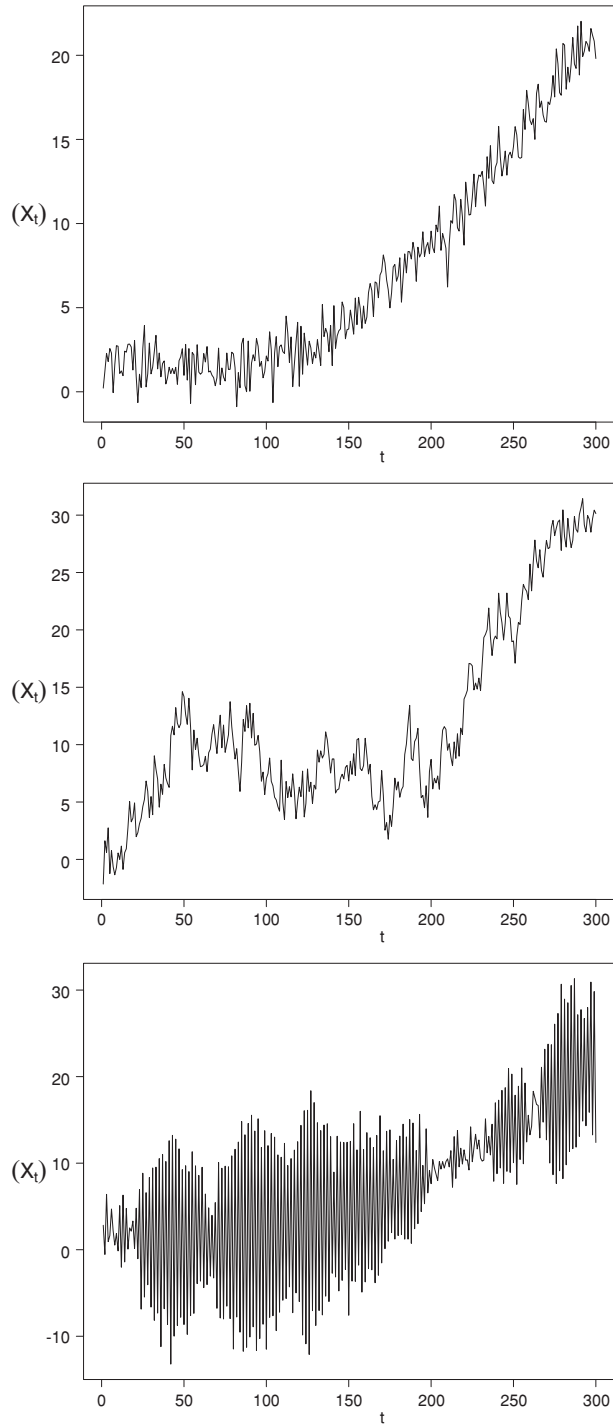


FIGURE 5. Example of simulations under  $\mathcal{H}_0$  (upper), under  $\mathcal{H}_1^+$  (middle) and under  $\mathcal{H}_1^-$  (lower) for  $T = 300$ ,  $p = 0$ ,  $\kappa \neq 0$ ,  $r = 2$  (with  $a_0 = 2$ ,  $a_1 = -10$  and  $a_2 = 30$ ) and standard Gaussian white noises.

under  $\mathcal{H}_1^+$  : “ $\sigma_\eta^2 > 0$  and  $\rho = 1$ ” and under  $\mathcal{H}_1^-$  : “ $\sigma_\eta^2 > 0$  and  $\rho = -1$ ”, using the configurations indicated in the captions. Clearly, a visual investigation is required to decide whether  $\rho = 1$  or  $\rho = -1$  is the most likely alternative, which is a crucial point to stationarize the process. In the whole experiments, the quantiles of the limit distribution of the test statistic under the null hypothesis, depending on  $\kappa$  and  $r$ , have been taken from Table 2 of [14].

The first observation is that, due to the alternation generated by  $\rho = -1$ , it seems quite intuitive to choose between  $\mathcal{H}_1^+$  and  $\mathcal{H}_1^-$  to conduct the test. Besides, it is perceptible on the simulations that heteroscedasticity is manifest. Such high-frequency signals (under  $\mathcal{H}_1^-$ ) are quite unusual in the econometric field, and yet it remains a nonstationary eventuality that a consistent test needs to handle. In the particular case where  $p = 0$ ,  $\kappa = 0$  and where  $(\varepsilon_t)$  and  $(\eta_t)$  are standard Gaussian white noises, we have conducted  $N = 10,000$  simulations, each time testing for stationarity using the KPSS and the LMC procedures. We have obtained the following results (Table 1). On the one hand, we observe that the size of each test is appropriate, since the procedures have been conducted with a significance level  $\alpha = 0.05$ . One also observes that each test is consistent under  $\mathcal{H}_1^+$  but, as one can notice in Table 1, they are misled under  $\mathcal{H}_1^-$  and do not detect this kind of nonstationarity.

TABLE 1. Frequency of rejection of the null hypothesis of stationarity on the basis of  $N = 10,000$  simulations, using the KPSS and LMC procedures.

	KPSS	LMC
Under $\mathcal{H}_0$	0.051	0.051
Under $\mathcal{H}_1^+$	0.989	0.998
Under $\mathcal{H}_1^-$	0.043	0.010

This phenomenon is a direct consequence of Theorem 1.2, in which we have proved that  $\widehat{K}_T$  converges to zero when the unit root of the integrated process is located at  $-1$ . To correct this misuse, we suggest to modify the rejecting rules of the usual procedures depending on whether the alternative is  $\mathcal{H}_1^+$  or  $\mathcal{H}_1^-$ . Let  $k_{r,\alpha}$  be the  $\alpha$ -quantile of the limiting distribution of Theorem 1.1 for a given  $r$ , with the convention that  $k_{r,\alpha} = k_\alpha$  if  $\kappa = 0$ . Then the corrected test takes the following form:

$$C_T = \mathbb{I}_{\{\widehat{K}_T \in \mathcal{R}_\alpha\}} \quad \text{with} \quad \begin{cases} \mathcal{R}_\alpha = ]k_{r,1-\alpha}, +\infty[ & \text{for } \mathcal{H}_0 \text{ vs. } \mathcal{H}_1^+, \\ \mathcal{R}_\alpha = [0, k_{r,\alpha}[ & \text{for } \mathcal{H}_0 \text{ vs. } \mathcal{H}_1^-. \end{cases}$$

Defined as above, the corrected test is exactly the LMC test for  $r \leq 1$  and  $\mathcal{H}_1^+$ , the generalization lies in  $r \geq 2$  and the correction lies in the whole situations under  $\mathcal{H}_1^-$ . In the particular case where  $p = 0$ , it is even possible to build a two-sided test for stationarity,

$$C_T = \mathbb{I}_{\{\widehat{K}_T \in \mathcal{R}_\alpha\}} \quad \text{with } \mathcal{R}_\alpha = [0, k_{r,\alpha/2}[ \cup ]k_{r,1-\alpha/2}, +\infty[,$$

which is adapted to test for  $\mathcal{H}_0$  against  $\mathcal{H}_1 = \mathcal{H}_1^- \cup \mathcal{H}_1^+$ . Figure 6 gives an overview

of the corresponding rejection areas. However, it is crucial to note that for  $p \neq 0$ , it may be problematic to get a consistent estimation of  $\theta$  since we cannot stationarize the process without any information on  $\rho$ . The two-sided procedure is therefore useful only for  $p = 0$ , i.e., in the KPSS framework.

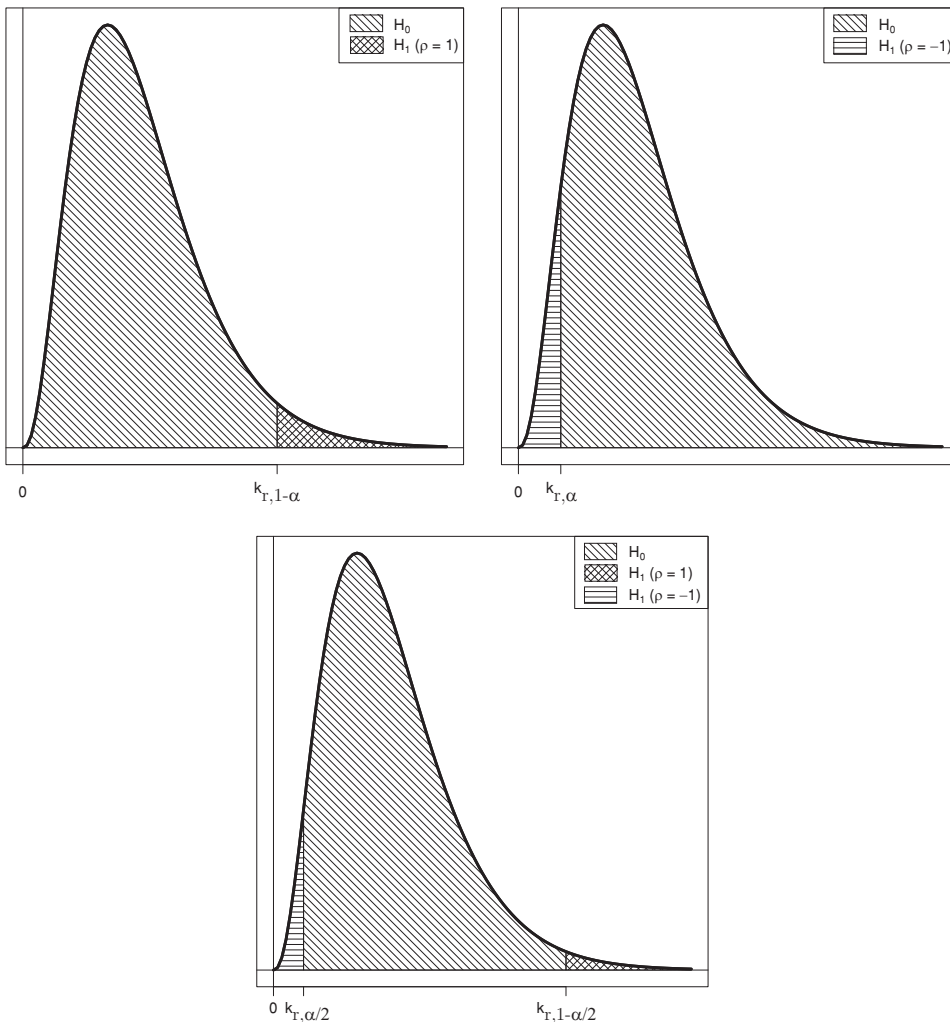


FIGURE 6. Schematic representation of the rejection areas of  $\mathcal{H}_0$  to decide  $\mathcal{H}_1^+$  (top left),  $\mathcal{H}_1^-$  (top right) and  $\mathcal{H}_1^- \cup \mathcal{H}_1^+$  (bottom) for a given significance level  $\alpha$ .

The application of the two-sided corrected test to the dataset used to fill Table 1 leads to 97.6 % of rejection of  $\mathcal{H}_0$ . With no doubt, this is a confirmation that  $\mathcal{H}_1^-$  is now correctly treated. The main corollary of the study is that our results should be rigorously driven to the KPSS procedure. Indeed, on the one hand, it is known that the LMC test suffers from size distortion for a stationary but strongly serially correlated process, as pointed out in [3] or [10] among others, not forgetting that  $p$  is always difficult to properly evaluate in an  $\text{ARMA}(p, q)$  process. On the other

hand, the corrected two-sided test could be conducted without choosing beforehand between  $\mathcal{H}_1^+$  and  $\mathcal{H}_1^-$  as the alternative. Such a test would be fully consistent for testing stationarity of ARMA processes, this is a trail for a future study.

#### 4. PROOF OF THE MAIN RESULTS

We are now going to prove our main results. In all the sequel, we will consider the design matrix  $A$  of order  $(r + 1) \times T$  defined as

$$(4.1) \quad A = \begin{pmatrix} 1 & 1 & \dots & 1 & \dots & 1 \\ 1_T & 2_T & \dots & k_T & \dots & 1 \\ \vdots & \vdots & & \vdots & & \vdots \\ 1_T^r & 2_T^r & \dots & k_T^r & \dots & 1 \end{pmatrix} \quad \text{with } k_T = k/T.$$

The Donsker's invariance principle and the Mann–Wald's continuity theorem being the cornerstone of all our reasonings, we found useful to remind them in this section. In order to lighten the proofs, we deliberately use the same notation for weak convergence and convergence of sequences of random elements in the space of right continuous functions on  $[0, 1]$  having left-hand limits, equipped with the Skorokhod topology (see [1]).

**THEOREM 4.1 (Donsker).** *Assume that  $(Z_T)$  is a sequence of independent and identically distributed random variables having mean zero and finite variance  $\sigma^2 > 0$ . Let  $S_0 = 0$  and  $S_T = Z_1 + \dots + Z_T$ . For a given  $0 < \tau \leq 1$ , let also*

$$S_T^{(\tau)} = \frac{1}{\sigma\sqrt{T}}(S_{[T\tau]} + (T\tau - [T\tau])Z_{[T\tau]+1}).$$

*Then, as  $T$  goes to infinity, we have the weak convergence*

$$S_T^{(\tau)} \xrightarrow{\mathcal{D}} W(\tau),$$

*where  $W(t)$  is the standard Wiener process.*

**THEOREM 4.2 (Mann–Wald).** *Assume that  $(Z_T, Z)$  is a sequence of random elements defined on a metric space  $\mathcal{S}$ . Assume that the application  $h : \mathcal{S} \rightarrow \mathcal{S}'$ , where  $\mathcal{S}'$  is also a metric space, has a set of discontinuity points  $\mathcal{D}_h$  such that  $\mathbb{P}(Z \in \mathcal{D}_h) = 0$ . Then, as  $T$  goes to infinity,*

$$Z_T \rightarrow Z \implies h(Z_T) \rightarrow h(Z).$$

*The implication holds for the convergence in distribution, the convergence in probability and the almost sure convergence.*

**Proof.** The Donsker's invariance principle is described and proved in Section 8 of [1]. The Mann–Wald's continuity theorem, usually called a *continuous mapping theorem*, is for example introduced in Theorem 2.7 of [1] and proved thereafter. ■



In addition, we need to introduce an invariance principle for the residuals of the regression of a random sequence on a polynomial trend in the case where the disturbance has an integrated component. This is an extension of Theorem 1(d) of [25]. For  $\kappa = 0$  but with a more general kind of perturbation, one can also find the foundations of this strategy in [8].

LEMMA 4.1. *Consider, for all  $1 \leq t \leq T$ , the model*

$$X_t = \alpha_0 + \alpha_1 t_T + \dots + \alpha_r t_T^r + S_t^{(d)} + \varepsilon_t$$

with  $d \geq 1$  and  $\kappa \neq 0$ . Let  $\hat{\alpha}_T$  be the least squares estimator of  $\alpha$ , and  $(\hat{\varepsilon}_t)$  the estimated residual set. Then, we have the weak convergence

$$\frac{\hat{\varepsilon}_{[T\tau]}}{\sigma_\eta T^{d-1/2}} \xrightarrow{\mathcal{D}} W_{r,d-1}(\tau),$$

where  $(W_{r,d-1}(t), t \in [0, 1])$  is the detrended Wiener process of order  $r \times (d-1)$ .

**P r o o f.** Recall that  $(S_t^{(d)})$  is a random walk of order  $d$  generated by a white noise sequence  $(\eta_t)$  of variance  $\sigma_\eta^2 > 0$ , that can be defined as

[illegible]

where we assume, to simplify the calculations, that  $S_0^{(1)} = \dots = S_0^{(d)} = 0$ . The least squares estimator of  $\alpha$  is given by

$$(4.3) \quad \hat{\alpha}_T = \left( \sum_{t=1}^T A_t A_t' \right)^{-1} \sum_{t=1}^T A_t X_t = R_T^{-1} \sum_{t=1}^T A_t X_t,$$

where  $A_t$  is the  $t$ -th column of  $A$  given by (4.1). It follows that

$$(4.4) \quad \hat{\alpha}_T - \alpha = R_T^{-1} P_T \quad \text{with } P_T = \sum_{t=1}^T A_t w_t$$

in which we define the residual  $w_t = S_t^{(d)} + \varepsilon_t$ . We start by establishing an invariance principle for  $(w_t)$ . First, Theorem 4.1 is sufficient to get

$$(4.5) \quad \frac{S_{[T\tau]}^{(1)}}{\sigma_\eta \sqrt{T}} = \frac{1}{\sigma_\eta \sqrt{T}} \sum_{t=1}^{[T\tau]} \eta_t \xrightarrow{\mathcal{D}} W(\tau).$$

By extension,

(4.6)

$$\frac{S_{[T\tau]}^{(2)}}{\sigma_\eta T^{3/2}} = \frac{1}{\sigma_\eta T^{3/2}} \sum_{t=1}^{[T\tau]} S_t^{(1)} = \sum_{t=1}^{[T\tau]} \int_{t/T}^{(t+1)/T} \frac{S_{[Ts]}^{(1)}}{\sigma_\eta T^{1/2}} ds \xrightarrow{\mathcal{D}} \int_0^\tau W(s) ds \equiv W^{(1)}(\tau)$$

from Theorem 4.2. Iterating the process, we obtain, for  $d \geq 2$ ,

$$(4.7) \quad \frac{S_{[T\tau]}^{(d)}}{\sigma_\eta T^{d-1/2}} \xrightarrow{\mathcal{D}} \int_0^\tau \int_0^{s_1} \dots \int_0^{s_{d-2}} W(s_{d-1}) ds_{d-1} \dots ds_1 \equiv W^{(d-1)}(\tau).$$

Since  $\varepsilon_{[T\tau]} = o(T^{d-1/2})$  a.s. from the strong law of large numbers, it follows that  $(w_t)$  also satisfies, for all  $d \geq 1$ , the invariance principle given by (4.7). For  $d = 1$ , one can identify the limiting distribution in (4.7) and  $\sigma_\eta$  with  $W$  and  $\sqrt{\omega}$  in Assumption 1(a) of [25]. In addition, the  $k$ -th line of  $P_T$  given in (4.4) is

$$(4.8) \quad P_{k,T} = \sum_{t=1}^T t^{k-1} w_t = \frac{1}{T^{k-1}} \sum_{t=1}^T t^{k-1} w_t.$$

We are now going to study the rate of convergence of  $P_{k,T}$ . For all  $1 \leq i \leq d$ , let us write  $\delta_k(i) = i + k - 1/2$ . We can use (4.7) to get

$$(4.9) \quad \frac{1}{\sigma_\eta T^{\delta_k(d)}} \sum_{t=1}^{[T\tau]} t^{k-1} w_t = \sum_{t=1}^{[T\tau]} \int_{t/T}^{(t+1)/T} \frac{[Ts]^{k-1} w_{[Ts]}}{\sigma_\eta T^{k-1} T^{\delta_0(d)}} ds \xrightarrow{\mathcal{D}} \int_0^\tau s^{k-1} W^{(d-1)}(s) ds.$$

By combining (4.8) and (4.9), we find that, for all  $d \geq 1$ ,

$$(4.10) \quad \frac{P_{[T\tau]}}{\sigma_\eta T^{d+1/2}} \xrightarrow{\mathcal{D}} P_d(\tau),$$

where the limiting distribution is given in (2.1). Moreover, by a direct calculation,

$$(4.11) \quad \lim_{T \rightarrow \infty} \frac{R_T}{T} = M \quad \text{and} \quad \lim_{T \rightarrow \infty} T R_T^{-1} = M^{-1},$$

where  $R_T$  is given in formula (4.3) and the nonsingular matrix  $M$  satisfies  $M_{ij} = 1/(i+j-1)$  for all  $1 \leq i, j \leq r+1$ . It follows from (4.4), (4.10) and (4.11) that

$$(4.12) \quad \frac{\hat{\alpha}_T - \alpha}{\sigma_\eta T^{d-1/2}} \xrightarrow{\mathcal{D}} M^{-1} P_d(1).$$

It only remains to notice that

$$(4.13) \quad \frac{\hat{\varepsilon}_{[T\tau]}}{T^{d-1/2}} = \frac{w_{[T\tau]}}{T^{d-1/2}} - \frac{(\hat{\alpha}_T - \alpha)' A_{[T\tau]}}{T^{d-1/2}}$$

and to combine (4.7) and (4.12) to conclude that, for  $d \geq 1$ ,

$$\frac{\widehat{\varepsilon}_{[T\tau]}}{\sigma_\eta T^{d-1/2}} \xrightarrow{\mathcal{D}} W^{(d-1)}(\tau) - P'_d(1)M^{-1}\Lambda(\tau) \equiv W_{r,d-1}(\tau)$$

from Theorem 4.2, where  $\Lambda(\tau) = (1 \ \tau \ \dots \ \tau^r)'$  is the limiting value of  $A_{[T\tau]}$ . For  $d = 1$ , the latter convergence is given in Theorem 1(d) of [25]. This completes the proof of Lemma 4.1. ■

**Proof of Theorem 1.1.** Denote by  $P = A'(AA')^{-1}A$  the projection matrix and by  $I_T$  the identity matrix of order  $T$ . We start by expressing  $(\widehat{\varepsilon}_t)$  in terms of  $(\varepsilon_t)$  to establish an invariance principle such as Theorem 4.1 on  $(S_t)$  given by (1.9). We first consider the general case where  $\kappa \neq 0$ . By (1.6) and (1.8), since  $\widehat{\alpha}_T$  is the least squares estimator of  $\alpha$ , a direct calculation shows that, for all  $1 \leq t \leq T$ ,

$$(4.14) \quad \widehat{\varepsilon}_t = \check{X}_t - \widehat{\alpha}_0 - \widehat{\alpha}_1 t_T - \dots - \widehat{\alpha}_r t_T^r = \sum_{i=1}^p (\theta_i - \check{\theta}_i) u_{i,t} + u_t,$$

where  $u_t$  is the  $t$ -th component of  $(I_T - P)\varepsilon$ , and, for  $1 \leq i \leq p$ ,  $u_{i,t}$  is the  $t$ -th component of  $(I_T - P)X_{-i}$  with  $X'_{-i} = (X_{1-i} \ \dots \ X_{T-i})$ . From Theorem 1 of [14], we have the weak convergence

$$(4.15) \quad \frac{1}{\sigma_\varepsilon \sqrt{T}} \sum_{t=1}^{[T\tau]} u_t \xrightarrow{\mathcal{D}} B_r(\tau).$$

In addition, for any  $1 \leq i \leq p$  and since  $\Theta$  is causal, equation (1.1) leads to

$$(4.16) \quad X_{t-i} = \Theta^{-1}(L)(\alpha_0 + \alpha_1(t-i)_T + \dots + \alpha_r(t-i)_T^r) + \mu_{t-i},$$

where  $(t-i)_T = (t-i)/T$  and  $\Theta(L)\mu_{t-i} = \varepsilon_{t-i}$ . The coefficients of the deterministic trend are identifiable *via* a tedious but straightforward calculation. It follows from (4.16) that  $(\mu_t)$  is a stable stationary  $\text{AR}(p)$  process which also satisfies an invariance principle, as it is stipulated, for example, in Theorem 1 of [5]. If we define the so-called long-run variance as

$$\sigma_\mu^2 = \mathbb{E}[\mu_0^2] + 2 \sum_{k=1}^{\infty} \mathbb{E}[\mu_0 \mu_k],$$

which is finite for a stable AR process (see Chapter 3 of [2]), then, for all  $1 \leq i \leq p$ ,

$$(4.17) \quad \frac{1}{\sigma_\mu \sqrt{T}} \sum_{t=1}^{[T\tau]} u_{i,t} \xrightarrow{\mathcal{D}} B_r(\tau),$$

by using again Theorem 1 of [5]. Convergence (4.17) and the consistency of  $\check{\theta}_T$  imply that

$$(4.18) \quad \frac{1}{\sigma_\mu \sqrt{T}} \sum_{i=1}^p (\theta_i - \check{\theta}_i) \sum_{t=1}^{[T\tau]} u_{i,t} \xrightarrow{\mathbb{P}} 0.$$

Noticing that  $(S_t)$  in (1.9) is the partial sum process of  $(\widehat{\varepsilon}_t)$ , we obtain

$$(4.19) \quad \frac{S_{[T\tau]}}{\sigma_\varepsilon \sqrt{T}} \xrightarrow{\mathcal{D}} B_r(\tau).$$

In addition, it is not hard to see that

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u_t^2 = \sigma_\varepsilon^2 \quad \text{a.s.}$$

since  $(u_t)$  can be seen as the residual of the regression of  $(\varepsilon_t)$  on a polynomial time trend with zero coefficients. The same kind of convergence can be reached for  $(u_{i,t})$  following a similar methodology to that in [22], since  $(u_{i,t})$  can be seen as the residual of the regression of a weakly stationary process  $(\mu_t)$  on a polynomial time trend also with zero coefficients. Hence, by the Cauchy–Schwarz inequality,

$$(4.20) \quad \lim_{T \rightarrow \infty} \frac{Q_T}{T} = \sigma_\varepsilon^2 \quad \text{a.s.},$$

where the process  $(Q_t)$  is given by (1.9). Finally,

$$\frac{1}{\sigma_\varepsilon^2 T^2} \sum_{t=1}^{[T\tau]} S_t^2 = \frac{1}{T} \sum_{t=1}^{[T\tau]} \left( \frac{S_t}{\sigma_\varepsilon \sqrt{T}} \right)^2 = \sum_{t=1}^{[T\tau]} \int_{t/T}^{(t+1)/T} \left( \frac{S_{[Ts]}}{\sigma_\varepsilon \sqrt{T}} \right)^2 ds \xrightarrow{\mathcal{D}} \int_0^\tau B_r^2(s) ds$$

by application of Theorem 4.2. This achieves the proof of Theorem 1.1, by using (4.19), (4.20), Slutsky's lemma and taking  $\tau = 1$ , in the case where there is a polynomial trend. On the other hand, for  $\kappa = 0$ ,  $P$  is the zero matrix and we merely have  $u_t = \varepsilon_t$  and  $u_{i,t} = X_{-i}$  in (4.14) for all  $1 \leq t \leq T$  and  $1 \leq i \leq p$ . Then, convergence (4.20) follows from the strong law of large numbers and, by Theorem 4.1, the invariance principle (4.19) becomes

$$(4.21) \quad \frac{S_{[T\tau]}}{\sigma_\varepsilon \sqrt{T}} \xrightarrow{\mathcal{D}} W(\tau).$$

The end of the proof follows the same reasoning as above. ■

**Proof of Theorem 1.2.** We now suppose that  $\sigma_\eta^2 > 0$ , implying that the process has a stochastic nonstationarity generated by the random walk  $(S_t^\eta)$

given by (1.4). We first consider the general case  $\kappa \neq 0$ . In the same way as for (4.14), we obtain

$$(4.22) \quad \hat{\varepsilon}_t = \check{X}_t - \hat{\alpha}_0 - \hat{\alpha}_1 t_T - \dots - \hat{\alpha}_r t_T^r = \sum_{i=1}^p (\theta_i - \check{\theta}_i) u_{i,t} + u_{\eta,t},$$

where  $u_{\eta,t}$  is the  $t$ -th component of  $(I_T - P)(S^\eta + \varepsilon)$ . In addition, for all  $1 \leq i \leq p$ ,  $u_{i,t}$  is the  $t$ -th component of  $(I_T - P)X_{-i}$ , and  $X_{-i}$  is given, for all  $1 \leq t \leq T$ , by

$$(4.23) \quad X_{t-i} = \Theta^{-1}(L)(\alpha_0 + \alpha_1(t-i)_T + \dots + \alpha_r(t-i)_T^r) + T_{t-i}^\eta,$$

and  $\Theta(L)T_{t-i}^\eta = S_{t-i}^\eta + \varepsilon_{t-i}$ , with the notation of (4.16). Hence,  $((I - \rho L)T_{t-i}^\eta)$  is second-order equivalent in moments to a stationary ARMA( $p, 1$ ) process. By Theorem 1 of [5], it satisfies an invariance principle in which its long-run variance is involved, and the rate is  $\sqrt{T}$ . Then, by Theorem 4.2 and standard calculations, one can see that  $(u_{i,t})$  behaves like  $(u_{\eta,t})$  since all invariance principles on  $(u_{\eta,t})$  can also be established on  $(u_{i,t})$ . However, as  $\check{\theta}_T$  is consistent, it appears that all asymptotic results will only be driven by  $(u_{\eta,t})$ ,  $(u_{\eta,t}^2)$  and their partial sum processes. First, by Theorem 4.1, in the case where  $\rho = 1$ , we have already seen in (4.5) that we have the invariance principle

$$(4.24) \quad \frac{S_{[T\tau]}^\eta}{\sigma_\eta \sqrt{T}} \xrightarrow{\mathcal{D}} W(\tau).$$

For  $\rho = -1$ , one cannot directly apply Theorem 4.1 since  $(S_t^\eta)$  is not built from identically distributed random variables. However, convergence (4.24) still holds by using, for example, Theorem 1 of [5]. Depending on the value of  $\rho$ , the end of the proof is totally different. On the one hand, for  $\rho = 1$ , from Lemma 4.1 with  $d = 1$ , we have the weak convergence

$$(4.25) \quad \frac{u_{\eta,[T\tau]}}{\sigma_\eta \sqrt{T}} \xrightarrow{\mathcal{D}} W_{r,0}(\tau).$$

It follows that

$$(4.26) \quad \frac{1}{\sigma_\eta T^{3/2}} \sum_{t=1}^{[T\tau]} u_{\eta,t} = \sum_{t=1}^{[T\tau]} \int_{t/T}^{(t+1)/T} \frac{u_{\eta,[Ts]}}{\sigma_\eta \sqrt{T}} ds \xrightarrow{\mathcal{D}} \int_0^\tau W_{r,0}(s) ds \equiv C_{r,1}(\tau)$$

by application of Theorem 4.2. Since the leading term of  $\hat{\varepsilon}_t$  is  $u_{\eta,t}$  as it is explained above and using convergence (4.25), we get an invariance principle for the partial sum process  $(S_t)$  in (1.9), given by

$$(4.27) \quad \frac{S_{[T\tau]}}{\sigma_\eta T^{3/2}} \xrightarrow{\mathcal{D}} C_{r,1}(\tau).$$

We can also reach the same convergence by using Theorem 1 of [14] combined with convergence (4.6), that is,

$$(4.28) \quad \frac{1}{\sigma_\eta T^{3/2}} \sum_{t=1}^{[T\tau]} S_t^\eta = \sum_{t=1}^{[T\tau]} \int_{t/T}^{(t+1)/T} \frac{S_{[Ts]}^\eta}{\sigma_\eta \sqrt{T}} ds \xrightarrow{\mathcal{D}} \int_0^\tau W(s) ds \equiv W^{(1)}(\tau).$$

Of course, (4.20) cannot hold under  $\mathcal{H}_1$ , and the asymptotic behavior of  $Q_T$  will now stem from (4.25). Indeed,

$$\frac{1}{\sigma_\eta^2 T^2} \sum_{t=1}^{[T\tau]} u_{\eta,t}^2 = \sum_{t=1}^{[T\tau]} \int_{t/T}^{(t+1)/T} \left( \frac{u_{\eta,[Ts]}}{\sigma_\eta \sqrt{T}} \right)^2 ds \xrightarrow{\mathcal{D}} \int_0^\tau W_{r,0}^2(s) ds,$$

implying that

$$(4.29) \quad \frac{Q_{[T\tau]}}{\sigma_\eta^2 T^2} \xrightarrow{\mathcal{D}} \int_0^\tau W_{r,0}^2(s) ds.$$

In addition, from (4.27),

$$\frac{1}{\sigma_\eta^2 T^4} \sum_{t=1}^{[T\tau]} S_t^2 = \frac{1}{T} \sum_{t=1}^{[T\tau]} \left( \frac{S_t}{\sigma_\eta T^{3/2}} \right)^2 = \sum_{t=1}^{[T\tau]} \int_{t/T}^{(t+1)/T} \left( \frac{S_{[Ts]}}{\sigma_\eta T^{3/2}} \right)^2 ds \xrightarrow{\mathcal{D}} \int_0^\tau C_{r,1}^2(s) ds.$$

The latter convergence together with (4.29) and Theorem 4.2 achieve the first part of the proof, by selecting  $\tau = 1$ . On the other hand, for  $\rho = -1$ , the summation (4.28) is different due to the phenomenon of compensation. As a matter of fact, it is not hard to see that, for any even and odd integer  $t \geq 1$ , respectively, we have

$$\sum_{k=1}^t S_k^\eta = \sum_{k=1}^{t/2} \eta_{2k} \quad \text{and} \quad \sum_{k=1}^t S_k^\eta = \sum_{k=1}^{(t+1)/2} \eta_{2k-1}.$$

Let  $(\zeta_t)$  be the sequence defined, for an even  $T$  and all  $1 \leq t \leq T/2$ , by

$$\zeta_t = \varepsilon_{2t-1} + \varepsilon_{2t} + \eta_{2t}$$

and, for an odd  $T$  and all  $1 \leq t \leq (T+1)/2$ , by

$$\zeta_t = \varepsilon_{2t-1} + \varepsilon_{2(t-1)} + \eta_{2t-1}.$$

Hence,  $\mathbb{E}[\zeta_t] = 0$ ,  $\mathbb{E}[\zeta_t^2] = 2\sigma_\varepsilon^2 + \sigma_\eta^2$  and all covariances are zero, since  $(\varepsilon_t)$  and  $(\eta_t)$  are not correlated. It follows that  $(\zeta_t)$  is a white noise and that it satisfies, by virtue of Theorem 1 of [5], the invariance principle

$$(4.30) \quad \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\tau]} \zeta_t \xrightarrow{\mathcal{D}} \sqrt{2\sigma_\varepsilon^2 + \sigma_\eta^2} W(\tau).$$

Thus, we obtain the invariance principles

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\tau]} (S_t^\eta + \varepsilon_t) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\tau/2]} \zeta_t \xrightarrow{\mathcal{D}} \sqrt{2\sigma_\varepsilon^2 + \sigma_\eta^2} W\left(\frac{\tau}{2}\right) \stackrel{\mathcal{D}}{=} \sqrt{\frac{2\sigma_\varepsilon^2 + \sigma_\eta^2}{2}} W(\tau)$$

and, by application of Theorem 1 of [14],

$$(4.31) \quad \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\tau]} u_{\eta,t} \xrightarrow{\mathcal{D}} \sqrt{\frac{2\sigma_\varepsilon^2 + \sigma_\eta^2}{2}} B_r(\tau).$$

Exploiting the latter convergence and the domination of  $u_{\eta,t}$  in  $\widehat{\varepsilon}_t$  (the estimator of  $\theta$  remaining consistent), we obtain

$$(4.32) \quad \frac{1}{T^2} \sum_{t=1}^{[T\tau]} S_t^2 = \sum_{t=1}^{[T\tau]} \int_{t/T}^{(t+1)/T} \left( \frac{S_{[Ts]}}{\sqrt{T}} \right)^2 ds \xrightarrow{\mathcal{D}} \frac{2\sigma_\varepsilon^2 + \sigma_\eta^2}{2} \int_0^\tau B_r^2(s) ds.$$

Restart now the reasoning developed in Lemma 4.1, but for  $d = 1$  and  $\rho = -1$ . We recall that, using the notation associated with (4.8), for all  $1 \leq k \leq r + 1$ ,

$$P_{k,T} = \sum_{t=1}^T t_T^{k-1} w_t = \frac{1}{T^{k-1}} \sum_{t=1}^T t^{k-1} (S_t^\eta + \varepsilon_t).$$

First, it is not hard to see that

$$M_T^k = \sum_{t=1}^T t^{k-1} \varepsilon_t$$

is a martingale adapted to the natural filtration of  $(\varepsilon_t)$ , whose increasing process is such that  $\langle M^k \rangle_T = O(T^{2k-1})$  a.s. with obviously

$$\lim_{T \rightarrow \infty} \langle M^k \rangle_T = +\infty \quad \text{a.s.}$$

The law of large numbers for scalar martingales (see, e.g., [6]) implies that  $M_T^k = o(T^k)$  a.s. Hence,

$$(4.33) \quad \frac{P_{k,T}}{T} = \frac{1}{T^k} \sum_{t=1}^T t^{k-1} S_t^\eta + o(1) \quad \text{a.s.}$$

In addition, denote by  $(\Sigma_t^\eta)$  the partial sum process associated with  $(\eta_t)$  for  $\rho = 1$ . Let also  $(\Lambda_t^\eta)$  and  $(\Pi_t^\eta)$  be the partial sum processes associated with  $(\eta_t)$ , for the even and odd subscripts, respectively. Explicitly,

$$\Lambda_{p_t}^\eta = \eta_2 + \eta_4 + \dots + \eta_{2p_t} = \sum_{\ell=1}^{p_t} \eta_{2\ell}$$

and

$$\Pi_{i_t}^\eta = \eta_1 + \eta_3 + \dots + \eta_{2i_t-1} = \sum_{\ell=1}^{i_t} \eta_{2\ell-1}$$

with  $i_t = [(t+1)/2]$  and  $p_t = t - [(t+1)/2]$ . A direct calculation shows that, for  $\rho = -1$  and all  $1 \leq k \leq r+1$ ,

$$(4.34) \quad \sum_{t=1}^T t^{k-1} S_t^\eta = \sum_{t=1}^T t^{k-1} \Sigma_t^\eta - 2 \sum_{t=1}^{p_T} (2t+1)^{k-1} \Lambda_t^\eta - 2 \sum_{t=1}^{i_T} (2t)^{k-1} \Pi_t^\eta + 2r_T,$$

where we have  $r_T = (T+1)^{k-1} \Pi_{(T+1)/2}^\eta$  for all odd  $T$  and  $r_T = (T+1)^{k-1} \Lambda_{T/2}^\eta$  for all even  $T$ . It is possible, *via* Theorem 4.1, to establish an invariance principle on the processes  $(\Lambda_t^\eta)$  and  $(\Pi_t^\eta)$ . As a matter of fact,

$$(4.35) \quad \frac{\Lambda_{[p_T \tau]}^\eta}{\sigma_\eta \sqrt{p_T}} \xrightarrow{\mathcal{D}} W(\tau) \quad \text{and} \quad \frac{\Pi_{[i_T \tau]}^\eta}{\sigma_\eta \sqrt{i_T}} \xrightarrow{\mathcal{D}} W(\tau).$$

It follows from Theorem 4.2 that

$$(4.36) \quad \frac{1}{\sigma_\eta p_T^{k+1/2}} \sum_{t=1}^{[p_T \tau]} (2t+1)^{k-1} \Lambda_t^\eta = \sum_{t=1}^{[p_T \tau]} \int_{t/p_T}^{(t+1)/p_T} \frac{(2[p_T s]+1)^{k-1} \Lambda_{[p_T s]}^\eta}{\sigma_\eta p_T^{k-1} \sqrt{p_T}} ds \\ \xrightarrow{\mathcal{D}} \int_0^\tau (2s)^{k-1} W(s) ds$$

and that

$$(4.37) \quad \frac{1}{\sigma_\eta i_T^{k+1/2}} \sum_{t=1}^{[i_T \tau]} (2t)^{k-1} \Pi_t^\eta = \sum_{t=1}^{[i_T \tau]} \int_{t/i_T}^{(t+1)/i_T} \frac{(2[i_T s])^{k-1} \Pi_{[i_T s]}^\eta}{\sigma_\eta i_T^{k-1} \sqrt{i_T}} ds \\ \xrightarrow{\mathcal{D}} \int_0^\tau (2s)^{k-1} W(s) ds$$

since it is not hard to see that  $p_T$  and  $i_T$  behave like  $T/2$ . Moreover, the convergences (4.35) and the definition of  $r_T$  directly lead to

$$(4.38) \quad \frac{r_T}{T^{k+1/2}} \xrightarrow{\mathbb{P}} 0.$$

In addition, the invariance principle (4.9) for  $\rho = 1$  and  $d = 1$ , here corresponding to the one associated with  $(\Sigma_t^\eta)$ , gives, together with (4.34) and (4.36)–(4.38),

$$\frac{1}{T^{k+1/2}} \sum_{t=1}^T t^{k-1} S_t^\eta = O_{\mathbb{P}}(1)$$



and thus, with the notation above, for all  $1 \leq k \leq r + 1$ , we obtain

$$\frac{P_{k,T}}{T^{3/2}} = O_{\mathbb{P}}(1) \quad \text{and} \quad \frac{u_{\eta,T}}{\sqrt{T}} = \frac{S_T^\eta + \varepsilon_T}{\sqrt{T}} + O_{\mathbb{P}}(1),$$

successively using (4.4) and (4.13). By virtue of Theorems 4.1, 4.2 and the strong law of large numbers, we deduce, following the same calculations, that the process  $(Q_t)$  grows with rate  $T^2$ , which achieves the proof for  $\rho = -1$  since (4.32) shows that the numerator of  $\widehat{K}_T$  also grows with the same rate. Finally, for  $\kappa = 0$ , the invariance principle (4.25) merely becomes

$$(4.39) \quad \frac{u_{\eta,[T\tau]}}{\sigma_\eta \sqrt{T}} \xrightarrow{\mathcal{D}} W(\tau)$$

from Theorem 4.1, and the end of the reasoning easily follows as above. ■

**Proof of Proposition 1.1.** This proof will be very succinct since all results have been established in the previous reasonings. Indeed, for  $\kappa = 0$  and  $\rho = -1$ , convergence (4.32) becomes

$$\frac{1}{T^2} \sum_{t=1}^{[T\tau]} S_t^2 \xrightarrow{\mathcal{D}} \sigma_\varepsilon^2 \int_0^\tau W_\varepsilon^2(s) ds + \frac{\sigma_\eta^2}{2} \int_0^\tau W_\eta^2(s) ds,$$

if we split the limiting distribution into two independent components, so as to easily deal with in the sequel. Without any trend fitted, we also have  $u_{\eta,t} = S_t^\eta + \varepsilon_t$  for all  $1 \leq t \leq T$ . It follows that, similarly,

$$\frac{Q_{[T\tau]}}{T^2} \xrightarrow{\mathcal{D}} \sigma_\eta^2 \int_0^\tau W_\eta^2(s) ds.$$

We complete the proof by choosing  $\tau = 1$  and applying Theorem 4.2. ■

**Acknowledgments.** The author thanks the anonymous referee for his suggestions and constructive comments which helped to improve substantially the paper.

## REFERENCES

- [1] P. Billingsley, *Convergence of Probability Measures*, Wiley Ser. Probab. Stat., New York 1999.
- [2] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, Springer, New York 1996.
- [3] M. Caner and L. Kilian, *Size distortions of tests of the null hypothesis of stationarity: Evidence and implications for the PPP debate*, J. Int. Money Finance 20 (2001), pp. 639–657.
- [4] R. M. De Jong, C. Amsler, and P. Schmidt, *A robust version of the KPSS test, based on indicators*, J. Econometrics 137 (2) (2007), pp. 311–333.
- [5] J. Dedecker and E. Rio, *On the functional central limit theorem for stationary processes*, Ann. Inst. Henri Poincaré B 36 (1) (2000), pp. 1–34.
- [6] M. Duflo, *Random Iterative Models*, Springer, New York 1997.

- [7] D. Harris, S. J. Leybourne, and B. P. M. McCabe, *Modified KPSS tests for near integration*, *Econometric Theory* 23 (2) (2007), pp. 355–363.
- [8] R. Ibragimov and P. C. B. Phillips, *Regression asymptotics using martingale convergence methods*, *Econometric Theory* 24 (4) (2008), pp. 888–947.
- [9] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin, *Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?*, *J. Econometrics* 54 (1992), pp. 159–178.
- [10] M. Lanne and P. Saikkonen, *Reducing size distortions of parametric stationarity tests*, *J. Time Series Anal.* 24 (2003), pp. 423–439.
- [11] S. J. Leybourne and B. P. M. McCabe, *On the distribution of some test statistics for parameter constancy*, *Biometrika* 76 (1989), pp. 167–177.
- [12] S. J. Leybourne and B. P. M. McCabe, *A consistent test for a unit root*, *J. Bus. Econom. Statist.* 12 (2) (1994), pp. 157–166.
- [13] S. J. Leybourne and B. P. M. McCabe, *Modified stationarity tests with data-dependent model-selection rules*, *J. Bus. Econom. Statist.* 17 (2) (1999), pp. 264–270.
- [14] I. B. MacNeill, *Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times*, *Ann. Statist.* 6 (2) (1978), pp. 422–433.
- [15] U. Müller, *Size and power of tests of stationarity in highly autocorrelated time series*, *J. Econometrics* 128 (2) (2005), pp. 195–213.
- [16] S. Nabeya and K. Tanaka, *Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative*, *Ann. Statist.* 16 (1) (1988), pp. 218–235.
- [17] P. Newbold, S. J. Leybourne, and M. E. Wohar, *Trend-stationarity, difference-stationarity, or neither: Further diagnostic tests with an application to U.S. Real GNP*, *J. Econ. Bus.* 53 (1) (2001), pp. 85–102.
- [18] J. Nyblom, *Testing for deterministic linear trend in time series*, *J. Amer. Statist. Assoc.* 81 (1986), pp. 545–549.
- [19] J. Nyblom and T. Makelainen, *Comparisons of tests for the presence of random walk coefficients in a simple linear model*, *J. Amer. Statist. Assoc.* 78 (1983), pp. 856–864.
- [20] M. M. Pelagatti and P. K. Sen, *A robust version of the KPSS test based on ranks*, Working Papers from Università degli Studi di Milano-Bicocca, Dipartimento di Statistica, No. 2009070 (2009).
- [21] P. C. B. Phillips, *Time series regression with a unit root*, *Econometrica* 55 (2) (1987), pp. 277–301.
- [22] P. C. B. Phillips and P. Perron, *Testing for a unit root in time series regression*, *Biometrika* 75 (2) (1988), pp. 335–346.
- [23] B. M. Pötscher, *Noninvertibility and pseudo-maximum likelihood estimation of misspecified ARMA models*, *Econometric Theory* 7 (1991), pp. 435–449.
- [24] P. Saikkonen and R. Luukkonen, *Testing for a moving average unit root in autoregressive integrated moving average models*, *J. Amer. Statist. Assoc.* 88 (1993), pp. 596–601.
- [25] J. Stock, *A class of tests for integration and cointegration*, in: *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, R. F. Engle and H. White (Eds.), Oxford University Press, Oxford 1999, pp. 135–167.

Frédéric Proïa

Laboratoire Angevin de REcherche en MATHématiques (UMR 6093)

Université d'Angers

Département de mathématiques, Faculté des Sciences

2 Boulevard Lavoisier, 49045 Angers cedex, France

E-mail: frederic.proia@univ-angers.fr

Received on 11.11.2015;

revised version on 21.8.2016