IMPROVING INFERENCE IN INTEGRATION AND COINTEGRATION TESTS

A Ph.D. Dissertation

by
BURAK ALPARSLAN EROĞLU

Department of
Economics
İhsan Doğramacı Bilkent University
Ankara
May 2016

to Gün and my Family

IMPROVING INFERENCE IN INTEGRATION AND COINTEGRATION TESTS

Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

BURAK ALPARSLAN EROĞLU

In Partial Fulfillment of the Requirements For the Degree of
DOCTOR OF PHILOSOPHY IN ECONOMICS

THE DEPARTMENT OF
ECONOMICS
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

May 2016

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Assoc. Prof. Dr. Taner Yiğit
Supervisor

Assist. Prof. Dr. Mirza Trokić
Co-supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Prof. Dr. Orhan Arıkan
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Assoc. Prof. Dr. Fatma Taşkın
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Assoc. Prof. Dr. Hakan Ercan
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Assist. Prof. Dr. Cem Çakmaklı
Examining Committee Member

Approval of the Graduate School of Economics and Social Sciences

Prof. Dr. Halime Demirkan
Director

ABSTRACT

IMPROVING INFERENCE IN INTEGRATION AND COINTEGRATION TESTS

EROĞLU, Burak Alparslan
Ph.D., Department of Economics
Supervisor: Assoc. Prof. Taner Yiğit
Co-supervisor: Asst. Prof. Mirza Trokiċ
May 2016

In this thesis, I address three different problems in unit root and cointegration models and I propose new methods to improve inference in testing procedures for these models. Two of these problems are related to unit root tests. First one is so-called nonstationary volatility issue, which causes severe size distortions in standard unit root tests. I try to resolve this problem with a nonparametric technique introduced first by Nielsen (2009). Second, I investigate the unit root testing under regulation, which constraints a time series process on a given interval. In this case, standard tests frequently fail to detect the presence of nonstationarity. I employ a similar methodology as in first part and provide correct inference in unit root testing for regulated series. The final problem is related to cointegration models. In these models, if innovations of the system are contaminated by MA type negative serial correlation, cointegration tests spuriously rejects the true null hypothesis. Combining wavelet theory and Nielsen's (2010) variance ratio testing procedure, I manage to reduce the impact of the problematic innovations on cointegration test. All three methods share the common feature of being nonparametric in sense that they do not require any regression or kernel type correction to handle serial correlation.

Keywords: Cointegration, Integration, Nonstationary volatility, Regulated time series, Wavelet filter

ÖZET

BÜTÜNLEŞME VE EŞGÜDÜM TESTLERİNDE ÇIKARIMSAL DÜZENLEMELER

EROĞLU, BURAK ALPARSLAN
Doktora, İktisat Bölümü
Tez Yöneticisi: Doç. Dr. Taner Yiğit
2. Tez Yöneticisi: Yard. Doç. Dr. Mirza Trokiĉ
Mayıs 2016

Bu tezde birim kök ve eşgüdüm modellerinde ortaya çıkan üç sorunu ele alınacak ve bu modellerde kullanılan testlerle yapılan çıkarımların geliştirilmesi için yeni yöntemler önerilecektir. Bu sorunlardan iki tanesi birim kök testleriyle ilgilidir. İncelenen ilk sorun durağan olmayan dalgalanma olarak adlandırılıp, standart birim kök testlerinde çok ciddi büyüklük sapmalarına neden olmaktadır. Bu sorunu Nielsen'in (2009) parametrik olmayan tekniği ile çözmeye çalışacağım. İkinci olarak zaman serilerini verili bir aralıkta sınırlayan düzenlemeler altında birim kök testlerini inceleyeceğim. Bu durumda, standard birim kök testleri durağan olmama durumunu yakalama konusunda sıklıkla başarısız olmaktadır. İlk kısımdakine benzer bir yöntem uygulayacak ve düzenlenmiş zaman serileri için birim kök testlerinde daha doğru çıkarımı sağlamaya çalışacağım. Bunlara ek olarak, son problem eşgüdüm modelleri ile alakalı olacak. Bu modellerde, eğer sistemin hata terimleri MA tipi negatif sıralı korelasyon ile kirlenmiş ise, bu testler gerçek boş hipotezi yanlışlıkla reddedecektir. Nielsen'in (2010) varyans oranı testini ve dalgacık teorisini birleştirerek, problemli hata terimlerinin eşgüdüm testlerine olan etkisini azaltmaya çalışacağım. Bu üç metot, sıralı korelasyonu ortadan kaldırmada herhangi bir regresyon ya da kernel tipi düzeltmeye ihtiyaç duymamaları dolaısıyla parametrik olmama ortak özelliğini taşımaktadırlar.

Anahtar Kelimeler: Bütünleşme, Dalgacık filtresi, Durağan olmayan varyans, Düzenlenmiş zaman serisi, Eşgüdüm

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Taner Yiğit and Mirza Trokiċ for their tremendous supervision throughout my Graduate career. Without their support and invaluable guidance the accomplishment of this thesis might not be possible. They are excellent mentors who made difference in my academic life.

I am also very grateful to Fatma Taşkın, Orhan Arıkan, Jesus Gonzalo, Carlos Velasco, for their insightful comments for my studies. I would like to thank my the examining committee members Hakan Ercan and Cem Çakmaklı, who provided very helpful comments and suggestions. I am also indebted to all of the professors at the Department of Economics, especially Cavit Pakel, Tarık Kara, Ayşe Özgür Pehlivan and Refet Gürkaynak for providing very supportive and friendly environment throughout my graduate and undergraduate years at the department. I need to mention Nimet Kaya, Meltem Sağtürk, Özlem Eraslan and Nilgün Çorapçıoğlu are very helpful with administrative matters.

I would like to thank TÜBİTAK for its financial support during my Phd and master studies and also for supporting me for research abroad program.

Special thanks to my graduate friends, especially Yıldız Akkaya, Nusret Doru, Koray Birdal, Kerim Keskin and Kemal Çağlar Göğebakan for their support and friendship.

Finally, I would like to thank to my family for their unconditional support and patience. Thank to my Mother Meryem, my Father Mustafa, my Brother Alperen, my sister Selenge, my aunt Selbi and my uncles Halil and Kürşat. Last but not least, for her unconditional love and unending support, I would like to express my thanks to Şeyma Gün Eroğlu.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The nonstationary dynamics in economic and financial data have established an important base for scientific research in time series analysis. In this base the main interest has been directed towards the development of models and methods to identify long run movements in data. From these methods and models, integration concept which brings flexibility for analyzing long run dynamics in univariate or multivariate setup, almost dominates the theoretical literature. Moreover, in empirical works, many authors such as Box and Pierce (1970), Nelson and Plosser (1982) and Baillie (1996) claim that the nonstationarity in macroeconomic and financial data can be explained by integration.

At the most elementary level, integration can be considered as aggregation of innovations or shocks through time. This aggregation is characteristically persistent for observed series, since the impact of a past shock is preserved for a long time. However, not all nonstationary series share a common persistence pattern. This kind of difference in persistence stems from different type of aggregation or additional dynamics. Within this context, integration literature has been populated by various studies which shed light on different dynamics in integrated series. Nonetheless, there is still much work to do so as to understand the nonstationary behavior of economic and financial data.

In this thesis, I first investigate two of nonstandard integration dynamics which are relatively new in the literature. These dynamics are important since they cause serious issues in classical nonstationary univariate time series analysis especially in testing the presence of nonstationarity or integration. The first dynamic is related with variance structure of the innovation or shocks of the integrated process. Many scholars (see Kim et al. (2002), Cavaliere and Taylor (2007), McConnell and Perez-Quiros (1998) etc.) report changing variance in macroeconomic and finance data. If neglected, this issue is shown to cause serious problems in nonstationarity testing (Cavaliere and Taylor, 2007), such as spurious rejection of the nonstationarity. To overcome this problem, I propose a new nonparametric unit root test. This test is nonparametric in sense that it does not necessitate any regression or kernel type of correction for serial correlation unlike the other tests in the literature. Moreover, I show that this new test also improves the small sample properties of the existing tests by means of size and power. This chapter is a joint work with Assoc. Prof. Taner Yiğit and published in Economics Letters on March 2016 with following reference: Eroğlu, B. A. and Yiğit, T. (2016). A Nonparametric Unit Root Test Under Nonstationary Volatility. Economics Letters, vol. 140, p. 6-10. The publishers "Elsevier Limited" provided permisson license for reusage and reprinting of this paper in thesis/dissertation as both electronic and print copy. Additionally license number is 3871830135669 and license date is 18 May 2016.

In the second part, I investigate the impact of range restriction on nonstationarity testing. Although range restrictions in time series econometrics are highly relevant in empirical data, they recently have attracted attention of scholars (see Cavaliere and Xu (2014), Trokić (2013) and Cavaliere (2005a)). These scholars first show how time series processes behave in nonstationary fashion while these processes are bounded between fixed intervals. Obviously, being bounded makes these series look like stationary, and this situation cause standard unit root tests to fail to differentiate the stationarity and nonstationarity. Consequently, under range constraints standard tests are size distorted. I seek to eliminate problems associated with unit root testing of bounded time series process. I utilize a similar nonparametric procedure as in the first part. Additionally I contribute to the literature as I introduce three types of dynamics for bounded time series. This chapter is joint work with my supervisor Assist. Prof. Mirza Trokiċ.

The final chapter of the thesis is about cointegration, which models the long run relations between integrated series. Cointegration can also be considered as multivariate analysis of nonstationary series. This analysis gives us chance to understand the presence of the long run relations and how they are constituted in a nonstationary setup. Moreover, cointegration framework is also used in empirical work, such as term structure of interest rates, consumption wealth relations and stock prices. In most of these work many different cointegration tests are utilized. Although literature is not scarce in cointegration tests, these tests are usually criticized by common problems. These problems are generally associated with small sample and short run dynamics in integrated systems. To handle these issues, I propose a new nonparametric method, which consists of a cointegration test based on wavelet filters and a bootsrapping routine using wavelet theory. Combined, these two method proves to be highly effective in reducing mentioned problems. This is a single author paper, which is not published yet.

The thesis is organized as follows: In chapter 2, I introduce the nonstationary volatility phenomena and propose a new method to deal with this problem in unit root testing. Chapter 3 is also concerned with unit root testing, but this time we cope with bounded or regulated time series. Chapter 4 is about combining cointegration and wavelet theory. All proofs are placed in appendix.

CHAPTER 2

NONPARAMETRIC UNIT ROOT TESTS UNDER NONSTATIONARY VOLATILITY

This chapter is published in the *Economics Letters*, March 2016 with reference: Eroğlu, B. A. and Yiğit, T. (2016). A Nonparametric Unit Root Test Under Nonstationary Volatility. Economics Letters, vol. 140, p. 6-10. The publisher of the paper is "Elsevier Limited". This company provided me permission license for reuse and reprinting of this paper in thesis/dissertation as both electronic and print copy. Additionally license number is given as 3871830135669 and license is obtained on 18 May 2016.

Recent body of evidence indicates that volatility shifts are common phenomena in macroeconomic and financial data; see Busetti and Taylor (2003), McConnell and Perez-Quiros (1998) and Sensier and Van Dijk (2004). For instance, McConnell and Perez-Quiros (1998) document a structural break in volatility of US GDP growth in the first quarter of 1984. Further, Sensier and Van Dijk (2004) report that many U.S macroeconomic data have a structural break in unconditional variance during 1959-1999. These empirical findings eventually led the researchers to investigate the impact of variance shifts on unit root tests. In one of the early studies, Kim et al. (2002) considers a single break in the innovation variance. In the case of a presence of the volatility break, they conclude that Dickey−Fuller tests can spuriously reject the null hypothesis of unit root Kim et al. (2002). Yet, Cavaliere and Taylor (2007) point out that single variance shifts may not be sufficient to explain abrupt changes in the volatility. They propose a more general model to investigate the variance shifts, referred as "Non-stationary volatility". This model allows volatility of the innovations to fluctuate as cadlag process in the limit. They also show that under these type of volatility shifts, the asymptotic distributions of standard unit root tests are altered by inclusion of a new nuisance parameter called "variance profile" of the innovations which ultimately leads to size distortions in these tests Cavaliere and Taylor (2007).

In order to achieve correct inference in non-stationary volatility models, Cavaliere and Taylor (2007) propose a two step procedure. First, they consistently estimate the variance profile. Second, They update the asymptotic distribution of Phillips and Perron (1988)'s tests with this estimator. Cavaliere and Taylor's (2007) method has several advantages over the standard unit root tests. To begin with, these new tests do not require a parametric model of volatility unlike standard methods proposed in the literature. Furthermore, one can easily simulate the asymptotic distributions after estimating the variance profile. Using these asymptotic distributions will lead to robust inference and the size distortion observed in standard tests are mostly eliminated.

Despite its advantages over previous tests, Cavaliere and Taylor's (2007) procedure relies on semi-parametric inference since they are modified versions of Phillips and Perron (1988) tests. As it is well known, these tests require the estimation of long run variance. This can be achieved by a semi-parametric kernel or a parametric ADF based regression estimation. The success of these methods highly depends on lag length, bandwidth and Kernel selection both in terms of finite sample power and size properties[1] Nielsen (2009b).

In this paper, we propose a non-parametric unit root test robust to non-stationary volatility problem. To compute the test statistic, we do not need to run any parametric regression or choose any tuning parameters such as lag length and bandwidth. The standard version of this test is developed as a ratio of the sample variance of the observed time series and that of a fractional partial sum of the series Nielsen (2009b). Nielsen's (2009) variance ratio statistic, by its construction, eliminates short run dynamics such as serial correlation without a parametric regression or a kernel estimation. Nonetheless, this test is exposed to non-stationary volatility problem, since it relies on the assumption of identically distributed innovations. So as to avoid this problem, we adopt a two step procedure similar to Cavaliere and Taylor (2007). As the first step, we utilize the non-parametric variance profile estimator of Cavaliere and Taylor (2007). [2] Afterwards, we apply the Nielsen's (2009) variance ratio statistic. We also show that the asymptotic distribution of the new test can be easily simulated after the estimation of variance profile. Furthermore, because our test includes fractional transformation

[1]Cavaliere et al. (2015) develop a bootstrap based approach to handle non-stationary volatility in unit root test. This approach also depends on consistent lag selection in ADF regression.

[2]This will not distort the non-parametric nature of our test.

of observed series, we derive the limiting distribution of fractionally integrated processes with non-stationary volatility. We also provide simulation based evaluation of this new object.

2.1 Model and Variance Ratio Test

2.1.1 Model

Let the time series process $\{x_t\}_{t=0}^{T}$ be generated according to the very standard autoregressive model with deterministic components.

$$x_t = y_t + \theta'\delta_t \tag{2.1}$$

$$y_t = \rho y_{t-1} + u_t \tag{2.2}$$

$$u_t = C(L)\epsilon_t = \sum_{i=0}^{t} c_j \epsilon_{t-j} \tag{2.3}$$

$$\epsilon_t \sim iid(0, \sigma_t^2) \tag{2.4}$$

The last line can be written as $\epsilon_t = \sigma_t e_t$ where $e_t \sim iid(0,1)$. This expression indicates that the innovations of the AR process defined above have time varying variance. Also note that $\theta'\delta_t$ is the deterministic term in equation (2.1) where $\theta = [\theta_0, \theta_1]$ is a 1×2 vector and $\delta_t = [1, t]'$. We consider three cases for deterministic term. First, whenever $\theta_0 = \theta_1 = 0$, x_t reduces to y_t without any deterministic terms. Second, when $\theta_1 = 0$ with $\theta_0 \neq 0$, x_t becomes y_t plus a mean. Lastly, when both $\theta_0 \neq 0$ and $\theta_1 \neq 0$, then x_t possess both a mean and a trend components. Further, $C(L)$ is the lag polynomial and the error term u_t is a linear process. Following assumption from Cavaliere and Taylor (2007) characterizes the dynamics of the innovations u_t:

Assumption 1. $\mathcal{A}.1$ The lag polynomial $C(L) \neq 0$ for all $|L| \leq 1$, and $\sum_{j=0}^{\infty} j|c_j| < \infty$. $\mathbb{E}|e_t|^r < K < \infty$ for some $r \geq 4$.

$\mathcal{A}.2$ ρ satisfies $|\rho| \leq 1$.

$\mathcal{A}.3$ The standard deviation of error term σ_t satisfies $\sigma_{\lfloor Ts \rfloor} := \omega(s)$ for all $s \in [0, 1]$, where $\omega(.) \in \mathscr{D}$ is non-stochastic and strictly positive with for $t < 0$, $\sigma_t \leq \sigma^* < \infty$ is uniformly bounded.

The assumptions $\mathcal{A}.1$ and $\mathcal{A}.2$, as indicated in Cavaliere and Taylor (2007), are very standard in unit root testing literature. $\mathcal{A}.1$ is needed for invertibility of the process u_t. It also indicates the error term is a stable linear process. $\mathcal{A}.2$ is required to avoid explosive processes for y_t, therefore the AR coefficient is inside the unit circle. Moreover, $\mathcal{A}.3$ is non-standard in the classical unit root testing. This assumption characterizes the dynamics of innovation variance, which should be bounded and display a countable number of jumps (Cavaliere and Taylor, 2007). For instance, if we have a single break in innovation variance the function $\omega(s)$ takes form of $\omega(s) = \sigma_0 + (\sigma_1 - \sigma_0)\mathbb{1}(s > \tau)$ for some $0 < \tau < 1$. In another example of admissible functional form, $\omega(s)$ has trending behaviour, where we can define $\omega(s) = \sigma_0 + s\sigma_1$ for $s \in [0, 1]$. We can also generalize this form as $\omega(s) = \sigma_0 + \sigma_1 f(s)$ where σ_0, σ_1 and $f(s)$ ensures $\omega(s) \in \mathscr{D}$ and. Finally, $w(s)$ can be represented as non negative function of any cadlag process, which again ensures $w(s) \in \mathscr{D}$ for all $0 \leq s \leq 1$. For instance, as in Cavaliere and Taylor (2007) $w(s) = \sigma_0 \exp(\nu J_c(s))$, where $J_c(s)$ is a Uhlenbeck-Orstein process with local to unity parameter c.

A fundamental object that is defined in Cavaliere and Taylor (2007) is given below:

$$\eta(s) := \left(\int_0^1 \omega(r)^2 dr\right)^{-1} \left(\int_0^s \omega(r)^2 dr\right) \tag{2.5}$$

This object will be referred as the *variance profile* of the process and $\eta(s) \in \mathscr{C}$. When $\omega(s) = \omega$ a constant, $\eta(s) = s$ under homoscedasticity. Further, Cavaliere and Taylor (2007) showed that $\int_0^1 \omega(r)^2 dr = \bar{\omega}^2$ is the limit of $T^{-1}\sum_{t=1}^T \sigma_t^2$, and may be called as asymptotic average innovation variance.

2.1.2 Unit root Asymptotics of Variance Ratio test under Non-stationary volatility

So as to devise the Variance Ratio test Nielsen (2009b) statistic We need to define the fractional partial sum operator:

$$\tilde{x}_t := \Delta_+^{-d_1} x_t, \; t = 0, 1, ..., \; d > 0 \tag{2.6}$$

$$\Delta_+^{-d_1} x_t = (1-L)_+^{-d_1} x_t = \sum_{k=0}^{t-1} \frac{\Gamma(k+d_1)}{\Gamma(d_1)\Gamma(k+1)} x_{t-k} = \sum_{k=0}^{t-1} \pi_k(d_1) x_{t-k} \tag{2.7}$$

where $\Gamma(.)$ is gamma function and $\pi_k(d)$ is $k-th$ fractional binomial expansion coefficient. Note that the operator $\Delta_+^{-d_1}$ filters the observed series x_t with positive indexed history. Under the assumptions $\mathcal{A}$, following Lemmas basically characterize the asymptotic behaviour of unit root and near unit root processes:

Lemma 1. Assume that $\{u_t\}_0^T$ is generated by (2.3)-(2.4). Further assume that

$\rho = 1 - c/T$ with $c \geq 0$.

i. $y_T(t) = T^{-1/2} \sum_{k=1}^{\lfloor Tt \rfloor} e^{-c(\lfloor Tt \rfloor - k} u_k \xrightarrow{w} \bar{\omega} C(1) J_\omega^c(t)$,

   where $J_\omega^c(t) = \int_0^t exp(-c(s-r)) dB_\omega(r)$ and $B_\omega(s) = \bar{\omega}^{-1} \int_0^s \omega(r) dB(r)$.

ii. $B_\omega(s) = B_\eta(s) := B(\eta(s))$ where $B_\eta(s)$ variance shifted Brownian motion, $\eta(s)$ is defined in (2.5). Thus, $J_\omega^c(t) = J_\eta^c(t) = \int_0^t exp(-c(s-r)) dB_\eta(r)$

iii. For all $d > 1$, $\tilde{y}_T(t) = T^{-d_1} \Delta_+^{-d_1} y_T(t) \xrightarrow{w} \bar{\omega} C(1) J_{\omega,d_1}^c(t)$, where $J_{\omega,d_1}^c(t) = \Gamma(d_1+1)^{-1} \int\limits_0^t (t-s)^{d_1} dJ_\omega^c(s)$. Further we can write $J_{\omega,d_1}^c(t) = J_{\eta,d_1}^c(t)$

Remark 1. Lemma 1.(i) and 1.(ii) are from Cavaliere (2005b) and Cavaliere and Taylor (2007). To our knowledge, Lemma 1.(iii) is a new one which establish weak convergence for fractionally integrated process with non-stationary volatility. Although Demetrescu and Sibbertsen (2014) models the fractional integrated process with non-stationary volatility, they do not establish weak convergence of this object.

Remark 2. Note that under the null hypothesis of $\rho = 1$ or $c = 0$ the above limits become Variance shifted Brownian motions, instead of Variance shifted Uhlenbeck-Orstein process. For instance, under the null the partial sum process $\tilde{y}_T(t)$ will converge to $\bar{\omega} C(1) \int_0^t (t-s)^{d_1} dB_\eta(s)$ where we can define $B_{\eta,d_1}(t) := \int_0^t (t-s)^{d_1} dB_\eta(s)$. This limiting distribution resembles the type II fractional Brownian motions defined by Marinucci and Robinson (2000), since $B_{\eta,d_1}(t)$ does not contain any pre-historic influence (see also Wang et al. (2002)).

With the help of these tools, we are ready to device the test statistic, which is first introduced by Nielsen (2009b), first consider the case of no deterministic trend.

$$T^{-2}\sum_{t=1}^{T} y_t^2 \xrightarrow{w} (\bar{\omega}C(1))^2 \int_0^1 B_\eta(s)^2 ds \tag{2.8}$$

$$T^{-2-2d_1}\sum_{t=1}^{T} \tilde{y}_t^2 \xrightarrow{w} (\bar{\omega}C(1))^2 \int_0^1 B_{\eta,d_1}(s)^2 ds \tag{2.9}$$

Dividing (2.8) with (2.9), we obtain the test statistic:

$$\rho_\eta(d_1) = T^{2d_1}\frac{\sum_{t=1}^{T} y_t^2}{\sum_{t=1}^{T} \tilde{y}_t^2} \tag{2.10}$$

This test statistic is the ratio of the variances of observed data x_t and the variance of fractional partial sum of x_t. From Lemma 1, applying Continuous mapping theorem (CMT), it is easy to see that

$$\rho_\eta(d_1) \xrightarrow{w} \frac{\int_0^1 B_\omega(s)^2 ds}{\int_0^1 B_{\omega,d_1}(s)^2 ds} = \frac{\int_0^1 B_\eta(s)^2 ds}{\int_0^1 B_{\eta,d_1}(s)^2 ds} \tag{2.11}$$

Remark 3. We denote Nielsen's (2009) test as $\rho(d_1)$ throughout the paper. The asymptotic distribution of this test under the null hypothesis is also very similar to the one we have for Non-stationary Volatility case, that is:

$$\rho(d_1) \xrightarrow{w} \frac{\int_0^1 B(s)^2 ds}{\int_0^1 B_{d_1}(s)^2 ds}$$

where $B(s)$ is standard Brownian motion and $B_{d_1}(s)$ is a fractional Brownian motion with integration order $d_1 + 1$. For further details, see Nielsen (2009b).

As in Nielsen (2009b), the limit of the test statistic does not contain any parameter associated with long run dynamics. This is because of the fact that both numerator and denumerator share same long run variance, that is $(\bar{\omega}C(1))^2$. However, this test statistic and its limiting distribution still contains fractional integration order d_1. This parameter is choice of econometricians and does not act as a nuisance parameter, the only restriction is $0 < d_1 \leq 1$. Further, if $d_1 = 1$ the test is equivalent to Breitung (2002). However, Nielsen (2009b) shows that the most power is attained at $d_1 = 0.1$. We will fallow Nielsen (2009b) and use this value in our test.

To adjust for non-zero mean and possible time trend, we assume the deterministic term is represented as $\theta'\delta_t$, where $\delta_t = 1$ when we have a mean and $\delta_t = [1, t]'$ when we have mean and time trend. To clear out the deterministic terms, we apply OLS de-trending mechanism to observed series x_t. Let $\hat{x}_t$ is residuals from regression of δ_t on x_t such that $\hat{x}_t = x_t - \hat{\theta}'\delta_t$. We can rewrite this as $\hat{x}_t = y_t - (\hat{\theta} - \theta)'\delta_t$. Moreover, define $\tilde{\hat{x}}_t = \Delta_+^{-d_1}\hat{x}_t$ as fractional partial sum of $\hat{x}_t$. Following Theorem will demonstrate the asymptotic behaviour of the de-trended variables.

Theorem 1. Assume that the time series $\{x_t\}$ is generated by equations (2.1)-(2.4) and $\rho = 1 - c/T$ for $c \geq 0$. Let $j = 1$ when $\delta_t = 1$ and $j = 2$ when $\delta_t = [1, t]'$ for $d_1 > 0$

i. $\hat{x}_T(t) \xrightarrow{w} J^c_{\eta,j}(s)$ where

$$J^c_{\eta,j}(s) = J^c_\eta(s) - \left(\int_0^1 J^c_\eta(s)D_j(s)'ds\right) \left(\int_0^1 D_j(s)D_j(s)'ds\right)^{-1} D_j(s)$$

for $j = 1, 2$, and $D_1(s) = 1$, $D_2(s) = [1, s]'$.

ii. $\tilde{\hat{x}}_T(t) \xrightarrow{w} J^c_{\eta,d_1,j}(s)$ where

$$J^c_{\eta,d_1,j}(s) = J^c_{\eta,d_1}(s) - \left(\int_0^1 J^c_\eta(s) D_j(s)' ds\right)\left(\int_0^1 D_j(s) D_j(s)' ds\right)^{-1}$$
$$\int_0^r \frac{(r-s)^{d-1}}{\Gamma(d)} D_j(s) ds$$

for $j = 1, 2$, and $D_1(s) = 1$, $D_2(s) = [1, s]'$.

iii. $\rho_\eta(d_1) = T^{2d} \frac{\sum_{t=1}^T \hat{x}_t^2}{\sum_{t=1}^T \tilde{\hat{x}}_t^2} \xrightarrow{w} U_{j,\eta}(d_1) = \frac{\int_0^1 J^c_{\eta,j}(s)^2 ds}{\int_0^1 J^c_{\eta,d_1,j}(s)^2 ds}$.

Remark 4. Theorem 1.(i) is called de-trended variance shifted Uhlenbeck-Orstein (U-0) process and can be derived easily from the results of Cavaliere and Taylor (2007). Theorem 1.(ii) and (iii) are new in literature. We call the limiting process $J^c_{\eta,d_1,j}(s)$ de-trended variance shifted Fractional Brownian motion.

Remark 5. Also note that if $c = 0$, variance shifted U-O processes above become variance shifted Brownian motions. For instance, under $c = 0$ $\tilde{\hat{x}}_T(t) \xrightarrow{w} B^c_{\eta,d_1,j}(s)$ where $B^c_{\eta,d_1,j}(s)$ can be obtained by plugging $B_{\eta,d_1}(s)$ instead of $J^c_{\eta,d_1}(s)$ in Theorem 1.(ii).

2.1.3 Simulated Asymptotic distribution

Note that all asymptotic distributions for test statistic in theorem 1 depends on *variance profile* $\eta(s)$. However, according to Cavaliere and Taylor (2007), this can be consistently estimated under the null hypothesis by following non-parametric estimator:

$$\hat{\eta}(s) := \frac{\sum_{t=1}^{\lfloor Ts \rfloor} (\Delta \hat{x}_t)^2 + (Ts - \lfloor Ts \rfloor)(\Delta \hat{x}_{\lfloor Ts \rfloor + 1})^2}{\sum_{t=1}^{T} (\Delta \hat{x}_t)^2} \tag{2.12}$$

After obtaining the consistent estimate for $\eta(s)$, we can simulate the asymptotic distribution and the critical value for the test statistic. Now define the partial sum $B_{\eta,T}(s) := T^{-1/2}\sum_{t=1}^{\lfloor(\eta(\lfloor Ts\rfloor/T)T\rfloor} e_t$ which satisfies $B_{\eta,T}(s) \xrightarrow{w} B_\eta(s)$ Cavaliere and Taylor (2007). Following Theorem establishes the convergence:

Theorem 2. Under the conditions of Theorem 1

i. (Cavaliere and Taylor (2007)) $B_{\hat{\eta},T}(s) := T^{-1/2}\sum_{t=1}^{\lfloor(\hat{\eta}(\lfloor Ts\rfloor/T)T\rfloor} e_t \xrightarrow{w} B_\eta(s)$

ii. $B_{\hat{\eta},d_1,T}(s) := T^{-d_1}\Delta_+^{-d_1} B_{\hat{\eta},T}(s) \xrightarrow{w} B_{\eta,d_1}(s)$

Remark 6. The part (i) of above theorem belongs to Cavaliere and Taylor (2007). We need this to simulate the numerator of our test statistic.

Remark 7. The part (ii) of above theorem illustrates how one can simulate variance shifted fractional Brownian motion and this is also a new construction.

With these two object in theorem 2 we can simulate the asymptotic distributions and find the critical values. The simulation for $B_{\hat{\eta},T}(s)$ is straightforward. First we need to choose a step level N. Let $s = j/N$ for $j = 1, 2, ..., N$, we compute $\eta(\lfloor Ts\rfloor/T)$ with (2.12). Then draw T random variable from $N(0,1)$, call this random variable $\{e_t\}_{t=1}^T$. Finally, $B_{\hat{\eta},T}(s)$ is given by formulation in Theorem 2 part(i). In order to compute $B_{\hat{\eta},d_1,T}(s)$, we first need $B_{\hat{\eta},T}(s)$. Then, we apply fractional integration operator $\Delta_+^{-d_1}$ to this object and scale it with T^{-d_1}.

The test rejects the null hypothesis for large values the test statistic. If α is confidence level, the critical value should be $1-\alpha$ quantile of the simulated distribution. Following, theorem demonstrates properties of critical values as well as size and power.

Theorem 3. Let the assumptions of Theorem 1 hold then, the test that reject H_0 when $\rho_\eta(d_1) > CV_{j,\alpha}(d_1)$ with $CV_{j,\alpha}(d_1) = P\left(U_{j,\eta}(d_1) > CV_{j,\alpha}\right) = \alpha$. Moreover, the test is consistent against H_1 which consists of stationary alternatives with the condition $\sum_{-\infty}^{\infty} |\gamma(k)| < \infty$ is satisfied where $\gamma(k) = \mathbb{E}[y_t y_{t-k}]$,

Theorem 3 indicates that under stationary alternatives the asymptotic power is achieved. Moreover, the nominal size is obtained by construction.

2.2 Finite Sample properties

In this section, we demonstrate the finite sample performance of proposed test via Monte Carlo simulations. In these simulations, we also report the size and power properties of Cavaliere and Taylor (2007) tests. Considering the space constraints, we do not give exact formulation for these tests, but they can be found in Cavaliere and Taylor (2007).

In the Monte Carlo simulations, data is generated according to equations (2.1)-(2.4) with $T = \{100, 500\}$. We consider following specifications for error term variance:

(i). *Constant volatility (CV)*: $\omega(s) = 1$ for $s \in [0,1]$ and $\sigma_0 > 0$

(ii). *Single break in volatility (SBV)*: $\omega(s) = 1 + 2 * \mathbb{1}(s > 0.2 * T)$ for $s \in [0,1]$.

(iii). *Trending volatility (TV)*: $\omega(s) = 1 + 2 * s$ for $s \in [0,1]$.

(iv). *Exponential integrated Stochastic volatility(EISV)* $\omega(s) = \sigma_0 \exp(4B(s))$ for $s \in [0,1]$ where $B(s)$ is standard Brownian process.

The innovations e_t are drawn from $N(0,1)$. All simulations are conducted $MC = 10000$ times. We fix the step size N to T in simulating the variance shifted Brownian motions. We consider four scenarios for serial correlation in innovations. First one does not contain any serial correlation. In second, u_t follows an AR(1) process with $u_t = 0.5u_{t-1} + e_t$, in third we consider ARMA(2,2) process: $u_t = 0.1u_{t-1} + 0.07u_{t-2} - 0.4e_{t-1} + 0.2e_{t-2} + e_t$. Last one follows a MA(2) process: $u_t = -0.2e_{t-1} + 0.15e_{t-2} + e_t$. We fix $\rho = \{1, 0.93, 0.86\}$. $\rho = 0$ indicates size and other values are for power evaluation. We also provide simulation for Cavaliere and Taylor (2007) MZ_t^S test.

For experiments with no serial correlation we set lag truncation parameter to 0 for Cavaliere and Taylor (2007) tests. The critical values for all Non-stationary Volatility tests are calculated by direct simulation, which is defined in Theorem 2. For step size in simulating the variance shifted Brownian motions, we use the sample size T. Note that Cavaliere and Taylor (2007) considers simulation of the test statistic with $N = 1000$ Monte Carlo. This implies that they need to simulate the data for $(N) \times MC$ which is too much time consuming[3]. Instead of this method we adopt a procedure which is similar to double fast bootstrap (DFB) technique of Davidson and MacKinnon (2007). This procedure only requires to simulate the data MC times. We can summarize the method as following:

1. For each Monte Carlo, compute the usual test statistic and compute one simulated asymptotic statistic which corresponds to $B_{\hat{\eta}}(.)$ and $B_{\hat{\eta},d_1}(.)$.

2. Do the step 1 for MC times to obtain MC test statistic and MC simulated

[3] We also run simulation with this setup for some scenarios. Results are very close to the DFB type of algorithm

asymptotic statistic and denote the α quantile of simulated test statistic as $Q^*_{test}(\alpha)$.

3. Obtain the critical values for Cavaliere and Taylor (2007) as $Q^*_{MSB^S}(0.05)$, $Q^*_{MZ^S_\alpha}(0.05)$ and $Q^*_{MZ^S_t}(0.05)$, and for our test $Q^*_{\rho_\eta(0.1)}(0.95)$ and for Nielsen' (2009) $Q^*_{\rho(0.1)}(0.95)$

Table 1 summarize the size and power results. When we consider the finite sample size, our test seems to have advantage over MZ^S_t test of Cavaliere and Taylor (2007). However, in some cases, such as AR(1) model with constant volatility and time varying volatility, MZ^S_t has slightly better power. In other case, we are always less size distorted. Further, power evaluations clearly favor our test against MZ^S_t. In every, except no serial correlation scenario with small sample size, we have better power. In most of the cases, we are twice powerful than MZ^S_t test. These findings indicate that we do not only provide simpler algorithm for unit root testing under nonstationary volatility, but we also provide a test with better finite sample properties.

2.3 Conclusion and Discussion

In this paper, we propose a new unit root testing mechanism when true data generation process has innovations with non-stationary volatility. We combine the results of Cavaliere and Taylor (2007) with Nielsen's (2009) non-parametric variance unit root test. Using this construction, our test enjoys all desirable features of Nielsen's (2009) test. In addition, the proposed method provides a unit root testing procedure robust to non-stationary volatility unlike Nielsen (2009b).

Table 1: **Emprical Size and Power of $\tau_\eta(0.1)$ and MZ_t^S**

		No serial correlation					
			$\tau_\eta(d_1)$			MZ_t^S	
		$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$	$\rho = 1$	$\rho = 0.93$	$\rho = 0.86$
CV	T=100	0.052	0.251	0.598	0.034	0.171	0.527
	T=500	0.051	0.987	1.000	0.048	1.000	1.000
SBV	T=100	0.052	0.284	0.656	0.031	0.311	0.738
	T=500	0.050	0.993	1.000	0.046	1.000	1.000
TV	T=100	0.054	0.257	0.588	0.029	0.261	0.654
	T=500	0.055	0.979	1.000	0.038	0.998	1.000
EISV	T=100	0.061	0.280	0.633	0.034	0.300	0.687
	T=500	0.053	0.949	1.000	0.044	0.986	1.000
				AR(1)			
			$\tau_\eta(d_1)$			MZ_t^S	
CV	T=100	0.016	0.237	0.507	0.035	0.140	0.290
	T=500	0.035	0.974	1.000	0.045	0.952	0.991
SVB	T=100	0.018	0.260	0.559	0.038	0.224	0.408
	T=500	0.043	0.982	1.000	0.041	0.976	0.995
TV	T=100	0.016	0.235	0.510	0.036	0.186	0.321
	T=500	0.036	0.966	1.000	0.043	0.940	0.986
EISV	T=100	0.017	0.175	0.419	0.014	0.079	0.182
	T=500	0.045	0.977	1.000	0.043	0.934	0.987
				ARMA(2,2)			
			$\tau_\eta(d_1)$			MZ_t^S	
CV	T=100	0.047	0.240	0.591	0.020	0.071	0.218
	T=500	0.045	0.985	1.000	0.040	0.940	0.991
SVB	T=100	0.049	0.286	0.656	0.011	0.215	0.490
	T=500	0.048	0.987	1.000	0.044	0.966	0.995
TV	T=100	0.047	0.269	0.607	0.014	0.144	0.332
	T=500	0.053	0.974	1.000	0.044	0.927	0.987
EISV	T=100	0.054	0.277	0.633	0.009	0.136	0.310
	T=500	0.049	0.884	0.996	0.046	0.654	0.868
				MA(2)			
			$\tau_\eta(d_1)$			MZ_t^S	
CV	T=100	0.055	0.240	0.576	0.026	0.094	0.279
	T=500	0.051	0.987	1.000	0.039	0.960	0.993
SVB	T=100	0.055	0.283	0.649	0.016	0.234	0.531
	T=500	0.054	0.989	1.000	0.042	0.977	0.995
TV	T=100	0.054	0.260	0.600	0.018	0.175	0.405
	T=500	0.046	0.982	1.000	0.037	0.955	0.987
EISV	T=100	0.053	0.301	0.669	0.011	0.249	0.557
	T=500	0.050	0.980	1.000	0.035	0.851	0.953

Note: The confidence level is $\alpha = 0.05$ and there is no trend and mean components. d_1 is fixed to 0.1 as recommended in Nielsen (2009b). For formula and asymptotic distribution of MZ_t^S test see (Cavaliere and Taylor, 2007). In fact Cavaliere and Taylor (2007) propose 3 different test statistic, but we only give the results of the best one from these tests. For selection of lag length we utilize MAIC proposed by Ng and Perron (2001)

Simulation evidence suggests that the proposed test has good size and power properties. On the outset, it has slightly less power than Cavaliere and Taylor (2007) tests when there is no serial correlation, but in other cases, we have better power. Besides, in terms of size performance, we can see that in general our test is less size distorted than MZ_t^S test. This result is also natural since we are using nonparametric adjustment for serial correlation.

Last but not least, we derive new theoretical tools for "Fractionally integrated process with non-stationary volatility" as the construction of the proposed unit root test involves fractional integration of the observed series. With these new tools, our approach can also be used for testing fractional integration order in $I(d+1)$ model with non-stationary volatility. This requires the limiting distribution of $y_T(s) = T^{-1/2-d} \sum_{k=1}^{\lfloor Tt \rfloor} \pi_k(d) u_t$ where u_t is innovation defined in (2.3)-(2.4). The proof should follow Lemma 1.iii. Further, the alternative against the null $H_0 : d = d_0$ is $H_a : -1/2 < d < d_0$. We just skip this testing procedure to save space.

CHAPTER 3

TESTING FOR UNIT ROOTS UNDER REGULATION AND SERIAL CORRELATION

3.1 Introduction

A salient feature of certain nonstationary time series which renders inference particularly challenging is the presence of inherent or artificial (policy control) bounds. Such series are said to be *regulated* (limited) or bounded, see Granger (2010), and are of significant practical importance to time series econometricians. For instance, macroeconomic series such as nominal interest rates, production, and unemployment rates, are important examples of processes that are inherently regulated below (at zero, say), above (full capacity), or both. Similarly, series exhibiting artificial regulation include price fixing (minimum wages, say), target zone exchange rates, or planning and inventory control problems in empirical microeconomics. What makes these series especially challenging to analyze is that the presence of bounds, particularly when they are tight, forces nonstationary series to increasingly resemble stationary ones as the series nears the bounds. This effect, usually overlooked in practice, renders traditional tools of analysis theoretically unjustifiable. This disregard for the nature of regulated data not only renders traditional integration methods theoretically unjustifiable, but by extension, inferential exercises can be shown to be significantly unreliable.

Until recently, Cavaliere (2005a), Granger (2010), and Cavaliere and Xu (2014) were the only serious efforts to develop a theory for regulated integrated processes. Particularly important is the seminal work of Cavaliere (2005a) in which he develops asymptotic distributions for well known unit root test statistics when the driving series is a regulated I(1) process. These results were further developed and expounded on in Cavaliere and Xu (2014) by broadening the framework to allow for more general innovation structures. They first introduce the serial correlation in innovations process, then use serially correlated innovations in regulation process. Since regulation process is a nonlinear transform of innovations, the introduction of the serial correlation in this way cannot be handled as a linear model as implicitly claimed in Cavaliere and Xu (2014) and Cavaliere (2005a). We first show the potential problem associated with this idea in Cavaliere and Xu (2014) methodology with a simple counterexample. Nonetheless, this counterexample does not mean one cannot introduce serial correlation in regulated time series.

In this study, we propose three different data generation processes which embed the serial correlation notion in regulated time series process. First data generation specification is a new one in the literature. With this specification, we can utilize the desired features of linear nonstationary models unlike Cavaliere and

Xu (2014). We do not use a model nonlinear in innovations, but instead use a linear model which includes regulated innovations. The second model concerns with a recent generalization of bounded integrated series. Trokić (2013) extends the regulated I(1) framework and develops the limiting distribution for regulated I($d+1$) processes under general innovation structures. He shows that unlike in the case of the regulated I(1) process where the limiting distribution is a regulated Brownian motion, the regulated I($d+1$) process tends in distribution to a regulated fractionally integrated Brownian motion. In the final specification, we utilize the data generation in Cavaliere (2005a). Although we are not using a new model, we provide a elegant way to handle the bound problem. The basic idea is removing the regulation components from the observed series by using local time concept, and test the unit root hypothesis as the series is not regulated.

For each data generation process, we derive the limiting distributions of the objects utilized in unit root testing. For testing procedure, we adopt Nielsen's (2009) variance ratio principle. This method utilizes nonparametric statistics different than the tests proposed in Cavaliere and Xu (2014). These proposed statistics have desirable power properties in addition to not suffering from ambiguous short-run dynamics estimators. Moreover, in order to carry over the results of Nielsen (2009a) to the framework of bounded series, this paper also develops the first theoretical justification for the limiting distribution of an integrated bounded series with a linear time trend. Additionally, limiting distributions are also obtained in the case of OLS detrended series.

This paper is organized as follows. The next section will state the problem of Cavaliere and Xu (2014) methodology with a counter example. In this section, we will introduce the three different serial correlation types for regulated series. The assumptions and theoretical results are outlined there as well. Section 3 presents simulation studies and Section 4 concludes. All proofs are placed in Appendix.

3.2 Regulated Integrated Processes with Serial Correlation

For some bounded interval $[\underline{b}, \overline{b}]$ with fixed bounds $\underline{b} < \overline{b}$, a series y_t is said to be *regulated* or *bounded* if $y_t \in [\underline{b}, \overline{b}]$ almost surely for all t. When y_t is regulated and integrated to order 1, these series are called bounded I(1), or BI(1) processes; see Cavaliere (2005a) and Cavaliere and Xu (2014). More generally, when y_t is regulated and integrated to order $(d+1) \in \mathbb{R}, d > -1/2$, such series are termed regulated fractionally integrated, or RFI$(d+1)$ processes; see Trokić (2013).

Abstracting momentarily from fractional integration and serially correlated errors, a BI(1) processes can be formalized thus:

$$x_t = \gamma_0 + y_t \tag{3.1}$$

$$y_t = y_{t-1} + u_t \tag{3.2}$$

$$u_t = \epsilon_t + \underline{\xi}_t - \overline{\xi}_t \tag{3.3}$$

where y_t is initialized at $y_0 = O_p(1)$, ϵ_t is a martingale difference sequence (MDS),

γ_0 is a constant deterministic component, and $\underline{\xi}_t$ and $\overline{\xi}_t$ are non-negative processes called *regulators* which satisfy the following relations:

$$\underline{\xi}_t > 0 \quad \text{iff} \quad y_{t-1} + \epsilon_t < \underline{b} - \gamma_0 \tag{3.4}$$

$$\overline{\xi}_t > 0 \quad \text{iff} \quad y_{t-1} + \epsilon_t > \overline{b} - \gamma_0 \tag{3.5}$$

This is in fact the Cavaliere and Xu (2014) specification of regulated integration where x_t is an I(1) process regulated on the interval $[\gamma_0 + \underline{b}, \gamma_0 + \overline{b}]$. Consider further the continuous time $t \in [0,1]$ approximant $x_T(t) = T^{-1/2}\left(x_{\lfloor Tt \rfloor} - x_0\right)$ of x_t on the càdlàg space $\mathcal{D}[0,1]$. The following assumptions guarantee functional central limit theorem (FCLT) results for $x_T(t)$.

Assumption 2. $\mathcal{A}$

1. $\{\epsilon_t, \mathcal{F}_t\}$ is a MDS with respect to some filtration $\mathcal{F}_t$.
2. $E(\epsilon_t^2|\mathcal{F}_t) = \sigma^2 < \infty$ and $E(|\epsilon_t|^p) < \infty$ for $p = 2/(2d+1)$ and $d > -1/2$.
3. $\sup_{t\in\mathbf{Z}} E\{|\underline{\xi}_t|^p\} < \infty$ and $\sup_{t\in\mathbf{Z}} E\{|\overline{\xi}_t|^p\} < \infty$.
4. $\max_{t=1,\ldots,T} |\underline{\xi}_t| = o_p(T^{d+1/2})$ and $\max_{t=1,\ldots,T} |\overline{\xi}_t| = o_p(T^{d+1/2})$
5. $\left(\frac{\sigma}{\Gamma(d+1)}T^{(d+1/2)}\right)^{-1}(\underline{b} - \gamma_0) = \underline{c} + o(1)$ and $\left(\frac{\sigma}{\Gamma(d+1)}T^{(d+1/2)}\right)^{-1}(\overline{b} - \gamma_0) = \overline{c} + o(1)$ for some constants $\underline{c} < \overline{c}$.

Assumptions $\mathcal{A}$ are in fact designed to accommodate the broader class of RFI($d+1$) processes. In particular, they are required to invoke fractional FCLTs in the results to follow. $\mathcal{A}.2$ is particularly delicate since the moment condition can be very strong when d is close to $-1/2$. Nevertheless, the condition has an established tradition since Davydov (1970) and is shown in Johansen and Nielsen

(2012) to be necessary. Moreover, it is readily verified that invoking assumption $\mathcal{A}$ with $d = 0$ renders conditions necessary to ensure asymptotic convergence of $x_T(t)$ when x_t is a BI(1) process. In this regard, let $\Longrightarrow$ denote weak convergence and define $\lfloor r \rfloor$ as the smallest integer not greater than r. Under Assumptions $\mathcal{A}$ with $d = 0$, if $\mathcal{D}[0, 1]$ is endowed with the Skorohod topology and $x_0 \in [\gamma_0 + \underline{b}, \gamma_0 + \overline{b}]$, Cavaliere (2005a); Cavaliere and Xu (2014) have shown that $x_T(t) \Longrightarrow \sigma B^{\underline{c},\overline{c}}(t)$, as $T \to \infty$, where $B^{\underline{c},\overline{c}}$ is a Brownian motion regulated on the interval $[\underline{c}, \overline{c}]$; see Harrison (1985). We now show that augmenting the model with serially correlated innovations results in a unique set of challenges.

3.2.1 BI(1) Processes with Serial Correlation: A Counterexample

Here we demonstrate that the regulated autoregressive (AR) model with serially correlated innovations in Cavaliere and Xu (2014) can result in the breakdown of the regulation process. Consider again equations eq. (3.1) – eq. (3.3) and express ϵ_t as a general linear process of the form $\epsilon_t = \psi(L)v_t$, where L is the lag operator, v_t is a MDS, and $\psi(z) = \sum_{j=1}^{\infty} \psi_j z^j$. $\psi(z)^{-1} = \alpha(z) = 1 - \sum_{j=1}^{\infty} \alpha_j z^j$ is well defined under the following assumptions.

Assumption 3. $\mathcal{B}$

1. $\sum_{j=0}^{\infty} |\psi_j| < \infty$, $\sum_{j=0}^{\infty} j|\psi_j| < \infty$, and $b_\psi = \sum_{j=0}^{\infty} \psi_j \neq 0$.

The noise processes is now defined as:

$$u_t = \psi(L)v_t^*$$
$$v_t^* = v_t + \underline{\xi}_t^* - \overline{\xi}_t^*$$
$$V_t^* = \sum_{i=1}^{t} v_i^*$$

where $\underline{\xi}_t^* = \psi(L)^{-1}\underline{\xi}_t$ and $\overline{\xi}_t^* = \psi(L)^{-1}\overline{\xi}_t$. Lemma A.1 in Cavaliere and Xu (2014) claims that $T^{-1/2}V_{\lfloor T\cdot \rfloor}^* \implies \sigma C(1)B^{\underline{c},\overline{c}}(\cdot)$. The proof should follow from the Harrison (1985) construction of the Skorohod (1961) equation for reflected Brownian motions. Specifically, for any stochastic process $x(t)$ in $\mathcal{C}$ and fixed bounds $[\underline{b}, \overline{b}]$, if there exist non-negative, non-decreasing functions $L(t), U(t) \in \mathcal{C}$ which increase only at the points of regulation, then, $z(t) = x(t) + L(t) - U(t) \in [\underline{b}, \overline{b}]$, provided $z(0) \in [\underline{b}, \overline{b}]$.

We demonstrate that the claimed convergence is falsified even for simple cases of a single lower bound. We show that regulators $\underline{\xi}_t^*$ can fail to satisfy non-negativity as in equation eq. (3.4), and that $L(t)$ is in fact decreasing at points where regulation is not active. In particular, suppose v_t follows a stationary AR(1) process so that $\psi(L)^{-1} = (1 - \phi L)$ with $0 < \phi < 1$. Suppose further that regulation occurs only once at some fixed time t_0. Accordingly, $\underline{\xi}_t = 0$ for all $t \neq t_0$, and $\underline{\xi}_{t_0} > 0$. While $\underline{\xi}_{t_0}^* = \underline{\xi}_{t_0} > 0$, observe that $\underline{\xi}_{t_0+1}^* = -\phi\underline{\xi}_{t_0} < 0$ is decreasing at a point where regulation is not imposed. Moreover, consider the process $L(t) = T^{-1/2}\sum_{i=1}^{\lfloor Tt \rfloor} \underline{\xi}_t^*$. While $l(t)$ certainly increases at the point of regulation t_0, it in fact decreases at the point of no regulation, $t_0 + 1$, thereby violating the Harrison (1985) construction.

3.2.2 BI(1) Processes with Serial Correlation

The failure of the model of serial correlation above is a consequence of a broader result that the class of regulating processes is not in general closed under linear transformations. Here we propose the first of three alternative BI(1) processes with serial correlation that have correct limiting distributions. In particular, consider the system below.

$$y_t = \rho y_{t-1} + \psi(L)\Delta z_t \tag{3.6}$$

$$z_t = z_{t-1} + u_t \tag{3.7}$$

$$u_t = \epsilon_t + \underline{\xi}_{z,t} - \overline{\xi}_{z,t} \tag{3.8}$$

$$\rho = 1 - c_\rho T^{-1} \tag{3.9}$$

where T is the sample size and $c_\rho \in [0, 2)$ is the localization constant which interprets y_t as the near unit root process of Phillips (1987a) whenever $c_\rho > 0$ and $|\rho| < 1$, and the pure unit process when $c_\rho = 0$ and $\rho = 1$. Note further that whereas regulation in Cavaliere and Xu (2014) bounds y_t, the regulating mechanism in eq. (3.8) acts on z_t instead. In other words,

$$\underline{\xi}_{z,t} > 0 \quad \text{iff} \quad z_{t-1} + \epsilon_t < \underline{b} \tag{3.10}$$

$$\overline{\xi}_{z,t} > 0 \quad \text{iff} \quad z_{t-1} + \epsilon_t > \overline{b} \tag{3.11}$$

so that $z_t \in [\underline{b}, \overline{b}]$. The idea here is to regulate y_t by virtue of eq. (3.6) – a random walk driven by a linear combination of increments of a BI(1) process, namely $\psi(L)u_t = \psi(L)\Delta z_t$. Moreover, when $\rho = 1$, an application of the Beveridge-Nelson (BN) decomposition to y_t implies that $y_t \in [\underline{b}_y, \overline{b}_y] = \left[\min\left(\psi(1)\underline{b}, \psi(1)\overline{b}\right), \max\left(\psi(1)\underline{b}, \psi(1)\overline{b}\right)\right]$, where the extrema functions accommodate cases where $\psi(1) < 0$. Theorem 4 derives the asymptotics.

Theorem 4. Let y_t be generated by eq. (3.6) – eq. (3.9). Under assumptions $\mathcal{A}$ and $\mathcal{B}$ with $d = 0$, if $\mathcal{D}[0,1]$ is endowed with the Skorohod topology and $y_0 \in [\underline{b}_y, \overline{b}_y]$, then, for $t \in [0,1]$,

$$y_T(t) = T^{-1/2}\left(y_{\lfloor Tt \rfloor} - y_0\right) \Longrightarrow \sigma\psi(1)J^{\underline{c},\overline{c}}(t)$$

as $T \to \infty$, where

$$J^{\underline{c},\overline{c}}(t) = B^{\underline{c},\overline{c}}(t) - c_\rho \int_0^t e^{-c_\rho(t-r)} B^{\underline{c},\overline{c}}(r)dr$$

The result holds for all $|\rho| \leq 1$. In particular, when $c_\rho = 0$, the expected unit root asymptotics hold with $y_T(t) \Longrightarrow \sigma\psi(1)B^{\underline{c},\overline{c}}(t)$. In contrast, when $c_\rho > 0$, near unit root asymptotics apply and $J^{\underline{c},\overline{c}}(t)$ denotes a regulated Ornstein-Uhlenbeck (OU) process with lower and upper time varying bounds $\underline{c}(1 - c_\rho \int_0^t e^{-c_\rho(t-r)}dr)$ and $\overline{c}(1 - c_\rho \int_0^t e^{-c_\rho(t-r)}dr)$, respectively; see Theorem 4 in Cavaliere (2005a).

We also consider the hypothesis pair $H_0 : \rho = 1$ versus $H_1 : |\rho| < 1$ to test for the presence of a unit root. As a testing mechanism, we modify the family of non-parametric variance ratio (VR) tests in Nielsen (2009a) to reflect the presence of regulation. The test statistic is a scaled ratio of the sample variances of the regulated process y_t and its fractional partial sum

$$\tilde{y}_t = \Delta_+^{-d_1} y_t = (1-L)_+^{-d_1} y_t = \sum_{k=0}^{t-1} \frac{\Gamma(d_1+k)}{\Gamma(d_1)\Gamma(k+1)} y_{t-k} = \sum_{k=0}^{t-1} \pi_k(d_1) y_{t-k}$$

where $d_1 > 0$ and $\triangle_+^{-d}$ is the truncated version of the binomial expansion in the lag operator L. The following result establishes the asymptotics $\widetilde{y}_T(t)$.

Theorem 5. Let y_t be generated by eq. (3.6) – eq. (3.9). Under assumptions $\mathcal{A}$ and $\mathcal{B}$ with $d = 0$, if $\mathcal{D}[0,1]$ is endowed with the Skorohod topology, $d_1 > -1/2$, and $y_0 \in [\underline{b}_y, \overline{b}_y]$, then, for $t \in [0,1]$,

$$\widetilde{y}_T(t) = T^{-1/2-d_1}\left(\tilde{y}_{\lfloor Tt \rfloor} - \tilde{y}_0\right) \Longrightarrow \sigma\psi(1)\widetilde{J}^{\underline{c},\overline{c}}(t,d_1)$$

as $T \to \infty$, where

$$\widetilde{J}^{\underline{c},\overline{c}}(t,d_1) = \frac{1}{\Gamma(d_1+1)} \int_0^t (t-s)^{d_1} dJ^{\underline{c},\overline{c}}(t)(s)$$

$\widetilde{J}^{\underline{c},\overline{c}}(t,d_1)$ is essentially a fractionally filtered regulated OU process which collapses to the fractionally filtered BI(1) process when $c_\rho = 0$. The result is also necessary in establishing the limiting distribution of the fractional variance ratio test, to which we turn next.

To reflect the presence of bounds, the fractional variance ratio statistic is augmented with bound parameters $\underline{b}$ and $\bar{b}$. It is formalized below.

$$\tau^{\underline{b},\bar{b}}(d_1) = T^{2d_1} \frac{\sum_{t=0}^{T} y_t^2}{\sum_{t=0}^{T} \tilde{y}_t^2} \tag{3.12}$$

When $\left(\underline{b}, \bar{b}\right) = (-\infty, \infty)$, $\tau^{\underline{b},\bar{b}}(d_1)$ clearly collapses to the VR statistic of Nielsen (2009a) and in general shares its appeal. In particular, the statistic does not require the estimation of the long-run variance of y_t, it has correct asymptotic size, and consistently discriminates between the null and alternative hypotheses; see also Müller (2008). Moreover, 5 shows that asymptotic distributions of $\tau^{\underline{b},\bar{b}}(d_1)$ under the null and alternative hypotheses are indexed by d_1. Accordingly, d_1 is not a tuning parameter, unlike lag length and bandwidth parameters in Dickey-Fuller and Phillips-Perron statistics.

Since $\tau^{\underline{b},\bar{b}}(d_1)$ has correct asymptotic size, it is natural to investigate whether choices of d_1 can improve power. We approach the problem through local asymptotic power. In particular, we establish limiting distributions for $\tau^{\underline{b},\bar{b}}(d_1)$ under the battery of local to unity hypotheses $H : c_\rho \in [0, 2]$ which collapse to $H_0 : \rho = 1$ when $c_\rho = 0$, and $H_1 : |\rho| < 1$ when $c_\rho > 0$.

Theorem 6. Let y_t be generated by eq. (3.6) – eq. (3.9). Under assumptions $\mathcal{A}$ and $\mathcal{B}$ with $d = 0$, if $\mathcal{D}[0, 1]$ is endowed with the Skorohod topology, $d_1 > -1/2$, and $y_0 \in [\underline{b}_y, \bar{b}_y]$, then, for $t \in [0, 1]$ and $T \to \infty$,

$$\tau^{\underline{b},\bar{b}}(d_1) \Longrightarrow \frac{\int_0^1 J^{\underline{c},\bar{c}}(s)^2 ds}{\int_0^1 \widetilde{J}^{\underline{c},\bar{c}}(s, d_1)^2 ds}$$

Simulation analysis in Section 3.3 will demonstrate that lower values of d_1 yield uniformly (in c_ρ) higher asymptotic local power at the cost of severe size distortions when d_1 is very close to zero. Accordingly, choosing $d_1 = 0.1$ seems like a reasonable compromise; see Nielsen (2009a).

3.2.3 RFI($d+1$) Processes with Serial Correlation

Our second model considers an alternative to linear stationary forms of serial correlation. In particular, we use fractional integration to generate a continuum of persistence effects intermediate between transience and permanence. This renders a broad class of autocorrelation functions consistent with stationarity but slower to decay than those associated with the ARMA class; see Baillie (1996). When used to drive the serial dependence of innovations in a model of regulated integration, the model collapses to the regulated fractionally integrated process RFI($d+1$). To also allow for linear trends, we augment the model with possibly time varying deterministic components. Specifically, consider the model below.

$$x_t = \gamma\delta_t + y_t \tag{3.13}$$

$$y_t = \rho y_{t-1} + u_t \tag{3.14}$$

$$u_t = \triangle_+^{-d}\epsilon_t + \underline{\xi}_t - \overline{\xi}_t, \quad d > -1/2 \tag{3.15}$$

$$\rho = 1 - c_\rho T^{-1}, \quad c_\rho \in [0, 2) \tag{3.16}$$

When $\gamma = [\gamma_0, \gamma_1]$ – a 1×2 vector – and conformably, $\delta_t = [1, t]^\top$, the $\gamma\delta_t$ term in eq. (3.13) captures common deterministic specifications. For instance, whenever $\gamma = 0$, x_t reduces to y_t without deterministic dynamics. Alternatively, whenever $\gamma_1 = 0$, eq. (3.13) reduces to a RFI($d+1$) process with a mean. Lastly, when both γ_0 and γ_1 are nonzero, x_t models a RFI$(d+1)$ process with mean and trend components. The latter case is particularly important since regulated processes with trends have received little formal discussion. The aim is to define a trend which neither dominates nor is dominated by the bounds. For instance, an upper bound imposed on a positively trending process will render the process constant at the bound. Alternatively, imposing a lower bound on a positively trending process will prove ineffective. To overcome these limitations, we explore the suggestion in Cavaliere and Xu (2014) to model the trend through the localized trend $\gamma_1 = T^{(d-1/2)}c_\gamma$, for some localization constant c_γ. This is interpreted as a scaled continuum of trends on $\mathbb{R}$ with $c_\gamma = 0$ as the special case of no trend. Scaling by $T^{(d-1/2)}$ now ensures that y_t and γ_1 are of identical orders of magnitude, effectively merging the impact of the local trend with the integrated process itself; see Haldrup and Hylleberg (1995). Since $\gamma\delta_t$ no longer dominates y_t, the bounded process is not subject to the aforementioned drawbacks. We therefore impose the following assumption.

Assumption 4. $\mathcal{C}$

1. $\gamma_1 = c_\gamma T^{(d-1/2)}$ where $c_\gamma \in \mathbb{R}$ is a fixed constant.

It was shown in Trokić (2013) that under assumptions $\mathcal{A}$, when $\gamma_1 = 0$ ($c_\gamma = 0$) and $\rho = 1$, the càdlàg process $x_T(t) = \left(\frac{\sigma}{\Gamma(d+1)}T^{(d+1/2)}\right)^{-1}\left(x_{\lfloor Tt \rfloor} - x_0\right)$ converges in distribution to $B^{\underline{c},\bar{c}}_{d+1}(t)$ – the type II fractional Brownian motion

$$B_{d+1}(t) = \int_0^t (t-s)^d dW(s)$$

where $W(t)$ denotes the standard Wiener process. Here we generalize this result to possibly non zero trend cases, $c_\gamma \in \mathbb{R}$. Furthermore, we simplify the exposition by maintaining that deterministic components $\gamma\delta_t$ are known *a priori*. We will relax this assumption in Section 3.2.4.

Theorem 7. Let x_t be generated by eq. (3.13) – eq. (3.16). Under assumptions $\mathcal{A}$ and $\mathcal{C}$, if $\mathcal{D}[0,1]$ is endowed with the Skorohod topology, $d, d_1 > -1/2$, and $x_0 \in [\gamma\delta_t + \underline{b}, \gamma\delta_t + \overline{b}]$, then, for $t \in [0,1]$ and $T \to \infty$,

1. $x_T(t) \Longrightarrow J^{\underline{c},\overline{c}}_{d+1}(t) + c_\gamma t$
2. $\tilde{x}_T(t) = \Delta_+^{-d_1} x_T(t) \Longrightarrow \widetilde{J}^{\underline{c},\overline{c}}_{d+1}(t, d_1) + \frac{c_\gamma s^{d_1+1}}{\Gamma(d_1+2)}$

where,

$$J^{\underline{c},\overline{c}}_{d+1}(t) = B_{d+1}(t) + L(t) - U(t) - c\int_0^t e^{-c(t-s)}\left(B_{d+1}(s) + L(s) - U(s)\right) ds$$

$$\widetilde{J}^{\underline{c},\overline{c}}_{d+1}(t, d_1) = \frac{1}{\Gamma(d_1+1)}\int_0^t (t-s)^{d_1} dJ^{\underline{c},\overline{c}}_{d+1}(s)$$

and $L(t)$ and $U(t)$ are non-negative, non-decreasing processes which increase only when $J^{\underline{c},\overline{c}}_{d+1}(t) = \underline{c}$ and $J^{\underline{c},\overline{c}}_{d+1}(t) = \overline{c}$, respectively.

Above, $J^{\underline{c},\overline{c}}_{d+1}(t)$ and $\widetilde{J}^{\underline{c},\overline{c}}_{d+1}(t, d_1)$ are respectively the regulated OU process $J^{\underline{c},\overline{c}}(t)$ and its fractionally filtered variant $\widetilde{J}^{\underline{c},\overline{c}}(t, d_1)$, generalized to fractional orders of integration. Moreover, when testing for unit roots under Model 2, Theorem 7 implies that the limiting distribution of the VR statistic can be formalized as follows.

Lemma 2. Let x_t be generated by eq. (3.13) – eq. (3.16). Under assumptions $\mathcal{A}$ and $\mathcal{C}$, if $\mathcal{D}[0,1]$ is endowed with the Skorohod topology and $x_0 \in [\gamma\delta_t + \underline{b}, \gamma\delta_t + \overline{b}]$, then, for $t \in [0,1]$ and $T \to \infty$,

$$\tau_d^{b,\bar{b}}(d_1) = T^{2d_1} \frac{\sum_{t=0}^{T} x_t^2}{\sum_{t=0}^{T} \tilde{x}_t^2} \Longrightarrow \frac{\int\limits_0^1 \left(J_{d+1}^{\underline{c},\bar{c}}(t) + c_\gamma t \right)^2 ds}{\int\limits_0^1 \left(\widetilde{J}_{d+1}^{\underline{c},\bar{c}}(t, d_1) + \frac{c_\gamma s^{d_1+1}}{\Gamma(d_1+2)} \right)^2 ds}$$

Observe here that $\tau_d^{b,\bar{b}}(d_1)$ is constructed using the raw series x_t. That is, since the deterministic components $\gamma\delta_t$ are assumed known, detrending x_t is unnecessary.

3.2.4 Regulated Integration with Serial Correlation

We have already seen that even serial correlation of known linear form can invalidate the Harrison (1985) construction of regulated integration. This is because regulating functions are in general not homogeneous in the innovation process. Inversion of serial correlation is therefore ineffective as regulation is not preserved in the inversion. This suggests a shift in focus from serial correlation attenuation, to correct specification of the regulating functions. We therefore propose a method to consistently estimate the regulating mechanism even in the presence of complete agnosticism concerning the serial dependence structure of the innovation process.

Recall that any regulated diffusion $Z(t) \in [\underline{b}, \bar{b}]$ is associated with a diffusion $X(t)$, and uniquely determined regulating functions $L(t)$ and $U(t)$ satisfying the following properties:

1. $L(t)$ and $U(t)$ are continuous nondecreasing processes with $L(0) = U(0) = 0$;

2. $Z(t) = X(t) + L(t) - U(t) \in [\underline{b}, \bar{b}]$ for all $t \geq 0$

3. $L(t)$ and $U(t)$ increase only when $Z(t) = \underline{b}$ and $U(t) = \overline{b}$, respectively.

It can further be shown that $L(t)$ and $U(t)$ are equivalent in law to $L_Z(t, \underline{b})$ and $L_Z(t, \overline{b})$ – the *local times* of $Z(t)$ at $\underline{b}$ and $\overline{b}$, respectively; see Ikeda and Watanabe (2014). The local time $L_X(t, x)$ of $X(t)$ at x; roughly interpreted as the length of time, up to time t, the process $X(t)$ spends in the immediate neighborhood of x, is formalized as

$$L_X(t, x) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{\{|X(s)-x|<\varepsilon\}} ds$$

Moreover, for any bounded, measurable function $f(\cdot)$, the *occupation times formula* (cf. Revuz and Yor (1999)) $\int_0^t f(X(s))ds = \int_{\mathbb{R}} f(x) L_X(t, x) dx$ justifies the use of local times as densities; the latter being consistently estimable using nonparametric kernel estimation. This suggests that $L(t)$ and $U(t)$ can be consistently estimated from $Z(t)$ alone. The implication is particularly salient when the object of inferential interest is $X(t) = Z(t) - L(t) + U(t)$, but the observable process is $Z(t)$. With most empirical exercises involving regulated process, including unit root detection, this is indeed the case. Moreover, the methodology accommodates arbitrary forms of serial dependence in the innovation process driving $X(t)$, and is also readily adapted to discrete time processes. Consider the specification below.

$$x_t = \gamma\delta_t + y_t \tag{3.17}$$

$$y_t = y_{t-1} + \Delta S_t + \underline{\xi}_t - \overline{\xi}_t \tag{3.18}$$

$$S_t = \rho S_{t-1} + u_t \tag{3.19}$$

$$u_t = \psi(L)\epsilon_t \tag{3.20}$$

$$\rho = 1 - c_\rho T^{-1}, \quad c_\rho \in [0, 2) \tag{3.21}$$

Here $\psi(L)$ generates serial correlation as in Section 3.2.1, while deterministic components $\gamma\delta_t$ were introduced in Section 3.2.3. Note that in the absence of serial correlation and deterministic components, the specification collapses to the Cavaliere (2005a) representation of regulated integration. Here we adapt to the presence of both $\gamma\delta_t$ and $\psi(L)$. In particular, we define some 'true' lower and upper integrated regulators $L_t = \sum_{i=i}^{t} \underline{\xi}_t$ and $U_t = \sum_{i=i}^{t} \underline{\xi}_t$, respectively, to ensure $y_t \in [\underline{b}, \overline{b}]$ and $x_t \in [\gamma\delta_t + \underline{b}, \gamma\delta_t + \overline{b}]$, even when $\psi(L) \neq 1$ and $\gamma\delta_t \neq 0$. While deterministic dynamics are readily handled in assumptions $\mathcal{A}.5$ and $\mathcal{C}$, the presence of serial correlation is inherently reflected in the variance of the limiting distribution of $y_T(t)$ (or $x_T(t)$), which is some function of $\lfloor Tt \rfloor$ when $\psi(L) \neq 1$; see Ng and Perron (2001). In this regard, it is readily verified that $y_T(t)$ (or $x_T(t)$) converges to a Gaussian process and the scaled càdlàg approximants $L_T(t)$ and $U_T(t)$ of L_t and U_t, respectively, converge in law to their respective processes $L(t)$ and $U(t)$, which increase only when $y_T(t) = \underline{c}$ and $y_T(t) = \overline{c}$ (or $x_T(t) = \underline{c} + c_\gamma t$ and $x_T(t) = \overline{c} + c_\gamma t$), respectively. Unfortunately, these 'true' regulators are typically unobservable and serial correlation correction is rendered inadmissible as argued in Section 3.2.1 in the case of the BN decomposition. Fortunately, these processes are consistently estimable.

For any integrated process x_t, if there exists a continuous Gaussian process $G_x(t)$ such that $T^{-1/2}x_{\lfloor Tt \rfloor} \Longrightarrow G_x(t)$ on $D[0,1]$ as $T \to \infty$, Wang et al. (2009) have shown that

$$\frac{1}{\sqrt{T}h}\sum_{k=0}^{\lfloor Tt \rfloor} K\left(\frac{T^{-1/2}\left(x_{\lfloor k \rfloor} - a\right)}{h}\right) \Longrightarrow L_G(t,\alpha) \tag{3.22}$$

where $T^{-1/2}a = \alpha + o(1)$ is a constant and $K(x)$ and h are respectively the kernel function and bandwidth parameter which satisfy the following assumptions:

Assumption 5. $\mathcal{D}$

1. The kernel function K is a symmetric and continuous density function with support $[-1,1]$.

2. The bandwidth parameter h satisfies $h \longrightarrow 0$ and $nh \longrightarrow \infty$.

Since $y_T(t)$ converges to a Gaussian process and $L_T(t)$ and $U_T(t)$ respectively converge to the local times of $y_T(t)$ at points $\underline{c}$ and $\overline{c}$, the result in eq. (3.22) guarantees that

$$\frac{1}{\sqrt{T}}\widehat{L}_y(t,\underline{b}) := \frac{1}{\sqrt{T}h}\sum_{k=0}^{\lfloor Tt \rfloor} K\left(\frac{y_{\lfloor k \rfloor} - T^{-1/2}\underline{b}}{h}\right) \Longrightarrow L_y(t,\underline{c})$$
$$\frac{1}{\sqrt{T}}\widehat{L}_y(t,\overline{b}) := \frac{1}{\sqrt{T}h}\sum_{k=0}^{\lfloor Tt \rfloor} K\left(\frac{y_{\lfloor k \rfloor} - T^{-1/2}\overline{b}}{h}\right) \Longrightarrow L_y(t,\overline{c})$$

Accordingly, $\widehat{L}_y(t,\underline{b})$ and $\widehat{L}_y(t,\overline{b})$ are consistent estimators of L_t and U_t, respectively. This naturally leads to consistent estimates of the unregulated integrated process S_t, X_t – the process S_t augmented by deterministic dynamics, and $\widehat{\widehat{X}}_t\left(\underline{b},\overline{b}\right)$ – the fractional partial sum transform of X_t.

$$\widehat{S}_t\left(\underline{b},\overline{b}\right) = y_t - \widehat{L}_y(t,\underline{b}) + \widehat{L}_y(t,\overline{b}) \tag{3.23}$$

$$\widehat{X}_t\left(\underline{b},\overline{b}\right) = \widehat{S}_t\left(\underline{b},\overline{b}\right) + \gamma\delta_t \tag{3.24}$$

$$\widehat{\widehat{X}}_t\left(\underline{b},\overline{b}\right) = \Delta_+^{-d_1}\widehat{X}_t\left(\underline{b},\overline{b}\right) \tag{3.25}$$

The following result establishes the limiting distributions for $\widehat{X}_t$ and $\widehat{\widehat{X}}_t$. We suppress the notation $\left(\underline{b},\overline{b}\right)$ since no confusion can arise.

Theorem 8. Let x_t be generated by Equation (3.17) – Equation (3.21). Under assumptions $\mathcal{A}$, $\mathcal{C}$ and $\mathcal{D}$, if $\mathcal{D}[0,1]$ is endowed with the Skorohod topology, $d_1 > -1/2$, and $x_0 \in [\gamma\delta_t + \underline{b}, \gamma\delta_t + \overline{b}]$, then, for $t \in [0,1]$, as $T \to \infty$,

1. $\widehat{X}_T(t) = \left(\sigma\psi(1)\sqrt{T}\right)^{-1}\left(\widehat{X}_{\lfloor Tt\rfloor} - \widehat{X}_0\right) \Longrightarrow J(t) + c_\gamma t$
2. $\widehat{\widehat{X}}_T(t) = \left(\frac{\sigma\psi(1)}{\Gamma(d_1+1)}T^{(d_1+1/2)}\right)^{-1}\left(\widehat{\widehat{X}}_{\lfloor Tt\rfloor} - \widehat{\widehat{X}}_0\right) \Longrightarrow \widetilde{J}(t,d_1) + \frac{c_\gamma s^{d_1+1}}{\Gamma(d_1+2)}$

where

$$J(t) = W(t) - c_\rho\int_0^t e^{-c_\rho(t-r)}W(r)dr$$
$$\widetilde{J}(t,d_1) = W_{d_1}(t) - c_\rho\int_0^t e^{-c_\rho(t-r)}W_{d_1}(r)dr$$

We exploit this result to conduct a test for unit roots in the model presented in Equation (3.17) – Equation (3.21). As before, we use the VR statistic to conduct a battery of local to unity hypotheses $H : c_\rho \in [0,2]$. Since the limiting distributions in Theorem 8 are standard OU processes without regulation, the limiting distribution of the VR statistic collapses to the limiting distribution shown in Nielsen (2009a). The following lemma formalizes this result.

Lemma 3. Let x_t be generated by Equation (3.17) – Equation (3.21). Under assumptions $\mathcal{A}$, $\mathcal{C}$ and $\mathcal{D}$, if $\mathcal{D}[0,1]$ is endowed with the Skorohod topology, $d_1 > -1/2$, and $x_0 \in [\gamma\delta_t + \underline{b}, \gamma\delta_t + \overline{b}]$, then, for $t \in [0,1]$, as $T \longrightarrow \infty$,

$$\tau(d_1) = T^{2d_1} \frac{\sum_{t=0}^{T} \widehat{X}_t^2}{\sum_{t=0}^{T} \widehat{\widetilde{X}}_t^2} \Longrightarrow \frac{\int\limits_0^1 \left(J(t) + c_\gamma t\right)^2 ds}{\int\limits_0^1 \left(\widetilde{J}(t, d_1) + \frac{c_\gamma s^{d_1+1}}{\Gamma(d_1+2)}\right)^2 ds}$$

3.3 Simulation Analysis

We next assess size and power properties of the regulated unit root tests proposed in Sections 3.2.2 to 3.2.4. In particular, finite sample performance and size-corrected local power of $\rho^{\underline{b},\overline{b}}(d_1)$, $\rho_d^{\underline{b},\overline{b}}(d_1)$ and $\rho(d_1)$ is presented and contrasted with other unit root tests in the literature – the statistics proposed in Cavaliere and Xu (2014) and standard Nielsen's (2009) VR test. All simulations are performed for sample sizes $T = \{100, 500\}$ over 10,000 MC replications and critical values derived at significance level $\alpha = 0.05$. Moreover, while our tests are designed to yield uniformly (in c_ρ) higher power as d_1 decreases, since size distortions are imminent when d_1 is very close to 0, we follow Nielsen (2009a) and fix $d_1 = 0.1$. For the tests in Section 3.2.2 and Section 3.2.3, we consider both symmetric bounds, where $-\underline{c} = \overline{c} = c_0 \in \{\infty, 0.8, 0.4\}$, and asymmetric bounds $\overline{c} = \infty$ and $\underline{c} = c_0 \in \{0, -0.4, -0.8\}$. Whereas, because of very high computational cost in Section 3.2.4, we utilize only symmetric bounds, where $-\underline{b} = \overline{b} = b_0 \in \{2.5, 15\}$. All tables for this chapter can be found in appendix.

We begin with the the model in Section 3.2.1 and generate the data according to Equation (3.6)-Equation (3.9). Table 2 demonstrates the size, finite sample power and size adjusted power of bounded VR test and bounded MZ_t test when we have two sided symmetric bounds and no serial correlation in innovations. In small sample (right panel of Table 2), VR test has better size properties, but also better finite sample power than MZ_t test of Cavaliere and Xu (2014). By means of size adjusted power, they are almost equivalent. This equivalence is broken when we have tighter bounds. In the left panel, we can observe the large sample properties. Both tests performs similar under all scenarios. However, we still have advantage in small samples.

Table 2: **Size and Power Comparison for Section 3.2.2: Symmetric Bounds and No Serial Correlation**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	Inf	1	0.0479	0.05	0.0402	0.05	500	Inf	1	0.0488	0.05	0.0482	0.05
100	Inf	0.95	0.1675	0.1779	0.0802	0.1034	500	Inf	0.95	0.9069	0.9069	0.9664	0.9642
100	Inf	0.9	0.4069	0.4001	0.2355	0.285	500	Inf	0.9	0.9994	0.9994	1	1
100	Inf	0.85	0.6379	0.654	0.5145	0.5791	500	Inf	0.85	1	1	1	1
100	Inf	0.8	0.8242	0.84	0.7683	0.8436	500	Inf	0.8	1	1	1	1
100	0.8	1	0.0524	0.05	0.0331	0.05	500	0.8	1	0.0504	0.05	0.0454	0.05
100	0.8	0.95	0.1014	0.1	0.0274	0.0472	500	0.8	0.95	0.8897	0.8876	0.8897	0.9018
100	0.8	0.9	0.2747	0.2692	0.1087	0.1626	500	0.8	0.9	1	1	1	1
100	0.8	0.85	0.5299	0.5347	0.2567	0.398	500	0.8	0.85	1	1	1	1
100	0.8	0.8	0.78	0.774	0.5477	0.6937	500	0.8	0.8	1	1	1	1
100	0.4	1	0.0549	0.05	0.0156	0.05	500	0.4	1	0.0488	0.05	0.0376	0.05
100	0.4	0.95	0.0856	0.0767	0.0139	0.068	500	0.4	0.95	0.9193	0.9369	0.8562	0.9259
100	0.4	0.9	0.2308	0.2181	0.0416	0.1924	500	0.4	0.9	1	1	1	1
100	0.4	0.85	0.5371	0.5032	0.1352	0.4711	500	0.4	0.85	1	1	1	1
100	0.4	0.8	0.8449	0.8122	0.3245	0.7652	500	0.4	0.8	1	1	1	1

Table 3: **Size and Power Comparison for Section 3.2.2: Symmetric Bounds and AR(1) Model**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	Inf	1	0.0169	0.05	0.0488	0.05	500	Inf	1	0.0412	0.05	0.0496	0.05
100	Inf	0.95	0.0674	0.1582	0.1084	0.1063	500	Inf	0.95	0.8161	0.8592	0.8515	0.8562
100	Inf	0.9	0.1562	0.3262	0.2052	0.2068	500	Inf	0.9	0.9949	0.9975	0.9515	0.9524
100	Inf	0.85	0.3254	0.5326	0.3057	0.3023	500	Inf	0.85	0.9998	0.9999	0.9424	0.9437
100	Inf	0.8	0.4692	0.7084	0.3286	0.3321	500	Inf	0.8	1	1	0.9238	0.9256
100	0.8	1	0.005	0.05	0.022	0.05	500	0.8	1	0.0161	0.05	0.0169	0.05
100	0.8	0.95	0.0266	0.1509	0.0396	0.081	500	0.8	0.95	0.7273	0.8778	0.6635	0.8441
100	0.8	0.9	0.0875	0.326	0.0983	0.1691	500	0.8	0.9	0.9927	0.9989	0.9301	0.9507
100	0.8	0.85	0.2247	0.5256	0.1734	0.2509	500	0.8	0.85	0.9998	1	0.921	0.9373
100	0.8	0.8	0.4117	0.6989	0.2249	0.2831	500	0.8	0.8	1	1	0.9064	0.9192
100	0.4	1	0.0006	0.05	0.0131	0.05	500	0.4	1	0.004	0.05	0.0115	0.05
100	0.4	0.95	0.0033	0.0923	0.0117	0.0497	500	0.4	0.95	0.4221	0.8257	0.4259	0.7602
100	0.4	0.9	0.0145	0.2117	0.0292	0.1041	500	0.4	0.9	0.9719	0.9991	0.8701	0.9319
100	0.4	0.85	0.055	0.3994	0.0701	0.1751	500	0.4	0.85	1	1	0.8935	0.9276
100	0.4	0.8	0.1554	0.6137	0.1028	0.2093	500	0.4	0.8	1	1	0.8906	0.9204

Table 4: **Size and Power Comparison for Section 3.2.2: Symmetric Bounds and MA(1) Model**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	Inf	1	0.0323	0.05	0.0527	0.05	500	Inf	1	0.045	0.05	0.0492	0.05
100	Inf	0.95	0.1246	0.1746	0.0941	0.0884	500	Inf	0.95	0.8849	0.8994	0.8322	0.8318
100	Inf	0.9	0.2751	0.3717	0.1641	0.1608	500	Inf	0.9	0.9992	0.9992	0.9416	0.9424
100	Inf	0.85	0.5031	0.6042	0.2758	0.2438	500	Inf	0.85	1	1	0.9252	0.926
100	Inf	0.8	0.7015	0.7831	0.3455	0.3169	500	Inf	0.8	1	1	0.9141	0.9133
100	0.8	1	0.0131	0.05	0.0298	0.05	500	0.8	1	0.0208	0.05	0.0253	0.05
100	0.8	0.95	0.0573	0.1534	0.0335	0.0572	500	0.8	0.95	0.8039	0.8886	0.6315	0.7613
100	0.8	0.9	0.1741	0.3452	0.0752	0.1112	500	0.8	0.9	0.9978	0.9994	0.9135	0.9282
100	0.8	0.85	0.3709	0.5716	0.1345	0.1772	500	0.8	0.85	1	1	0.9065	0.9171
100	0.8	0.8	0.5984	0.7622	0.2062	0.2286	500	0.8	0.8	1	1	0.8913	0.8979
100	0.4	1	0.0114	0.05	0.013	0.05	500	0.4	1	0.0207	0.05	0.0172	0.05
100	0.4	0.95	0.0159	0.07	0.0155	0.0555	500	0.4	0.95	0.6487	0.8301	0.4189	0.7061
100	0.4	0.9	0.0492	0.1746	0.03	0.0933	500	0.4	0.9	0.9996	0.9999	0.881	0.9265
100	0.4	0.85	0.1692	0.3993	0.056	0.1514	500	0.4	0.85	1	1	0.8998	0.9237
100	0.4	0.8	0.3741	0.637	0.0739	0.196	500	0.4	0.8	1	1	0.9004	0.9152

Table 5: **Size and Power Comparison for Section 3.2.2: Asymmetric Bounds and No Serial Correlation**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	0	1	0.0513	0.05	0.0172	0.05	500	0	1	0.0519	0.05	0.0408	0.05
100	0	0.95	0.091	0.0943	0.0157	0.0704	500	0	0.95	0.7053	0.6954	0.6136	0.6903
100	0	0.9	0.2101	0.2204	0.0409	0.1621	500	0	0.9	0.9892	0.99	0.9995	0.9999
100	0	0.85	0.4194	0.3985	0.101	0.3292	500	0	0.85	0.9999	0.9999	1	1
100	0	0.8	0.5876	0.5967	0.1914	0.5545	500	0	0.8	1	1	1	1
100	-0.8	1	0.0472	0.05	0.0328	0.05	500	-0.8	1	0.0514	0.05	0.046	0.05
100	-0.8	0.95	0.1337	0.1319	0.0455	0.0704	500	-0.8	0.95	0.8944	0.8919	0.9182	0.9334
100	-0.8	0.9	0.3239	0.3241	0.149	0.2124	500	-0.8	0.9	0.9996	0.9996	1	1
100	-0.8	0.85	0.5732	0.5811	0.3762	0.4799	500	-0.8	0.85	1	1	1	1
100	-0.8	0.8	0.8	0.7974	0.6361	0.7631	500	-0.8	0.8	1	1	1	1
100	-0.4	1	0.0483	0.05	0.0211	0.05	500	-0.4	1	0.0512	0.05	0.0397	0.05
100	-0.4	0.95	0.0989	0.0992	0.0254	0.0767	500	-0.4	0.95	0.7707	0.7836	0.8057	0.8634
100	-0.4	0.9	0.2233	0.2486	0.0726	0.2036	500	-0.4	0.9	0.9973	0.9974	1	1
100	-0.4	0.85	0.474	0.477	0.2123	0.4355	500	-0.4	0.85	1	1	1	1
100	-0.4	0.8	0.6632	0.6979	0.4154	0.7072	500	-0.4	0.8	1	1	1	1

Table 6: **Size and Power Comparison for Section 3.2.2: Asymmetric Bounds and AR(1) Model**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	0	1	0.0056	0.05	0.0331	0.05	500	0	1	0.0202	0.05	0.0185	0.05
100	0	0.95	0.0075	0.0882	0.0526	0.0864	500	0	0.95	0.4525	0.6589	0.3893	0.6308
100	0	0.9	0.0199	0.1741	0.0953	0.1431	500	0	0.9	0.8977	0.9662	0.7896	0.8956
100	0	0.85	0.0435	0.2953	0.1549	0.2074	500	0	0.85	0.9874	0.9969	0.8614	0.9095
100	0	0.8	0.0849	0.4133	0.1854	0.2556	500	0	0.8	0.9984	0.9997	0.8647	0.9021
100	-0.8	1	0.0005	0.05	0.0328	0.05	500	-0.8	1	0.0061	0.05	0.0176	0.05
100	-0.8	0.95	0.0046	0.107	0.0657	0.089	500	-0.8	0.95	0.5167	0.8225	0.5533	0.7452
100	-0.8	0.9	0.0125	0.2222	0.1151	0.1496	500	-0.8	0.9	0.9555	0.9942	0.886	0.929
100	-0.8	0.85	0.0468	0.3796	0.1936	0.2253	500	-0.8	0.85	0.9981	0.9998	0.9107	0.9315
100	-0.8	0.8	0.1023	0.5521	0.2593	0.2976	500	-0.8	0.8	1	1	0.9018	0.9173
100	-0.4	1	0.0029	0.05	0.0362	0.05	500	-0.4	1	0.0139	0.05	0.0191	0.05
100	-0.4	0.95	0.0016	0.0703	0.0591	0.0741	500	-0.4	0.95	0.4378	0.7353	0.4614	0.7128
100	-0.4	0.9	0.0077	0.1669	0.0973	0.1308	500	-0.4	0.9	0.9392	0.9861	0.8694	0.9228
100	-0.4	0.85	0.0221	0.3063	0.1638	0.1961	500	-0.4	0.85	0.9938	0.9988	0.8943	0.9239
100	-0.4	0.8	0.0615	0.476	0.2441	0.2594	500	-0.4	0.8	0.9995	1	0.8997	0.9226

Table 7: **Size and Power Comparison for Section 3.2.2: Asymmetric Bounds and MA(1) Model**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	0	1	0.0166	0.05	0.0165	0.05	500	0	1	0.0376	0.05	0.0226	0.05
100	0	0.95	0.0322	0.0983	0.0272	0.0824	500	0	0.95	0.6106	0.6674	0.3685	0.5568
100	0	0.9	0.0783	0.207	0.0441	0.132	500	0	0.9	0.9704	0.9813	0.7898	0.8654
100	0	0.85	0.1696	0.3593	0.0687	0.1941	500	0	0.85	0.9991	0.9993	0.8437	0.8819
100	0	0.8	0.278	0.5356	0.0873	0.2454	500	0	0.8	1	1	0.8482	0.8732
100	-0.8	1	0.0102	0.05	0.0353	0.05	500	-0.8	1	0.0188	0.05	0.0255	0.05
100	-0.8	0.95	0.0357	0.1212	0.0463	0.0605	500	-0.8	0.95	0.7355	0.8667	0.5886	0.7253
100	-0.8	0.9	0.1147	0.2859	0.082	0.0946	500	-0.8	0.9	0.9949	0.9981	0.887	0.9162
100	-0.8	0.85	0.2579	0.5043	0.1221	0.1459	500	-0.8	0.85	1	1	0.8928	0.9072
100	-0.8	0.8	0.4718	0.6976	0.1832	0.1967	500	-0.8	0.8	1	1	0.8792	0.8887
100	-0.4	1	0.015	0.05	0.0255	0.05	500	-0.4	1	0.03	0.05	0.0263	0.05
100	-0.4	0.95	0.0253	0.0855	0.0322	0.0621	500	-0.4	0.95	0.637	0.7589	0.4989	0.6829
100	-0.4	0.9	0.0635	0.1998	0.054	0.1034	500	-0.4	0.9	0.9858	0.9929	0.8638	0.9092
100	-0.4	0.85	0.1935	0.4008	0.0957	0.1596	500	-0.4	0.85	0.9998	0.9998	0.8874	0.9046
100	-0.4	0.8	0.3405	0.5896	0.1305	0.2185	500	-0.4	0.8	1	1	0.8832	0.9

When we introduce one sided bounds, both tests suffer some size distortion, but they are still not oversized (Table 5). Further, by means of power, our test has clear superiority against Cavaliere and Xu (2014) test.

Now, we also introduce serial correlation to innovations by letting them follow either $AR(1)$ process $u_t = \psi u_{t-1} + \epsilon_t$ or $MA(1)$ process $u_t = \theta\epsilon_{t-1} + \epsilon_t$ with $\theta = \psi = 0.5$. While Table 3-Table 4 illustrate the size and power properties under respectively $AR(1)$ and $MA(1)$ serial correlation for two sided bounds,

Table 6-Table 7 exhibits size and power for one sided bound case. In all these tables, we can notice that adding serial correlation does not change the previous result stating that our test has better finite sample properties than Cavaliere and Xu (2014) in almost all cases.

For our second specification, we use the data generation in Equation (3.13)-Equation (3.16) with $d \in 0.2, 0.4, -0.2$ and same bound specification utilized in previous section. Results for different scenarios are summarized in Table 8-Table 13. For all different bound specifications and different values for d parameter, our test produce uniformly better size and power performance relative to MZ_t test. Actually, this result is natural, since the standard spectral estimators used in MZ_t test or ADF type correction fails to capture true serial correlation structure. Combined with nonparametric long memory parameter estimators, our testing procedure is properly sized and capturing the absence of unit roots better than Cavaliere and Xu (2014)tests.

Table 8: **Size and Power Comparison for Section 3.2.3: Symmetric Bounds and $d = 0.2$ Model**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	Inf	1	0.0473	0.05	0.0304	0.05	500	Inf	1	0.0526	0.05	0.0237	0.05
100	Inf	0.95	0.1763	0.1769	0.0151	0.0323	500	Inf	0.95	0.7976	0.7929	0.2664	0.513
100	Inf	0.9	0.3774	0.371	0.0209	0.0543	500	Inf	0.9	0.9799	0.9802	0.8816	0.9618
100	Inf	0.85	0.5762	0.5683	0.0574	0.1197	500	Inf	0.85	0.9984	0.9984	0.9575	0.9799
100	Inf	0.8	0.7304	0.7295	0.118	0.2454	500	Inf	0.8	1	0.9999	0.9576	0.9832
100	0.8	1	0.048	0.05	0.062	0.05	500	0.8	1	0.0521	0.05	0.0618	0.05
100	0.8	0.95	0.1874	0.1833	0.0131	0.0068	500	0.8	0.95	0.7913	0.7827	0.276	0.1989
100	0.8	0.9	0.3558	0.3734	0.0224	0.0117	500	0.8	0.9	0.9809	0.9802	0.863	0.8056
100	0.8	0.85	0.5655	0.5734	0.0471	0.0198	500	0.8	0.85	0.9983	0.9983	0.9569	0.9417
100	0.8	0.8	0.7203	0.7324	0.115	0.056	500	0.8	0.8	1	1	0.9575	0.9402
100	0.4	1	0.0444	0.05	0.0585	0.05	500	0.4	1	0.0468	0.05	0.06	0.05
100	0.4	0.95	0.0524	0.0683	0.0087	0.0058	500	0.4	0.95	0.5569	0.6042	0.1944	0.09
100	0.4	0.9	0.1294	0.1601	0.0073	0.0041	500	0.4	0.9	0.9247	0.9351	0.8292	0.6226
100	0.4	0.85	0.2695	0.3156	0.0225	0.0089	500	0.4	0.85	0.9895	0.9913	0.9546	0.8955
100	0.4	0.8	0.4224	0.4906	0.0662	0.024	500	0.4	0.8	0.998	0.9984	0.9491	0.9029

Table 9: **Size and Power Comparison for Section 3.2.3: Symmetric Bounds and $d = 0.4$ Model**

T	c_0	ρ	VR	VRsa	MZT	MZTsa	T	c_0	ρ	VR	VRsa	MZT	MZTsa
100	Inf	1	0.0443	0.05	0.0652	0.05	500	Inf	1	0.05	0.05	0.155	0.05
100	Inf	0.95	0.1697	0.1747	0.0164	0.0108	500	Inf	0.95	0.6605	0.6636	0.0005	0
100	Inf	0.9	0.3302	0.3377	0.0041	0.0021	500	Inf	0.9	0.8862	0.8909	0.0047	0
100	Inf	0.85	0.4833	0.4861	0.0054	0.0016	500	Inf	0.85	0.9544	0.9568	0.1239	0
100	Inf	0.8	0.6078	0.6294	0.0117	0.004	500	Inf	0.8	0.9808	0.982	0.5602	0
100	0.8	1	0.0481	0.05	0.1152	0.05	500	0.8	1	0.0488	0.05	0.2372	0.05
100	0.8	0.95	0.1737	0.1681	0.0185	0.0019	500	0.8	0.95	0.6515	0.66	0.0007	0
100	0.8	0.9	0.3213	0.3278	0.0053	0.0003	500	0.8	0.9	0.8811	0.8788	0.0051	0
100	0.8	0.85	0.4673	0.4823	0.0044	0	500	0.8	0.85	0.9578	0.9593	0.1196	0
100	0.8	0.8	0.6193	0.6198	0.0096	0.0001	500	0.8	0.8	0.9834	0.9821	0.5915	0
100	0.4	1	0.047	0.05	0.162	0.05	500	0.4	1	0.0458	0.05	0.2896	0.05
100	0.4	0.95	0.1328	0.1363	0.0216	0.0022	500	0.4	0.95	0.6007	0.6044	0.0005	0
100	0.4	0.9	0.2665	0.2685	0.0044	0	500	0.4	0.9	0.8602	0.8614	0.0056	0
100	0.4	0.85	0.3942	0.4053	0.0042	0	500	0.4	0.85	0.9433	0.9418	0.1246	0
100	0.4	0.8	0.5168	0.5292	0.0111	0	500	0.4	0.8	0.9706	0.9717	0.5584	0

Table 10: **Size and Power Comparison for Section 3.2.3: Symmetric Bounds and $d = -0.2$ Model**

T	c_0	ρ	VR	VRsa	MZT	MZTsa	T	c_0	ρ	VR	VRsa	MZT	MZTsa
100	Inf	1	0.0526	0.05	0.1035	0.05	500	Inf	1	0.0479	0.05	0.2264	0.05
100	Inf	0.95	0.1461	0.1486	0.2175	0.1033	500	Inf	0.95	0.9622	0.9689	0.9601	0.7865
100	Inf	0.9	0.3832	0.3639	0.4326	0.253	500	Inf	0.9	1	1	0.9963	0.961
100	Inf	0.85	0.689	0.6577	0.6053	0.4335	500	Inf	0.85	1	1	0.9984	0.9873
100	Inf	0.8	0.8746	0.8751	0.6836	0.5731	500	Inf	0.8	1	1	0.9994	0.9917
100	0.8	1	0.0706	0.05	0.0254	0.05	500	0.8	1	0.0642	0.05	0.1409	0.05
100	0.8	0.95	0.1386	0.118	0.0385	0.1048	500	0.8	0.95	0.7217	0.7294	0.5956	0.4857
100	0.8	0.9	0.2498	0.2517	0.0882	0.2027	500	0.8	0.9	0.9978	0.9984	0.8323	0.8173
100	0.8	0.85	0.429	0.4481	0.1626	0.3287	500	0.8	0.85	1	1	0.9082	0.9184
100	0.8	0.8	0.6092	0.6709	0.3052	0.4531	500	0.8	0.8	1	1	0.9442	0.9552
100	0.4	1	0.4483	0.05	0.0016	0.05	500	0.4	1	0.3611	0.05	0.1322	0.05
100	0.4	0.95	0.5223	0.0726	0.0023	0.0767	500	0.4	0.95	0.639	0.2474	0.2217	0.1701
100	0.4	0.9	0.5678	0.1255	0.0019	0.1215	500	0.4	0.9	0.8629	0.6529	0.3392	0.3662
100	0.4	0.85	0.6372	0.1775	0.002	0.1565	500	0.4	0.85	0.9716	0.9275	0.471	0.5549
100	0.4	0.8	0.6848	0.2598	0.0016	0.2166	500	0.4	0.8	0.9977	0.996	0.5901	0.6861

Table 11: **Size and Power Comparison for Section 3.2.3: Asymmetric Bounds and $d = 0.2$ Model**

T	c_0	ρ	VR	VRsa	MZT	MZTsa	T	c_0	ρ	VR	VRsa	MZT	MZTsa
100	0	1	0.0504	0.05	0.0145	0.05	500	0	1	0.0485	0.05	0.014	0.05
100	0	0.95	0.1283	0.1217	0.0269	0.1087	500	0	0.95	0.6674	0.6815	0.3272	0.7226
100	0	0.9	0.2633	0.2523	0.0585	0.2232	500	0	0.9	0.9573	0.959	0.8248	0.9426
100	0	0.85	0.4503	0.424	0.1286	0.3798	500	0	0.85	0.9956	0.9957	0.844	0.9583
100	0	0.8	0.6333	0.5991	0.2166	0.5017	500	0	0.8	0.9998	0.9998	0.8424	0.9579
100	-0.8	1	0.0536	0.05	0.046	0.05	500	-0.8	1	0.0534	0.05	0.0445	0.05
100	-0.8	0.95	0.1778	0.1801	0.0125	0.0157	500	-0.8	0.95	0.7804	0.7859	0.2894	0.33
100	-0.8	0.9	0.363	0.3619	0.0213	0.0235	500	-0.8	0.9	0.9796	0.9793	0.8684	0.8933
100	-0.8	0.85	0.5601	0.5709	0.0441	0.0542	500	-0.8	0.85	0.9983	0.9982	0.9527	0.9617
100	-0.8	0.8	0.7242	0.7298	0.113	0.1284	500	-0.8	0.8	0.9997	0.9997	0.9586	0.9655
100	-0.4	1	0.0516	0.05	0.0443	0.05	500	-0.4	1	0.049	0.05	0.0441	0.05
100	-0.4	0.95	0.0876	0.0919	0.0101	0.0109	500	-0.4	0.95	0.6559	0.6738	0.2344	0.2304
100	-0.4	0.9	0.2111	0.2193	0.0146	0.0139	500	-0.4	0.9	0.9553	0.9573	0.8605	0.8301
100	-0.4	0.85	0.3644	0.3891	0.0284	0.0279	500	-0.4	0.85	0.9927	0.9938	0.9514	0.9436
100	-0.4	0.8	0.5292	0.5623	0.0752	0.077	500	-0.4	0.8	0.9993	0.9995	0.9557	0.9472

Table 12: **Size and Power Comparison for Section 3.2.3: Asymmetric Bounds and $d = 0.4$ Model**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	0	1	0.0499	0.05	0.0337	0.05	500	0	1	0.0445	0.05	0.0233	0.05
100	0	0.95	0.0714	0.0734	0.0319	0.0554	500	0	0.95	0.2179	0.2561	0.1146	0.262
100	0	0.9	0.1059	0.1044	0.0439	0.0831	500	0	0.9	0.4795	0.5256	0.2772	0.5503
100	0	0.85	0.1253	0.1449	0.0633	0.1102	500	0	0.85	0.688	0.7306	0.4933	0.7406
100	0	0.8	0.1833	0.1911	0.0795	0.1473	500	0	0.8	0.8256	0.8411	0.614	0.7989
100	-0.8	1	0.0477	0.05	0.0306	0.05	500	-0.8	1	0.0539	0.05	0.0218	0.05
100	-0.8	0.95	0.1768	0.1818	0.0253	0.0416	500	-0.8	0.95	0.6464	0.6351	0.1023	0.2395
100	-0.8	0.9	0.3255	0.3414	0.0336	0.0578	500	-0.8	0.9	0.8976	0.882	0.4324	0.6512
100	-0.8	0.85	0.4874	0.4961	0.0412	0.0778	500	-0.8	0.85	0.9584	0.9509	0.7242	0.8809
100	-0.8	0.8	0.6	0.6209	0.053	0.0992	500	-0.8	0.8	0.981	0.976	0.8483	0.9218
100	-0.4	1	0.0459	0.05	0.0658	0.05	500	-0.4	1	0.0548	0.05	0.0561	0.05
100	-0.4	0.95	0.1494	0.1446	0.0222	0.0075	500	-0.4	0.95	0.6189	0.6058	0.0968	0.0505
100	-0.4	0.9	0.2795	0.2938	0.0286	0.0064	500	-0.4	0.9	0.8694	0.8583	0.4305	0.2788
100	-0.4	0.85	0.4413	0.4325	0.0409	0.0086	500	-0.4	0.85	0.9438	0.9402	0.7387	0.6068
100	-0.4	0.8	0.5682	0.5702	0.0465	0.0082	500	-0.4	0.8	0.9775	0.9769	0.8438	0.7686

Table 13: **Size and Power Comparison for Section 3.2.3: Asymmetric Bounds and $d = -0.2$ Model**

T	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**	**T**	c_0	ρ	**VR**	**VRsa**	**MZT**	**MZTsa**
100	0	1	0.0486	0.05	0.021	0.05	500	0	1	0.0423	0.05	0.089	0.05
100	0	0.95	0.4031	0.3968	0.1817	0.2894	500	0	0.95	0.9999	0.9999	0.9559	0.9327
100	0	0.9	0.7927	0.7868	0.4314	0.5319	500	0	0.9	1	1	0.9763	0.9618
100	0	0.85	0.9553	0.9515	0.5314	0.6296	500	0	0.85	1	1	0.9727	0.957
100	0	0.8	0.991	0.9904	0.5916	0.6506	500	0	0.8	1	1	0.9692	0.9534
100	-0.8	1	0.0456	0.05	0.0322	0.05	500	-0.8	1	0.0515	0.05	0.1017	0.05
100	-0.8	0.95	0.1213	0.1383	0.0791	0.1298	500	-0.8	0.95	0.9434	0.9358	0.8496	0.7428
100	-0.8	0.9	0.3506	0.3488	0.2057	0.272	500	-0.8	0.9	1	1	0.9694	0.9361
100	-0.8	0.85	0.5926	0.6136	0.3655	0.4433	500	-0.8	0.85	1	1	0.9908	0.9776
100	-0.8	0.8	0.7982	0.8246	0.5129	0.568	500	-0.8	0.8	1	1	0.9927	0.9864
100	-0.4	1	0.0558	0.05	0.0253	0.05	500	-0.4	1	0.0498	0.05	0.0855	0.05
100	-0.4	0.95	0.2567	0.2419	0.1201	0.2192	500	-0.4	0.95	0.996	0.9953	0.911	0.86
100	-0.4	0.9	0.5689	0.549	0.2936	0.4171	500	-0.4	0.9	1	1	0.9732	0.9556
100	-0.4	0.85	0.8197	0.8094	0.4626	0.5667	500	-0.4	0.85	1	1	0.9882	0.9803
100	-0.4	0.8	0.9467	0.9403	0.5627	0.6433	500	-0.4	0.8	1	1	0.9896	0.9836

The other noteworthy result from these simulation exercise is that both test suffer size distortions when we have negative d parameter and two sided bounds. Although this distortions are present for MZ_t test in all two sided bound specifications, our test is only influenced when bounds are tight. There may be two explanation for this issue: First, when $d = -0.2$ (Table 10) we introduce overdifferencing in innovations. This potentially creates overrejection under the null hypothesis, since the observed series looks like stationary process. Second, combined with tight bounds the impact of overdifferencing increases. To overcome this issue some bootstrapping methods can be employed.

Finally, we investigate the performance of the test introduces in Section 3.2.3. The data is generated by Equation (3.17)-Equation (3.21). As discussed earlier, we utilize only two sided symmetric bounds with $-\underline{b} = \overline{b} = b_0 \in \{2.5, 15\}$ because of high computational cost. Since we clear out the regulators L_T and U_t, the asymptotic parameter $\underline{c}$ and $\overline{c}$ are irrelevant for asymptotic distribution. We report the results for our test denoted as VR, Nielsen's VR test (2009) and MZ_t test of Cavaliere and Xu (2014) with three different innovation structure: 1) no serial correlation (Table 14), 2) positive MA(1) type serial correlation where $u_t = \epsilon_t + 0.3\epsilon_{t-1}$ (Table 16) and 3) negative MA(1) type of serial correlation $u_t = \epsilon_t - 0.2\epsilon_{t-1}$ (Table 16).

Table 14: **Size and Power Comparison for Section 3.2.4: Symmetric Bounds and No Serial Correlation**

T	c_0	ρ	VR	VRs	MZT	T	c_0	ρ	VR	VRs	MZT
100	2.5	1	0.0431	0.593	0.0307	500	2.5	1	0.0453	0.9999	0.0229
100	2.5	0.95	0.2217	0.941	0.1297	500	2.5	0.95	0.7953	1	0.3528
100	2.5	0.9	0.5073	0.9852	0.2869	500	2.5	0.9	0.9234	1	0.8072
100	2.5	0.85	0.7416	0.9939	0.4726	500	2.5	0.85	0.9576	1	0.9264
100	2.5	0.8	0.8779	0.9962	0.6227	500	2.5	0.8	0.9785	1	0.919
100	15	1	0.0531	0.053	0.0428	500	15	1	0.051	0.0586	0.0428
100	15	0.95	0.2372	0.2372	0.2843	500	15	0.95	0.9434	0.9435	0.9991
100	15	0.9	0.526	0.526	0.6893	500	15	0.9	0.9997	0.9997	0.9998
100	15	0.85	0.7727	0.7727	0.8752	500	15	0.85	1	1	0.9996
100	15	0.8	0.8975	0.8975	0.91	500	15	0.8	1	1	0.9997

Table 15: **Size and Power Comparison for Section 3.2.4: Symmetric Bounds and Negative MA(1) model**

T	b_0	ρ	VR	VRs	MZT	T	b_0	ρ	VR	VRs	MZT
100	2.5	1	0.0579	0.546	0.0639	500	2.5	1	0.0516	0.9994	0.0389
100	2.5	0.95	0.2871	0.9218	0.2029	500	2.5	0.95	0.8689	1	0.4093
100	2.5	0.9	0.6044	0.9792	0.366	500	2.5	0.9	0.9648	1	0.7615
100	2.5	0.85	0.8286	0.9924	0.4913	500	2.5	0.85	0.9875	1	0.8517
100	2.5	0.8	0.924	0.9961	0.5702	500	2.5	0.8	0.9943	1	0.8144
100	15	1	0.0639	0.0639	0.0577	500	15	1	0.0528	0.0531	0.0518
100	15	0.95	0.2894	0.2894	0.3199	500	15	0.95	0.9654	0.9654	0.9974
100	15	0.9	0.6079	0.6079	0.6669	500	15	0.9	0.9993	0.9993	0.9993
100	15	0.85	0.8384	0.8384	0.8164	500	15	0.85	1	1	0.9997
100	15	0.8	0.9445	0.9445	0.855	500	15	0.8	1	1	0.9998

Table 16: **Size and Power Comparison for Section 3.2.4: Symmetric Bounds and Positive MA(1) model**

T	$[\underline{b} = \bar{b} = b_0]$	ρ	**VR**	**VRs**	**MZT**	**T**	$[\underline{b} = \bar{b} = b_0]$	ρ	**VR**	**VRs**	**MZT**
100	2.5	1	0.0401	0.7151	0.0105	500	2.5	1	0.0426	1	0.0207
100	2.5	0.95	0.1932	0.9736	0.0197	500	2.5	0.95	0.7836	1	0.0803
100	2.5	0.9	0.4561	0.9945	0.0385	500	2.5	0.9	0.8897	1	0.1073
100	2.5	0.85	0.6604	0.998	0.0758	500	2.5	0.85	0.9246	1	0.2491
100	2.5	0.8	0.8124	0.9988	0.164	500	2.5	0.8	0.9444	1	0.6583
100	15	1	0.0487	0.0487	0.0435	500	15	1	0.046	0.0905	0.0807
100	15	0.95	0.2183	0.2183	0.254	500	15	0.95	0.9339	0.9369	0.9968
100	15	0.9	0.4913	0.4913	0.5532	500	15	0.9	0.9986	0.9986	0.9996
100	15	0.85	0.7257	0.7257	0.773	500	15	0.85	1	1	0.9999
100	15	0.8	0.86	0.86	0.8906	500	15	0.8	1	1	1

Our modified VR test has absolute advantage in all cases to both standard VR test and modified MZ_t test by means of size and power in small and large sample sizes. Furthermore, while MZ_t test still shows promise in terms of size and local power, standard Nielsen test totally fails when we have tight bounds. Results are very similar for different serial correlation structures, but with negative MA(1) innovations, both modified VR and MZ_t tests are slightly over-rejecting the null hypothesis when null is true. This result is also expected as in previous section simulations with $d = -0.2$.

To sum, we evaluate small sample properties of three sections seperately. Nonetheless, all sections share a common feature: our test is properly sized with better power relative to competing methods. Moreover, it is also seen that without further correction for serial correlation, our methodology provides better results. This may be stemed from the fact that in small samples regression type corrections are not performing well.

3.4 Conclusion

Regulated time series are of tremendous practical importance. They characterize a number of important time series and are particularly salient in macroeconomic literature. As demonstrated in Cavaliere (2005a), Cavaliere and Xu (2014), and Trokić (2013), these series have asymptotic distributions which are quite different than those characterizing unbounded series. Unfortunately, the unit root literature has only recently begun to take this into consideration. This article contributes in this direction by extending the fractional variance ratio statistic of Nielsen (2009a) to three different types of regulated $I(1)$ time series. This extension relies on the theory of regulated fractionally integrated series developed in Trokić (2013) which is used here to develop the regulated fractional variance ratio statistic along with its limiting distribution. Simulation evidence for local (asymptotic) power curves confirms the results of Nielsen (2009a) and illustrates that practitioners should exercise caution when applying the fractional variance ratio statistic to series which may be regulated in different forms. We believe that especially the third method can bring a new understanding in regulated time series, since, to our knowledge, this is the first paper which deals with estimation of regulators in discrete time unit root models.

CHAPTER 4

WAVELET VARIANCE RATIO TEST AND WAVESTRAPPING FOR DETERMINATION OF COINTEGRATION RANK

4.1 Introduction

After the seminal work of Granger (1981), researchers were extensively concerned with the cointegration tests to find the number of the long run relations among integrated variables. Generally, barring a few exceptions, most of these tests are devised in time domain. However, the frequency domain properties of the integrated variables also have consequential implications in statistical analysis. For instance, Granger (1966) had observed that such variables have typical power spectra in which low frequency components have dominating importance. Following Granger's (1966) observation, a few scholars used Fourier transform to construct spectral unit root tests such as in Choi and Phillips (1993), Robinson (1994) and Chambers et al. (2014), and frequency domain cointegration tests such as Levy (2002), Morana (2004) and Nielsen (2004). Additionally, Fan and Gencay (2010) and Trokic (2014) develop spectral approaches for testing unit roots by using wavelets. In these studies, authors utilize low frequency components of the economic variables to improve the existing unit root tests. In this regard, we expand the results of Fan and Gencay (2010) and Trokic (2014) to multivariate setting and propose a wavelet based nonparametric method for testing cointegration in fractionally integrated systems.

The proposed method combines the Nielsen's (2010) variance ratio cointegration test and the frequency domain tools, specifically wavelets. This new testing design has several advantages over the existing cointegration tests. First, it is a fully nonparametric test and does not necessitate the estimation of a parametric model as in Johansen (1988), or requires the selection of an optimal bandwidth as in Phillips and Ouliaris (1988). Second, the asymptotic distribution of the test is nuisance and tuning parameter free, and is invariant to the short run dynamics of the variables. Third, the test can be applicable to a wide range of fractionally integrated series as well as standard I (1) case. Finally, by utilizing the wavelets in the cointegration framework, we propose a multivariate version of the wavelet based bootsrapping method particularly known as *wavestrapping*. This method is also nonparametric.

In related recent studies, the use of spectral tools, such as wavelets, in unit root and cointegration analysis has attracted some attention. In particular, Fan and Gencay (2010) developed a wavelet based unit root test. Unlike Fourier transform, wavelets operate in both time and frequency domain and adapt themselves to capture nonstationarity features of variables across a wide range of frequencies

(Fan and Gencay, 2010). This makes the wavelet transform a proper instrument for nonstationary time series and also cointegration analysis. More recently still, Trokic (2014) proposes a wavelet version of Nielsen's (2009) variance ratio (VR hereafter) unit root test. Similar to the VR test, Trokic's (2014) test is also non-parametric and tuning parameter free and offers two advantages against the standard VR test: first, using wavelets, the test is strengthened against short run dynamics. Second, using wavestraping, size distortions in unit root tests are effectively removed.

In a recent work, Fernandez-Macho et al. (2013) develops a parametric cointegration test by using wavelet theory. This test relies on the discrepancy between coefficients of the regression with low pass and high pass filtered series. Hence, the procedure is based on the estimation of two separate parametric regressions. Although Fernandez-Macho et al. (2013) makes an important attempt to mesh cointegration and wavelets, it may possibly suffer from the problems associated with parametric models, such as consistent estimation of cointegrating vectors and the impacts of short run dynamics. To avoid these shortcomings, we try to extend the ideas in Trokic (2014) and Fan and Gencay (2009) in cointegration analysis to establish a fully non-parametric wavelet based cointegration test for general fractionally integrated processes.

The test proposed in this paper generally has good size and power features in small samples. Nevertheless, the cointegration tests are negatively effected by complicated short run dynamics in the system. An important cause of such dynamics is the highly negative MA roots in the innovation structure of common stochastic factors. This in fact has been a widely recognized problem in the unit root testing literature. Since testing for cointegration is a multivariate generalization of unit root testing, it is not surprising that the problems such as these persist in the cointegration literature as well. Moreover, from simulation evidence, we observe Nielsen's VR test (2010) suffers severe size distortions in the aforementioned case. Our wavelet based test can partially remove these size distortions. We attribute this feature to the ability of wavelets to approximately filter out the short run dynamics in the common stochastic factors. In this regard, in order to remove remaining size distiortions, we design an additional wavelet based bootstrapping routine for our test without forsaking its non-parametric nature. In particular, Trokic (2014) showed how wavelet based bootstrapping, known as wavestrapping, can be used to effectively reduce size distortions in the case of unit root tests. Here we adopt a similar method in the multivariate setting. This method is conducted via two different wavelet filtering technique, namely Discrete Wavelet Transform (DWT) and Discrete Wavelet Packet Transform (DWPT). These techniques generate different size and power features. As DWPT provides excellent reduction in size distortion with low power, DWT has desirable size distortion adjustment and very good power. To the best of our knowledge, this is the first paper which implements wavestrapping in a multivariate cointegration setup.

The rest of the article is organized as follows: In section 2 we review Variance ratio cointegration rank test of Nielsen (2010). After briefly introducing discrete wavelet transform in section 3, we propose the new wavelet based variance ratio cointegration test in section 4. Section 5 will illustrate the small sample properties of the proposed test under different scenarios. Section 6 will be devoted to wavestrapping. The proof of the theorems can be found in appendix.

4.2 Variance Ratio Cointegration Test

To construct our cointegration test, we first define the concepts of fractional integration and cointegration. For this part, we mostly follow Nielsen (2010).

Definition 1. 1 The p vector time series Y_t is fractionally integrated of order d, denoted by $Y_t \sim I(d)$ if

$$Y_t = \triangle_+^{-d} v_t, \ t = 1, 2, 3, ... \tag{4.1}$$

where v_t has a continuous spectral density matrix that is bounded positive semi-definite, and bounded away from the zero matrix at all frequencies, and $\triangle_+^{-d}$ can be defined as:

$$\triangle_+^{-d} y_t = \sum_{j=0}^{t} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} y_{t-j} \tag{4.2}$$

such that only past values of y_t with a positive index enter the integration.

With this specification, the right hand side of (4.2) is always well defined for all values of d (Nielsen, 2010). Note that although the series in (4.1) are never stationary for any $d \neq 0$, they are asymptotically stationary for $d < 1/2$. As in Nielsen (2010), we also use the term stationary for this case. In addition, if $d = 1$, Y_t becomes a standard unit root process, but negative indexed history does not have any impact on the process.

Definition 2. 2 The p vector $Y_t \in I(d)$ is said to be cointegrated with rank r if there exist a full rank $p \times r$ matrix β such that $\beta' Y_t \in I(d-b)$ for $b > 0$. Further, we assume $0 < d - b < 1/2 < d$.

This definition generalizes the concept of cointegration to the fractional systems. Note that if we choose $d = b = 1$ this characterization collapses to the familiar I(1)-I(0) cointegration framework. This representation therefore gives us a chance to deal with many different types of integrated time series which we encounter in the economics and finance literature. Furthermore, remark that there is an extra assumption on the fractional order of the cointegrating residuals which is restricted to be less than 1/2. This implies that the cointegrating residuals are stationary for any value of d. This assumption is more technical and our proofs are based on it. Still, it is not too restrictive to impose such condition which says that equilibrium error is stationary.

To connect the above definition of fractional integration and cointegration with testing and asymptotic theory, Nielsen (2010) suggests the following assumption:

Assumption 6. (Assumption of Cointegration) Y_t is p vector of observable variables which are I(d). There exists a full rank orthonormal $p \times p$ matrix $R = [R_{p-r}, R_r]$, where R_{p-r} has $p-r$ columns and R_r has r columns with $0 \leq r \leq p-1$ such that

$$R'Y_t = \begin{bmatrix} \triangle_+^{-d} I_{p-r} & 0_{(p-r)\times r} \\ 0_{(p-r)\times(p-r)} & \triangle_+^{-(d-b)} I_r \end{bmatrix} u_t \tag{4.3}$$

where I_m is an m-dimensional identity matrix and u_t is generated by a linear stationary process $u_t = \Psi(L)\epsilon_t = \sum_{k=0}^{\infty} \Psi_k \epsilon_{t-k} \; \forall \; t = 1, 2, ...$ The $p \times p$ coefficient matrices Ψ_k satisfies $\sum_{k=0}^{\infty} k^{1/2} \|\Psi_k\| < \infty$, $rank(\Psi(1)) = n$ if $r = 0$, $rank(\Psi(1)) \geq p - r + 1$ and $rank(\Psi_{11}(1)) = p - r$ if $r \geq 1$, where $\Psi_{11}(1)$ is the upper left $(p - r) \times (p - r)$ block of $\Psi(1) = \sum_{k=0}^{\infty} \Psi_k$; the remaining blocks being $\Psi_{12}(1)$ and $\Psi_{21}(1)$, and $\Psi_{22}(1)$. Finally ϵ_t are i.i.d with $\mathbb{E}[\epsilon_t] = 0$, $\mathbb{E}[\epsilon_t \epsilon_t'] = I_p$, $\mathbb{E} \|\epsilon_t\|^q < \infty$ for some $q > max(4, 2/(2d - 1))$.

This is the driving assumption of Variance Ratio (VR) test of Nielsen (2010) and also of our test. Essentially, the assumption states that when $r \geq 1$ the system is cointegrated. On the other hand, when $r = 0$, all variables in Y_t follow distinct $I(d)$ processes which do not constitute any long run equilibrium. More importantly, equation (4.3) depicts a factorization for cointegrated system. The orthonormal matrix R factorizes the observed times series Y_t into two components which asymptotically behave very differently from each other. The first component can be interpreted as a $(p - r)$-dimensional common non-stationary factors. These factors are not cointegrated but they construct the base of the long run equilibrium among the variables in Y_t. As a result, these factors can be linked to common stochastic trend representations of cointegrated variables (see Stock and Watson (1988)). On the other hand, the second component of the factorization represents the stationary component of the equilibrium or the short run dynamics of the system. These dynamics are associated with the cointegrating residuals or the equilibrium error which is illustrated in definition 2. With these two separate set of processes, the cointegration system is characterized.

Under the assumption (6), Nielsen devises his non-parametric cointegration test as follows. First he defines the fractional integration of Y_t as $\tilde{Y}_t = \triangle_+^{-d_1} Y_t$ with the fractional integration parameter d_1, then he defines the moment matrices $A_T = \sum_{t=1}^{T} Y_t Y_t'$ and $B_T = \sum_{t=1}^{T} \tilde{Y}_t \tilde{Y}_t'$. The next step requires the computation and sorting the eigenvalues of $A_T B_T^{-1}$ which turns out to be the ratio of the variances of the observed series and its fractionally transformed counterpart. Let these eigenvalues be denoted as $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_p$. The test statistic for testing the null $H_0 : \; r = r_0$ is given by $\Lambda_{p,r_0}(d_1) = T^{2d_1} \sum_{i=1}^{p-r_0} \lambda_i$. In his paper, Nielsen (2010) explains that under the null of r_0 cointegrating relations, only the largest $p-r_0$ roots of the eigenvalue problem are non-zero. These roots are in the direction of the common factors or associated with the nonstationary part in assumption (6). The remaining r eigenvalues are in the direction of the cointegrating vector. More detailed discussion can be found in Nielsen (2010). The following Theorem illustrates the asymptotic distribution of the test statistic when the observed time series process does not contain any deterministic process:

Theorem 9. (Nielsen, 2010) Suppose that assumption (6) holds. For $d_1 > 0$ and $r = 0, 1, 2, ..., p-1$

$$\Lambda_{p,r}(d_1) \xrightarrow{D} tr \left\{ \int_0^1 B_d^{p-r}(s) B_d^{p-r}(s)' ds \left(\int_0^1 B_{d+d_1}^{p-r,d}(s) B_{d+d_1}^{p-r,d}(s)' ds \right)^{-1} \right\}$$

as $T \to \infty$ where $B_d^{p-r}(s)$ is $p-r$ dimensional vector of independent Fractional Brownian motions with fractional order d and $B_{d+d_1}^{p-r,d}(s)$ is $p-r$ vector of independent Fractional Brownian motion with fractional order $d + d_1$.

Observe here that the asymptotic distribution of the VR test statistic depends on d and d_1 where d_1 is chosen by the practitioner. Nielsen (2010) suggests a value close to 0.1. The fractional integration parameter can be estimated a priori to testing. If one has a consistent estimate for d, then the result in theorem (9) are valid. Additionally, the asymptotic distribution does not contain any information about the stationary component of the cointegrating system nor the short run dynamics of error structure. These components become asymptotically irrelevant. Accordingly, the test statistic does not depend on b and consequently, the limiting distribution of the VR test statistic is free from any nuisance parameters. As demonstrated in Nielsen (2010), this test has good size properties and considerable power under the alternative hypothesis and its is very easy to compute.

In this paper, we modify Nielsen's (2010) VR test by using wavelet theory. Next section briefly explains the basic properties of wavelets, which we utilize in our analysis. This section will mostly follow Gençay et al. (2001)and Trokic (2014).

4.3 Wavelet Transform

A wavelet $\psi(t)$ is a real valued function of time with two basic identities:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \quad and \quad \int_{-\infty}^{\infty} \psi(t)^2 dt = 1$$

These two properties implies, a wavelet function must take non-zero value through time, but all departures from zero should be cancelled out (Gencay, et al., 2000). By using a wavelet function, which satisfy above conditions, we can construct continuous time wavelet transform (CWT) of a time series x_t as follows:

$$W(u,s) = \int_{-\infty}^{\infty} x_t \psi_{u,s}(t) dt$$

where $\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right)$ is translated by u and dilated by s. Note that u is location and s is scale parameters. The parameter s allows wavelets to work under different scales. As it can be observed, this transformation is useful for the continuous time processes. Hence, CWT is not an appropriate tool when working with empirical data, which is rarely given in continuous time (Trokic, 2014). However, the fundamental properties of CWT filters have discrete time counterparts. Let $h = (h_0, h_1, ..., h_{L-1})$ denote a discrete wavelet (or high pass) filter with finite length L. The high pass filter h_l integrates to 0, $\sum_{l=0}^{L-1} h_l = 0$ and its variance is 1, $\sum_{l=0}^{L-1} h_l^2 = 1$ like CWT filters. Additionally this filter is orthogonal to its even shifts, that is:

$$\sum_{l=0}^{L-1} h_l h_{l+2n} = \sum_{l=-\infty}^{\infty} h_l h_{l+2n} = 0 \; for \; all \; n \in \mathbb{Z}^+$$

In DWT, we also have an additional filter g (low pass filter) which complements the high pass filter. This object can be obtained by quadrature mirror relationship:

$$g_l = (-1)^{l+1} h_{L-l-1} \; for \; all \; l = 0, 1, ..., L-1$$

Unlike the high pass filter, g integrates to $\sqrt{2}$, $\sum_{l=0}^{L-1} g_l = \sqrt{2}$, but it has unit variance $\sum_{l=0}^{L-1} g_l^2 = 1$. Moreover, low pass filter is orthogonal to its even shifts as well as the even shifts of high pass filter:

$$\sum_{l=0}^{L-1} g_l g_{l+2n} = \sum_{l=-\infty}^{\infty} g_l g_{l+2n} = 0$$
$$\sum_{l=0}^{L-1} g_l h_{l+2n} = \sum_{l=-\infty}^{\infty} g_l h_{l+2n} = 0$$

Applying these two filters to given time series we decompose the time series process in to its high frequency and low frequency components. Let $x = \{x_t\}_{t=1}^{T}$ be observed time series process with dyadic length $T = 2^J$ for some integer J. Then we define the matrix of DWT coefficients $\mathbf{w} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_J, \mathbf{v}_J]'$, where for $j = 1, 2, ..., J$ $\mathbf{w}_j$ is j-th level wavelet coefficients associated with the changes on scale of length $\lambda_j = 2^{j-1}$ and $\mathbf{v}_J$ is J-th level scaling (approximation) coefficients associated with the averages on scale $2\lambda_J = 2^J$. Moreover, note that $\mathbf{v}_J$ has length $T/2^J$ and $\mathbf{w}_j$ has length $T/2^{j-1}$ for each $j = 1, 2..., J$.

In practice, wavelet coefficients can be obtained by the pyramid algorithm first proposed by Mallat (1989). The algorithm starts with filtering the observed series by high pass and low pass filters to obtain first level wavelet and scaling coefficients:

$$v_{1,t} = \sum_{l=0}^{L-1} g_l x_{2t-l \mod T} \tag{4.4}$$

$$w_{1,t} = \sum_{l=0}^{L-1} h_l x_{2t-l \mod T} \tag{4.5}$$

where $t = 1, 2, ..., T/2$. Let $\mathbf{w}_1 = \{w_{1,t}\}_{t=1}^{T/2}$ and $\mathbf{v}_1 = \{v_{1,t}\}_{t=1}^{T/2}$ denote first level wavelet coefficients and scaling coefficients respectively. Then, the matrix $[\mathbf{w}_1, \mathbf{v}_1]'$ constitutes first level partial DWT of the time series x_t. For $j = 2, ..., J$

we can obtain j-th level coefficients by simply filtering the scaling coefficient of level $j-1$:

$$v_{j,t} = \sum_{l=0}^{L-1} g_l v_{j-1,2t-l \mod T}$$

$$w_{j,t} = \sum_{l=0}^{L-1} h_l v_{j-1,2t-l \mod T}$$

In this study, we only use first level DWT of observed time series processes. As indicated in Trokic (2014), the scale λ_j is inversely proportional to frequency. Consequently, for higher values of j λ_j corresponds to lower frequencies. This indicates that first level wavelet coefficients are associated with shortest run components. Then, since v_1 and w_1 are obtained by complement filters, we can conclude that v_1 corresponds to longest run patterns of the observed series. These long run components determine the characteristics of the cointegrating relations as in unit root case.

Next section will combine the VR testing with wavelet theory to build up a powerful nonparametric cointegration test.

4.4 Wavelet Variance Ratio Cointegration Test

Using the results of Nielsen (2010) and Trokic (2014), we construct a wavelet based cointegration test for fractionally integrated time series. In our analysis, we also consider type II fractionally integrated processes with deterministic components. These components are restricted to mean and time trend for simplicity. The following assumption depicts the characteristics of observed time series processes with these deterministic terms:

Assumption 7. The observed time series $\{W_t\}$ is generated by:

$$W_t = \alpha'\delta_t + Y_t$$

where for $j = 0$ $\delta_t = 0$, for $j = 1$ $\delta_t = 1$ and for $j = 2$ $\delta_t = [1, t]'$

In order to remove the deterministic components from the observed time series, we apply a OLS demeaning or detrending procedure. Hence, we use the residuals from the regression of δ_t on observed time series W_t instead of Y_t in above equations. The residuals are given by $\hat{Y}_t = Y_t - (\hat{\alpha} - \alpha)'\delta_t$. Now define the following demeaned ($j = 1$) and detrended ($j = 2$) versions of the Fractional Brownian motion:

$$\begin{aligned} B^{p-r}_{j,d}(s) \quad &= W^{p-r}_d(s) - \left(\int_0^1 W^{p-r}_d(s) D_j(s)' ds\right) \\ &\times \left(\int_0^1 D_j(s) D_j(s)' ds\right)^{-1} D_j(s) \qquad (4.6) \\ B^{p-r}_{j,d,d_1}(s) \quad &= \left[W^{p-r}_{d+d_1}(s) - \left(\int_0^1 W^{p-r}_d(s) D_j(s)' ds\right)\right. \\ &\left.\times \left(\int_0^1 D_j(s) D_j(s)' ds\right)^{-1} \int_0^s \frac{(s-r)^{d_1-1}}{\Gamma(d_1)} D_j(s) ds\right] \qquad (4.7) \end{aligned}$$

where $D_j(s) = 1$ if $j = 1$, and $D_j(s) = [1, s]'$ if $j = 2$ for $0 \leq s \leq 1$. These processes have same form as in Nielsen (2010).

Now applying the low pass filter to each element of the observable time series Y_t we obtain $V_{1,t}$. At the first stage we use Haar wavelet which can be defined as:

$$\hat{V}_{1,t} = \frac{1}{\sqrt{2}} \left(\hat{Y}_{2t} + \hat{Y}_{2t-1}\right) \; for \; t = 1, 2, .., T/2$$

We also apply fractional integration operator to Haar wavelet coefficients of the residual process, $\tilde{\hat{Y}}_t$ to obtain:

$$\tilde{\hat{V}}_{1,t} = \triangle^{-d_1} \hat{V}_{1,t} \; for \; t = 1, 2, .., T/2$$

We next redefine the objects $A_T^H = \sum_{t=1}^{T/2} \hat{V}_{1,t} \hat{V}_{1,t}'$ and $B_T^H = \sum_{t=1}^{T/2} \tilde{\hat{V}}_{1,t} \tilde{\hat{V}}_{1,t}'$ and let $T_1 = T/2$. Rest of the procedure is the same as Nielsen (2010). Let's define the test statistic as $\Lambda_{p,r_0}^H(d_1) = T_1^{2d_1} \sum_{i=1}^{p-r_0} \lambda_i^H$ where λ_i^Hs are the ordered eigen values of

$$A_T^H (B_T^H)^{-1} \tag{4.8}$$

The following theorem summarizes the main result of the paper.

Theorem 10. Suppose that the assumption 6 holds. For $d_1 > 0$ and $r = 0, 1, 2, ..., p-1$

$$\begin{aligned} &\Lambda_{p,r}^H(d_1) \xrightarrow{D} U_{p-r}(d, d_1) \\ &U_{p-r}(d, d_1) = tr \left\{ \int_0^1 B_{j,d}^{p-r}(s) B_{j,d}^{p-r}(s)' ds \left(\int_0^1 B_{j,d+d_1}^{p-r}(s) B_{j,d+d_1}^{p-r}(s)' ds \right)^{-1} \right\} \end{aligned}$$

as $T \to \infty$, where $B_{j,d}^{p-r}(s)$ is $p-r$ dimensional vector of independent standard ($j = 0$) or demeaned ($j = 1$) or detrended ($j = 2$) Fractional Brownian motions with fractional order d and $B_{j,d,d_1}^{p-r}(s)$ is $p-r$ vector of independent Fractional Brownian motion with fractional order $d + d_1$ for all j=0,1,2.

Theorem 10 indicates that the new wavelet based test statistic has very similar asymptotic properties as Nielsen's VR (2010) test. However, the detrending and demeaning part of this new test differs from Nielsen (2010) as mentioned above.

This new test also carries over the most important features of Nielsen's (2010) test such as being nuisance and tuning parameter free. Moreover, the test statistic only depends on d and d_1 and so there are no short run dynamics in the limiting distribution.

Up to now, the proposed wavelet based test statistic was obtained using the traditional Haar wavelets. The following theorem derives the asymptotics for the general compactly supported Daubechy class wavelets which were defined in 4.4. Here we first define $V_{1,t}^{L} = \sum\limits_{l=0}^{L-1} g_l y_{2t-l \mod T}$ which is the first level wavelet coefficient of the observed time series. Additionally, if we use the demeaned or the detrended series we can set $\hat{V}_{1,t}^{L} = \sum\limits_{l=0}^{L-1} g_l \hat{Y}_{2t-l \mod T}$. Accordingly, this object has the following form:

$$\hat{V}_{1,t}^{L} = \sum_{l=0}^{L-1} g_l Y_{2t-l \mod T} - (\hat{\alpha} - \alpha)' \sum_{l=0}^{L-1} g_l \delta_{2t-l \mod T}$$

Now we can redefine the objects $A_T^L = \sum\limits_{t=1}^{T/2} \hat{V}_{1,t}^{L} \hat{V}_{1,t}^{L'}$ and $B_T^L = \sum\limits_{t=1}^{T/2} \tilde{\hat{V}}_{1,t}^{L} \tilde{\hat{V}}_{1,t}^{L'}$. Let's define the test statistic as $\Lambda_{p,r_0}^{L}(d_1) = T_1^{2d_1} \sum_{i=1}^{p-r_0} \lambda_i^L$ where λ_i^Ls are ordered eigen values of

$$A_T^L (B_T^L)^{-1}$$

Theorem 11. Suppose that assumption 6 holds. For $d_1 > 0$ and $r = 0, 1, ..., p-1$

$$\Lambda_{p,r}^{L}(d_1) \xrightarrow{D} U_{p-r}^{L}(d, d_1)$$

$$U_{p-r}^{L}(d, d_1) = tr\left\{ \int_0^1 B_{j,d}^{p-r}(s) B_{j,d}^{p-r}(s)' ds \left(\int_0^1 B_{j,d+d_1}^{p-r}(s) B_{j,d+d_1}^{p-r}(s)' ds \right)^{-1} \right\}$$

as $T \to \infty$ where $B^{p-r}_{j,d}(s)$ is $p-r$ dimensional vector of standard $(j=0)$ or demeaned $(j=1)$ or detrended $(j=2)$ independent Fractional Brownian motions with fractional order d and $B^{p-r}_{j,d,d_1}(s)$ is $p-r$ vector of independent Fractional Brownian motion with fractional order $d+d_1$ for all j=0,1,2.

Theorem 10 and 11 clearly establish that using Haar or general Daubechy type wavelets give asymptotically equivalent results.

Finally we have additional theorem for asymptotic power properties of the proposed test.

Theorem 12. Under the assumptions of Theorem 10, the test that rejects the null $H_0:\ r=r_0$ when $\Lambda^H_{p,r}(d_1) > CV_{\xi,p-r_0}(d,d_1)$, where $CV_{\xi,p-r_0}(d,d_1)$ is found from

$$P\left(U_{p-r_0}(d,d_1) > CV_{\xi,p-r_0}(d,d_1)\right) = \xi$$

has assymptotic size ξ and is consistent against $H_1:\ r>r_0$

In theorem 6 $CV_{\xi,p-r_0}(d,d_1)$ is asymptotic critical value for the null hypothesis $H_0 : r = r_0$. This theorem is important for constructing a testing strategy to determine the cointegration rank as indicated in Nielsen (2010). Since the asymptotic power is against the alternative $H_1:\ r>r_0$, we adopt a recursive testing scheme. Starting from the null $H_0:\ r=0$, if H_0 is rejected, we continue with testing the null $H_0:\ r=1$. This continues until we cannot reject the new null hypothesis or $r=p$.

4.5 Small Sample Properties of the Wavelet VR Test

In this section we demonstrate size and power comparisons of the proposed test with Nielsen's (2010) test via simulations. Here we restrict ourselves to the use of Daubechy 2 wavelets[1] and test the null hypotheses $r = 0$, $r = 1$ and $r = 2$. We initially set $p - r = 2$ and generate two independent unit roots as common stochastic factors. Using these factors we obtain the observed time series. Detailed illustration of DGP's for each hypothesis will be given below. Further, we use the same critical values in all simulations. Critical values are also obtained by simulation[2]. After we find critical values we perform power and size exercise for cointegration rank $r = 0, 1, 2$.

Let ϵ_{jt} be independent standard normal draws for $j = 1, 2$. We obtain the common stochastic factors as $z_{j,t} = \sum_{t=1}^{T} \epsilon_{j,t}$. Simply applying the wavelet VR testing procedure defined in section 4, we can find the value of the test statistic in each simulation. Also note that we reject the null when the test statistic is larger than the appropriate critical value. Consider now the DGPs for each test of interest:

[1] We also implement our method with Haar Db4 Db8 Db16 wavelets, but by means of size and power, best results are obtained by using db2

[2] In all of the cases we use same common stochastic factors, thus we use these factors in critical value computations. We first generate independent standard normal distributions with unit variance, then from these normally distributed variables we generate unit roots. Using these observations we compute the WVR test statistic many times. Since all of the simulations are size corrected, we use the same sample size throughout the simulations. Number of iterations for the simulations are set to 10000.

(a) $H_0 : \; r = 0$

$$\begin{aligned} y_{1,t} &= z_{1,t} \\ y_{2,t} &= z_{1,t} + z_{2,t} \\ z_{2,t} &= \rho_c z_{2,t-1} + \epsilon_{2,t} \end{aligned}$$

where $\rho_c = (1 - c/T)$. Note that if $c = 0$ then $y_{1,t}$ and $y_{2,t}$ are not cointegrated and so $r = 0$ holds with common stochastic trends $z_{1,t}$ and $z_{2,t}$. However, if $c > 0$ then $u_{2,t}$ is a stationary process and these two variables are cointegrated thereby satisfying the alternative hypothesis.

(b) $H_0 : \; r = 1$

$$\begin{aligned} y_{1,t} &= z_{1,t} \\ y_{2,t} &= z_{1,t} + u_{2,t} \\ y_{3,t} &= z_{1,t} + u_{2,t} + z_{2,t} \\ z_{2,t} &= \rho_c z_{2,t-1} + \epsilon_{2,t} \end{aligned}$$

where $u_{2,t}$ is standard normal draw which is independent of $\epsilon_{j,t}$ for $j = 1, 2$. Similar to case above, if $c = 0$. $z_{2,t}$ will be nonstationary and we have only one cointegrating relation which is between $y_{1,t}$ and $y_{2,t}$. However, if $c > 0$ then $z_{2,t}$ will be stationary and we have cointegration between $y_{1,t}$ and $y_{3,t}$. Consequently, $r = 2$.

(c) $H_0: \ r = 2$

$$\begin{aligned} y_{1,t} &= z_{1,t} \\ y_{2,t} &= z_{1,t} + u_{2,t} \\ y_{3,t} &= z_{1,t} + u_{2,t} + u_{3,t} \\ y_{4,t} &= z_{1,t} + u_{2,t} + u_{3,t} + z_{2,t} \\ z_{2,t} &= \rho_c z_{2,t-1} + \epsilon_{2,t} \end{aligned}$$

where $u_{2,t}$ and $u_{3,t}$ are standard normal draws which are independent of $\epsilon_{j,t}$ for $j = 1, 2$. As before, if $c = 0$. $z_{2,t}$ will be nonstationary and we have only two cointegrating relations which are between $y_{1,t}$ and $y_{2,t}$, and $y_{1,t}$ and $y_{3,t}$. However, if $c > 0$, then $z_{2,t}$ will be stationary and we have cointegration between $y_{1,t}$ and $y_{4,t}$. As a result, $r = 3$. Table 1 demonstrates the power and size performance of WVR test and VR test.

From table 17 we can observe that Nielsen's (2010) VR and the proposed wavelet VR tests are exhibiting similar performances by means of both size and power. For both tests, size is close to 0.05 and as c increases the power approaches to 1. Consequently, the newly proposed wavelet variance ratio test enjoys similar small sample properties as Nielsen's Test (2010). One critical note about the small sample behavior of these tests is that Nielsen's (2010) test almost always exhibits slightly better power. This appears mostly in form of slow convergence of wavelet based tests. As it can be seen in technical proofs, wavelet test have some extra components which do not appear in Nielsen's test. Although these components are asymptotically not important, they may have some impact on the test statistic for small samples.

Table 17: **Size and Power for Variance Ratio and Wavelet Variance Ratio Tests for Standard Cointegration**

	r=0		r=1		r=2	
ρ_c	WVR	VR	WVR	VR	WVR	VR
1	0.0571	0.0562	0.0514	0.0536	0.0469	0.0489
0.95	0.1706	0.1755	0.1745	0.1808	0.168	0.1802
0.9	0.3416	0.3639	0.3284	0.3566	0.3201	0.3514
0.85	0.5164	0.5631	0.4985	0.553	0.4809	0.547
0.8	0.6674	0.7362	0.6491	0.7274	0.6294	0.7163
0.75	0.7932	0.8622	0.7841	0.8636	0.7598	0.8558
0.7	0.885	0.9433	0.8665	0.9372	0.859	0.9396
0.65	0.9378	0.9801	0.9266	0.9764	0.9104	0.9714
0.6	0.9696	0.9932	0.9579	0.9923	0.9491	0.9885
0.55	0.9846	0.9987	0.9765	0.9979	0.9725	0.9982
0.5	0.9925	0.9995	0.9877	0.9995	0.9839	0.9994

Note: VR is Nielsen's Variance ratio (2010) and WVR is wavelet Variance ratio test. Sample size is 128 for each scenario. We repeat each test for 100000 times. We use size corrected critical values for both Variance ratio and wavelet variance ratio tests.

The next table will demonstrate power curves for Nielsen and Wavelet Variance Ratio tests for fractional cointegration. In this case, instead of using and AR(1) near integrated process for $z_{2,t}$, we generate it as fractionally integrated process i.e $z_{2,t} = \Delta_+^{-d}\epsilon_{2,t}$, where $\epsilon_{2,t}$ is defined as before. Results are similar to standard cointegration case, but with improved power. This result is expected since fractional noises establish stronger persistence than AR(1) processes.

Table 18: **Size and Power for Variance Ratio and Wavelet Variance Ratio Tests for Fractional Cointegration**

	r=0		r=1		r=2	
d	WVR	VR	WVR	VR	WVR	VR
0	0.0567	0.0562	0.0471	0.0482	0.0485	0.052
0.1	0.00.0738	0.0807	0.0725	0.0797	0.0655	0.0762
0.2	0.1164	0.1396	0.1123	0.1403	0.1094	0.1372
0.3	0.2037	0.261	0.1812	0.2398	0.1805	0.2476
0.4	0.3165	0.4283	0.301	0.4239	0.2955	0.4201
0.5	0.4984	0.6715	0.4839	0.6668	0.4597	0.654
0.6	0.7124	0.8847	0.6959	0.8807	0.6751	0.8736
0.7	0.895	0.9815	0.8871	0.9816	0.8597	0.9778
0.8	0.9833	0.9998	0.9749	0.9995	0.9691	0.9997
0.9	0.9989	1.000	0.9984	1.000	0.9976	1.000
1	1.000	1.000	1.000	1.000	0.9997	1.000

Note: VR is Nielsen's Variance ratio (2010) and WVR is wavelet Variance ratio test. Sample size is 128 for each scenario. We repeat each test for 100000 times. We use size corrected critical values for both Variance ratio and wavelet variance ratio tests.

In order to analyse size properties of these two tests in detail, we run another series of simulations in which we include scenarios where cointegration tests may suffer more severe size distortions. In this regard, one popular exercise is to check for the size properties when the innovations of the cointegrating system follow a moving average process with a highly negative MA parameter. For instance, in the simple case of unit roots, if x_t is a unit root with $x_t = x_{t-1} + \epsilon_t$ and $\epsilon_t = \eta_t + \theta\eta_{t-1}$ with a highly negative θ, x_t will look like a stationary process since ϵ_t acts like an over-differenced series. As a result. the unit root test will overreject the null of unit root. Mallor and Lence (2012) investigate this issue in a cointegration framework by allowing a negative MA root in the cointegration system. We will use a similar but simple procedure by allowing MA root in error terms of common stochastic trends. Due to space limitations, we will only consider one common stochastic trend that includes MA innovations. Adopting the same setup as in the previous exercises, we set $c = 0$ and $\epsilon_{2,t} = \eta_t + \theta\eta_{t-1}$ where η_t is drawn from standard normal distribution. Hence, with this DGP we operate under the null hypothesis of $r = r_0$ where $r_0 = 0, 1, 2$. Results are given in table 2.

Table 2 indicates that both tests suffer from size distortions arising from the presence of highly negative MA roots in the common stochastic trends. Note however that the Wavelet based test seems to be affected less than Nielsen's test (2010). This is a natural consequence of the wavelet transform. The wavelet transform used in this paper is constructed to filter out high frequency components which contains short run problematic dynamics such as MA process. Additionally, the table 2 clearly shows that the wavelet based test has better size properties even in the presence of positive MA roots.

Table 19: **Size distortion of Variance Ratio and Wavelet Variance Ratio Tests in Presence of MA roots**

	r=0		r=1		r=2	
θ	WVR	VR	WVR	VR	WVR	VR
-0.9	0.7274	0.9042	0.6989	0.9018	0.669	0.8887
-0.7	0.1843	0.3331	0.1716	0.3256	0.1677	0.3268
-0.5	0.0924	0.1385	0.0785	0.1299	0.0759	0.1274
-0.3	0.0624	0.0753	0.0613	0.0733	0.0589	0.0726
-0.1	0.0527	0.0546	0.0526	0.0559	0.0511	0.0563
0	0.0537	0.0532	0.0478	0.0483	0.0505	0.0535
0.1	0.0497	0.0481	0.0507	0.0498	0.0468	0.048
0.3	0.0492	0.0469	0.0476	0.0459	0.049	0.0476
0.5	0.0505	0.0456	0.0441	0.0415	0.0473	0.0456
0.7	0.0492	0.0447	0.0485	0.0448	0.0473	0.0464
0.9	0.0529	0.0464	0.0476	0.0433	0.0469	0.0456

Note: VR is Nielsen's Variance ratio (2010) and WVR is wavelet Variance ratio test. Sample size is 128 for each scenario. We repeat each test for 100000 times. We use size corrected critical values for both Variance ratio and wavelet variance ratio tests without MA root adjustment.

4.6 Wavestrapping for Cointegration Test

Although the wavelet test still performs better in the presence of MA innovations, size distortions are still problematic, particularly for highly negative MA roots.

In order to address this issue, we exploit the wavelet nature of the proposed test to further eliminate any remaining size distortions. A common practice in the literature of unit roots is to apply a sieve parametric bootstrap method, such as in Nielsen (2009). Nevertheless, such bootstrap methods rely on parametric techniques such as specifying the parameter governing the sieve length. Even though Nielsen (2009) proposes a non-parametric unit root test, he relies on parametric bootstrap for fixing size distortion issues. This of course implies that the bootstrapped test no longer enjoys the advantages of being fully nonparametric, which of course could be somewhat of a disadvantage. Our purpose in this paper is to construct a wavelet based non-parametric bootstrap method which does not rely on any parametric estimation of the cointegration system. Further, apart from reducing size distortions, this nonparametric wavestrappping method will be an alternative to parametric sieve bootstrap for cointegration rank determination (see Cavaliere et al. (2010)). The method proposed here will closely follow Trokic (2014) and Percival and Walden (2006)).

As its name indicates, wavestrapping is a bootstrap like procedure in which resampling is done via the wavelet transform of the observed series instead. The main idea behind this procedure stems from the fact that discrete wavelet transform approximately decorrelates long memory time series (Percival and Walden, 2006). After obtaining and resampling decorrelated coefficients, one can obtain the resampled process via reconstruction filter. We summarize the DWT based wavestrapping below:

1. Fix the Monte Carlo replication as MC and nominal size as α. Given the

observed time series $Y_{p\times T}$ with length $T = 2^M$, set the level of DWPT as $J_0 = M - 2$. Also compute the VR test statistic (either VR or WVR) $\Lambda_{p,r_0}(d_1)$.

2. First of all, we will apply the procedure to the common stochastic factors, which can be derived from the eigen value problem defined in section 2. According to Nielsen (2010), if we define $\eta(p - r_0)$ as the matrix in which each column is the eigen vector corresponding to the $p - r_0$ smallest eigen values, these eigen vectors are in direction of the non-cointegrated component (Nielsen, 2010). Then, if we multiply $\eta(p - r_0)$ with Y we obtained the common stochastic factors upto a linear transformation, say Z_t. However, to implement wavestrapping properly we need to have approximately independent common factors. Then, we apply Cholesky decomposition to Z_t. Let $\Omega = chol(var(Z_t))$, we can obtain the independent common factors as $\hat{Z} = \eta(p - r)'Y * \Omega^{-1}$.

3. Similar to bootstrapping, we need to resample the time series under the null hypothesis in which we have $p - r_0$ independent unit root processes[3]. First, define $\hat{Z}_j$ as j-th column of $\hat{Z}$ for $j = 1, ..., p - r_0$. Then, under the null $\hat{Z}_{j,t} = \hat{Z}_{j,t-1} + \eta_{j,t}$ for all $j = 1, ..., p - r_0$. We will wavestrap each η_j individually by constructing a resampling scheme only for high pass filter part of the decomposition as stated in Tang et al. (2008). Obtain the wavelet decomposition $\{V_{J_0,j}, W_{1,j}, W_{2,j}, ..., W_{J_0,j}\}$, then resample each $W_{i,j}$ with replacement for $i = 1, 2, ..., J_0$ to have the resampled decomposition

[3] In case of Fractional Cointegration, our method also works but it is slightly different. Instead of assuming we have $p - r$ unit roots, we assume we have $p - r$ I(d) series under the null. Consequently instead of wavestrapping first differenced series, we need to fractionally difference the observe series and apply wavestrapping routine described.

coefficients $\left\{V_{J_0,j}, W^*_{1,j}, W^*_{2,j}, ..., W^*_{J_0,j}\right\}$. Finally applying reconstruction filter we obtain $\left\{e^*_{j,t}\right\}$ and $\hat{Z}^*_t = \sum_{i=1}^{t} e^*_{j,t}$. After integrating and stacking these variables in $\hat{Z}^*$, we compute the test statistic for new wavestrapped sample $\hat{Z}^*$, say $\Lambda^*_{p,r_0}(d_1)$.

4. Pick a large bootstrap number B and repeat the steps 3-7 B times.

5. Calculate the wavestrap p value with:

$$p^*_m = \frac{1}{B}\sum_{i=1}^{B} \mathbb{1}_{\left\{\Lambda_{p,r_0}(d_1) > \Lambda^*_{p,r_0}(d_1)\right\}}$$

   for $m = 1, ..., MC$

6. The wavestrap rejection probability can be computed as :

$$RP^* = \frac{1}{MC}\sum_{m=1}^{MC} \mathbb{1}_{\{p^*_m < \alpha\}}$$

This approximate decorrelation feature allows one to apply bootstrap procedure in which the independent replicates of the wavelet coefficients are generated by resampling. Applying inverse transform to these resampled wavelet coefficients, an independent replicate of the observed time series can be obtained (Trokic, 2014). Nonetheless, Percival et al (2000) showed that this procedure does not work well for the short memory series, such as MA process since DWT of such processes may not produce adequate decorrelation for wavelet coefficients (Percival, et al., 2000). To handle this issue properly, Percival et al. (2000) proposed another wavelet based method for obtaining independent replicates of time series with arbitrary short run dynamics. This method, called wavestrapping, is based on Discrete wavelet packet transform (DWPT) which offers a way to search for the least correlated wavelet coefficients.

DWPT is a similar procedure as DWT's pyramid algorithm. The main difference is that in DWPT we apply high pass and low pass filters to each wavelet coefficient and detail level. Therefore, we have following tree structure for 3 level DWPT: Each node in this transform corresponds to the wavelet coefficients ob-

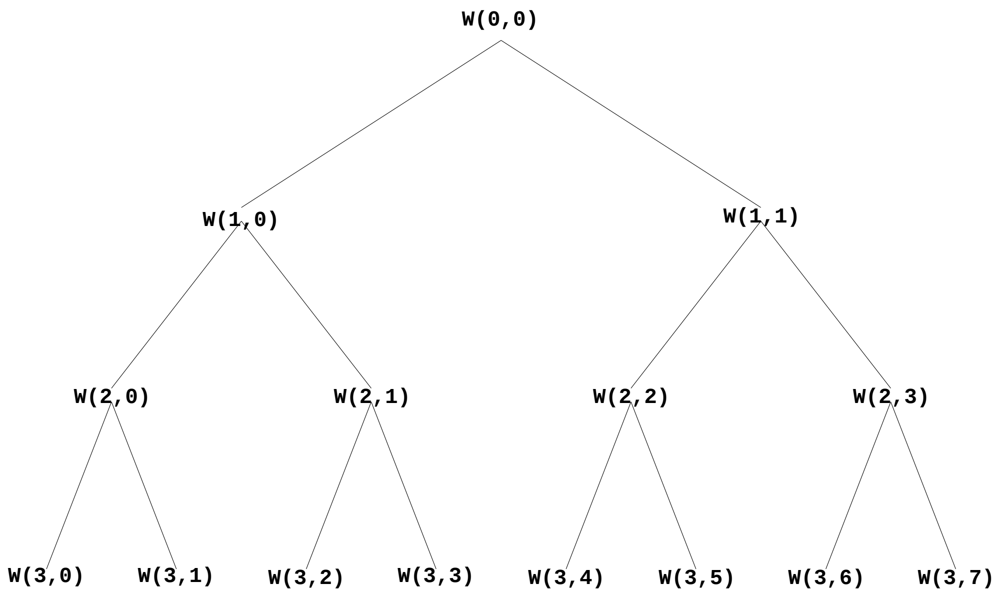

Figure 1: **DWPT Tree Representation**

tained in each step. Starting from the node (0,0) (this is observed time series) we apply the low pass filter to obtain $W(1,0)$ and high pass filter to obtain $W(1,1)$. Repeating this procedure for each node we obtain the wavelet packet tree with $\sum_{j=1}^{J} 2^{j-1}$ wavelet coefficients. Note that the left hand side is coefficients associated with the low pass filter, right hand side is the wavelet coefficients associated with the high pass filter.

After obtaining DWPT, we can proceed the wavestrapping by first searching for which coefficients we need to re-sample. This search is done by a white noise test applied to wavelet coefficients. Any white noise test can be applicable. In this study we apply Ljung-Box-Q test. Following procedure will clarify the idea behind the wavestrapping cointegration test with DWPT Under the null H_0 $r = r_0$:

1. We follow step 1-2 of DWT based wavestraping
2. We will wavestrap each e_j (defined in step 3 of DWT wavestrapping) individually by constructing a resampling scheme only for high pass filter part of the DWPT transform. This will include the elements only on right hand side of the DWPT tree. This is required since DWT offers decorrelation of the detail levels (high pass coefficients). As a result, we keep $W(1, 0)$ and do not go below in that branch.
3. Starting from $W(1, 1)$, if $j = J_0$, then keep $W(j, k)$ for all $k = 2^{j-1}, .., 2^j - 1$. If $j < M_0$ level run a white noise test for $W(j, k)$ for $k = 2^{j-1}, .., 2^j - 1$, if white noise hypothesis is rejected than decompose the series into $W(j + 1, 2k)$ with low pass filter and $W(j + 1, 2k + 1)$ with high pass filter.[4]
4. If the white noise test is not rejected than keep that coefficient and do not go below in that branch. We will call these nodes as Terminal node.
5. After obtaining the terminal nodes as in step 4 resample coefficients in each terminal node without replacement to obtain a bootstrap sample for these coefficients.

[4] One can prefer to apply white noise test only to high pass filtered coefficients, However, current routine is also doing a good job by means of removing size distortions

6. Reconstruct the DWPT with these terminal nodes in right hand side and single node in left hand side to obtain wavestrapped time series η_j^* for $j = 1, ..., p - r_0$. After integrating and stacking these variables in $\hat{Z}^*$, we compute the test statistic for new wavestrapped sample $\hat{Z}_b^*$, say $\Lambda_{p,r_0}^*(d_1)$.

7. We follow step 4-6 of DWT based wavestrapping.

The steps 1-6 describes the wavestrapping procedure for a particular sample. Step 7 is for size distortion and power evaluation exercise and it requires to compute $MC(B+1)$ test statistic (Trokic, 2014). Trokic (2014) state that steps 1-6 of DWPT and DWT resampling is a double bootstrap exercise which is relatively expansive to compute. However, Trokic (2014) indicates that one can use Fast double bootstrap (wavestrap) procedure in which we only need to calculate $2MC$ number of test statistic to inexpensively compute the rejection probability in step 7. in order to achieve this, first set $B = 1$ and estimate the RP^* as:

$$RP^* \backsimeq RP^*_{FDW} = \frac{1}{MC}\sum_{m=1}^{MC} \mathbb{1}_{\left\{\Lambda_{p,r_0}(d_1) > Q^*(1-\alpha)\right\}}$$

where $Q^*(1-\alpha)$ is the $(1-\alpha)$-th quantile of the wavestrap test statistic $\Lambda_{p,r_0}^*(d_1)$.

Next tables illustrates the implementation of wavestrapping to both VR and WVR test for size and power evaluation. We adopt same DGP defined in Chapter 5. We only report the results for $r = 0$. However, the results for $r = 1$ are also similar[5]. As in chapter 5, we only consider Daubechy 2 wavelets in our exercises.

[5]Results for $r = 1$ are available upon request from author

Table 20: **Size Distortion for Wavelet Variance Ratio and Wavestrapped Tests Under the Null $r = 0$**

	T=128			T=1024		
θ	WVR	WVR^{ws1}	WVR^{ws2}	WVR	WVR^{ws1}	WVR^{ws2}
-0.9	0.7142	0.0681	0.2027	0.412	0.038	0.1615
-0.7	0.1778	0.0434	0.0542	0.0947	0.0429	0.0471
-0.5	0.0811	0.0471	0.0483	0.0728	0.048	0.0503
-0.3	0.0586	0.0494	0.0544	0.0638	0.0485	0.0561
-0.1	0.0527	0.0496	0.0553	0.0669	0.0502	0.0596
0.1	0.0509	0.0487	0.0613	0.0596	0.0496	0.0545
0.3	0.0496	0.0484	0.0569	0.0651	0.0509	0.0612
0.5	0.0512	0.0516	0.0616	0.0625	0.0507	0.0596
0.7	0.0526	0.0524	0.0602	0.0604	0.0511	0.0536
0.9	0.0493	0.0502	0.0618	0.0615	0.0486	0.055

Note: WVR is wavelet Variance ratio test. We also have WVR^{ws1} as DWPT based wavestrapped version of WVR test and WVR^{ws2} as DWT based wavestrapped version of WVR tests. Sample size is 128 for each scenario. 4 and 7 level DWPT and DWT are utilized for $T = 128$ and $T = 1024$ respectively. We repeat each test for 10000 times. We use size corrected critical values for wavelet variance ratio tests without MA root adjustment.

Table 21: **Power of Wavelet Variance Ratio and Wavestrapped Tests for Diffirent Values of θ Under the Null $r = 0$ I**

$\theta = -0.9$				
	T=128		T=1024	
ρ_c	WVR^{ws1}	WVR^{ws2}	WVR^{ws1}	WVR^{ws2}
1	0.0653	0.1783	0.037	0.1663
0.85	0.2461	0.7844	0.4134	1.000
0.7	0.3409	0.9565	0.5786	1.000
0.55	0.4125	0.9921	0.8077	1.000
0.4	0.4772	0.9971	0.97	1.000
0.25	0.5537	0.999	0.9997	1.000
0.1	0.6903	0.9996	1.000	1.000
0	0.7645	0.9993	1.000	1.000
$\theta = -0.5$				
	T=128		T=1024	
ρ_c	WVR^{ws1}	WVR^{ws2}	WVR^{ws1}	WVR^{ws2}
1	0.0485	0.0445	0.0474	0.0536
0.85	0.079	0.1955	0.1425	0.9999
0.7	0.131	0.3989	0.2322	1.000
0.55	0.1893	0.5895	0.316	1.000
0.4	0.2519	0.7797	0.4055	1.000
0.25	0.3187	0.9095	0.5534	1.000
0.1	0.4008	0.9542	0.841	1.000
0	0.5038	0.9751	0.9464	1.000

Table 22: **Power of Wavelet Variance Ratio and Wavestrapped Tests for Diffirent Values of θ Under the Null $r = 0$ II**

$\theta = 0$				
	T=128		*T=1024*	
ρ_c	WVR^{ws2}	WVR^{ws1}	WVR^{ws2}	
1	0.0518	0.0564	0.0505	0.0555
0.85	0.0593	0.1434	0.0882	0.9957
0.7	0.0836	0.2525	0.1451	1.000
0.55	0.1123	0.3645	0.1936	1.000
0.4	0.1374	0.4435	0.2457	1.000
0.25	0.1625	0.5611	0.2599	1.000
0.1	0.199	0.6257	0.3311	1.000
0	0.2112	0.6814	0.3685	1.000
$\theta = 0.5$				
	T=128		*T=1024*	
ρ_c	WVR^{ws1}	WVR^{ws2}	WVR^{ws1}	WVR^{ws2}
1	0.0522	0.0584	0.0518	0.0571
0.85	0.0566	0.14	0.0695	0.9924
0.7	0.0639	0.2529	0.1011	1.000
0.55	0.0779	0.3422	0.1493	1.000
0.4	0.0918	0.4244	0.1873	1.000
0.25	0.1174	0.4939	0.2096	1.000
0.1	0.1266	0.5368	0.2408	1.000
0	0.1395	0.5839	0.2572	1.000

The results of size distortion exercise indicate that the new wavestrapped cointegration tests are doing good job in reducing size distortion for even in extreme cases. DWPT version appears to have better size distortion reduction than DWT based wavestrapping. Moreover, the simulations demonstrate that Wavestrapped WVR test is more robust to negative MA root case than standard version of VR and WVR tests. Large sample properties of wavestrapped tests are also desirable. As a result, this procedure can be a good alternative to existing methods used for determining cointegration rank. By using steps described above, one can compute bootstrap distribution of the test statistic. This distribution does not include any nuisance or tuning parameter and it is easy to compute. Unlike other bootstrap methods in the cointegration literature, this method is nonparametric by means of test statistic and bootstrap routine.

Our proposed wavetsrap procedure also exhibits power against local alternatives. Table 22 depicts the power curves for different values of MA coefficient of the innovations. These simulations are performed in order to show our wavestrapping test maintain Power properties of the original test. In all cases, DWT based wavestrapping has better power than DWPT based wavestrapping. In spite of the fact that for negative MA coefficient DWPT becomes a good alternative by means of power and size , DWT is more stable in sense that it is always more powerful and has sufficient size adjustment. For the cases $\theta = 0$ and $\theta = 0.5$ DWPT produces at most %30 power even under pure stationary case and large sample size. Large sample power properties of DWT is more desirable. It seems we lose a little power when $\theta = 0$ and $\theta = 0.5$, but this kind of power loss is also observed in sieve bootstrapped version of Nielsen's (2009) unit root test.

4.7 Conclusion

In this paper we have proposed a non-parametric wavelet based cointegration testing method for fractionally integrated time series processes as well as the usual unit root variables. This method can be regarded as an alternative version of Nielsen's (2010) variance ratio test. In fact, these two tests share many common features. First, both tests are fully nonparametric and do not require the estimation of a regression model to remove short run dynamics. They also possess nuisance and tuning parameter free limiting distributions if one has a consistent estimate of the fractional integration parameter of the observed variables. Furthermore, the asymptotic distributions of these tests are same except for the detrended version of the test statistic.

Although these two tests enjoy similar properties asymptotically, they may exhibit slightly different small sample characteristics in particular scenarios. Monte Carlo experiments indicate that when we compare power performances, the standard VR test yields better results. On the other hand, if we study size distortions, the wavelet based test has superiority over the standard VR test under a wide range of specifications. Even so, size distortions in extreme cases, such as in the presence of negative MA roots in common stochastic factors cannot be eliminated effectively by either of the tests. To handle such scenarios we propose a multivariate version of the wavestrapping procedure of Trokic (2014). Instead of wavestrapping the observed time series vector however, we wavestrap the common stochastic factors which in fact carries all statistical features of the cointegration test in the Nielsen (2010) setup. Simulations demonstrate tremendous improvement in size distortion even in extreme negative MA root cases. Further, DWT based procedure exhibits local power under different innovation structures.

Removing size distortions in particular cases is not the only benefit of the wavestrapping procedure. It also provides an alternative bootstrapping method for the determination of cointegration rank. To our knowledge this is this first attempt to implement a wavelet based bootstrapping algorithm in multivariate time series, particularly in a cointegration framework. Furthermore, unlike the sieve bootstrapping methods which require the sieve length tuning parameter and estimation of VECM model(Cavaliere et al. (2012) and SSwensen (2006)), this new method remains fully nonparametric and tuning parameter free. In order to apply wavestrapping in cointegration framework, we only need to estimate common stochastic trends. These trends can be nonparametrically estimated from the eigen value problem of Nielsen's (2010) VR test explained in section 2.

Finally, this paper has expanded the wavelet framework for unit root testing to multivariate models by providing derivations for not only multiple time series processes but also for fractionally integrated systems. We believe that these findings may shed light on the properties of the wavelets transform of fractionally integrated variables in a multivariate setup.

BIBLIOGRAPHY

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73(1):5–59.

Billingsley, P. (1968). *Convergence of Probability Measures.* New York: Wiley.

Box, G. E. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332):1509–1526.

Breitung, J. (2002). Nonparametric tests for unit roots and cointegration. *Journal of Econometrics*, 108:342–363.

Busetti, F. and Taylor, A. R. (2003). Testing against stochastic trend and seasonality in the presence of unattended breaks and unit roots. *Journal of Econometrics*, 117(1):21–53.

Cavaliere, G. (2005a). Limited time series with a unit root. *Econometric Theory*, 21(05):907–945.

Cavaliere, G. (2005b). Unit root tests under time-varying variances. *Econometric Reviews*, 23(3):259–292.

Cavaliere, G., Phillips, P. C. B., Smeekes, S., and Taylor, A. M. R. (2015). Lag length selection for unit root tests in the presence of nonstationary volatility. *Econometric Reviews*, 34(4):512–536.

Cavaliere, G., Rahbek, A., and Taylor, A. (2012). Bootstrap determination of the co-integration rank in vector autoregressive models. *Econometrica*, 80(4):1721–1740.

Cavaliere, G., Rahbek, A., and Taylor, A. R. (2010). Testing for co-integration in vector autoregressions with non-stationary volatility. *Journal of Econometrics*, 158(1):7–24.

Cavaliere, G. and Taylor, A. R. (2007). Testing for unit roots in time series models with non-stationary volatility. *Journal of Econometrics*, 140(2):919–947.

Cavaliere, G. and Xu, F. (2014). Testing for unit roots in bounded time series. *Journal of Econometrics*, 178:259–272.

Chambers, M. J., Ercolani, J. S., and Taylor., A. M. (2014). Testing for seasonal unit roots by frequency domain regression. *Journal of Econometrics*, (178):243–258.

Choi, I. and Phillips, P. (1993). Testing for a unit root by frequency domain regression. *Journal of Econometrics*, 59():263286.

Davidson, R. and MacKinnon, J. G. (2007). Improving the reliability of bootstrap tests with the fast double bootstrap. *Computational Statistics & Data Analysis*, 51(7):3259–3281.

Davydov, Y. A. (1970). The invariance principle for stationary processes. *Theory of Probability & Its Applications*, 15(3):487–498.

Demetrescu, M. and Sibbertsen, P. (2014). Inference on the long-memory properties of time series with non-stationary volatility. Technical report, Discussion Paper, Wirtschaftswissenschaftliche Fakultät, Leibniz University of Hannover.

Eroğlu, B. A. and Yiğit, T. (2016). A nonparametric unit root test under non-stationary volatility. *Economics Letters*, 140:6–10.

Fan, Y. and Gencay, R. (2010). Unit root tests with wavelets. *Econometric Theory*, 05(26):1305–1331.

Fernandez-Macho, J. et al. (2013). A wavelet approach to multiple cointegration testing. Technical report.

Gençay, R., Selçuk, F., and Whitcher, B. J. (2001). *An introduction to wavelets and other filtering methods in finance and economics*. San Diego: Academic press.

Granger, C. (1966). The typical spectral shape of an economic variable. *Econometrica: Journal of the Econometric Society*, pages 150–161.

Granger, C. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, (16):121*130.

Granger, C. (2010). Some thoughts on the development of cointegration. *Journal of Econometrics*, 158(1):3–6.

Haldrup, N. and Hylleberg, S. (1995). A note on the distribution of the least squares estimator of a random walk with drift: Some analytical evidence. *Economics Letters*, 48(3-4):221–228.

Harrison, J. (1985). *Brownian motion and stochastic flow systems*. New York: Wiley.

Ikeda, N. and Watanabe, S. (2014). *Stochastic differential equations and diffusion processes*, volume 24. Elsevier.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Contro*, 1(12):231–254.

Johansen, S. and Nielsen, M. Ø. (2012). A necessary moment condition for the fractional functional central limit theorem. *Econometric Theory*, 28(03):671–679.

Kim, T.-H., Leybourne, S., and Newbold, P. (2002). Unit root tests with a break in innovation variance. *Journal of Econometrics*, 109(2):365–387.

Levy, D. (2002). Cointegration in frequency domain. *Journal of Time Series Analysis*, 23(3):333–339.

Mallat, S. (1989). A theory for multiresolution signal decomposition: The wavelet representation. pattern analysis and machine intelligence. *IEEE Transactions*, 11(7):674–693.

Mallor, M. and Lence, S. H. (2012). Testing for cointegration in the presence of moving average errors. *Journal of Time Series Econometrics*, 4(2):.

Marinucci, D. and Robinson, P. M. (2000). Weak convergence of multivariate fractional processes. *Stochastic Processes and their applications*, 86(1):103–120.

McConnell, M. M. and Perez-Quiros, G. (1998). Output fluctuations in the united states: what has changed since the early 1980s? *FRB of New York Staff Report*, 41.

Morana, C. (2004). Frequency domain principal components estimation of fractionally cointegrated processes. *Applied economics letters*, 11(13):837–842.

Müller, U. (2008). The impossibility of consistent discrimination between i (0) and i (1) processes. *Econometric Theory*, 24(03):616–630.

Nelson, C. R. and Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. *Journal of monetary economics*, 10(2):139–162.

Ng, S. and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, pages 1519–1554.

Nielsen, M. . (2004). Spectral analysis of fractionally cointegrated systems. *Economics Letters*, 83(2):225–231.

Nielsen, M. (2009a). A powerful test of the autoregressive unit root hypothesis based on a tuning parameter free statistic. *Econometric Theory*, 25(06):1515–1544.

Nielsen, M. (2010). Nonparametric cointegration analysis of fractional systems with unknown integration orders. *Journal of Econometrics*, 155(2):170–187.

Nielsen, M. Ø. (2009b). A powerful test of the autoregressive unit root hypothesis based on a tuning parameter free statistic. *Econometric Theory*, 25(06):1515–1544.

Percival, D. B. and Walden, A. T. (2006). *Wavelet methods for time series analysis*, volume 4. New York: Cambridge university press.

Phillips, P. and Ouliaris, S. (1988). Testing for cointegration using principal components methods. *Journal of Economic Dynamics and Control*, (12):205–230.

Phillips, P. C. (1987a). Towards a unified asymptotic theory for autoregression. *Biometrika*, 74(3):535–547.

Phillips, P. C. (1987b). Towards a unified asymptotic theory for autoregression. *Biometrika*, 74(3):535–547.

Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335346.

Reed, J., Ward, A., and Zhan, D. (2012). The generalized drift skorokhod problem in one dimension: Its solution and application to the gi/gi/1+ gi and m/m/n/n transient distributions. *working paper*.

Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion, vol. 293 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Berlin: Springer-Verlag.

Robinson, P. M. (1994). Efficient tests of nonstationary hypotheses. *Journal of the american statistical association*, 89(428):1420–1437.

Sensier, M. and Van Dijk, D. (2004). Testing for volatility changes in us macroeconomic time series. *Review of Economics and Statistics*, 86(3):833–839.

Skorohod, A. (1961). Stochastic differential equations for a bounded region. *Theor. Probability Appl*, 6:264–274.

Stock, J. H. and Watson, M. W. (1988). Testing for common trends. *Journal of the American statistical Association*, 83(404):1097–1107.

Swensen, A. R. (2006). Bootstrap algorithms for testing and determining the cointegration rank in var models. *Econometrica*, 74(6):1699–1714.

Tang, L., Woodward, W. A., and Schucany, W. R. (2008). Undercoverage of wavelet-based resampling confidence intervals. *Communications in StatisticsSimulation and Computation®*, 37(7):1307–1315.

Trokić, M. (2013). Regulated fractionally integrated processes. *Journal of Time Series Analysis*, 34(5):591–601.

Trokic, M. (2014). Wavelet power: Wavelet energy ratio unit root tests. *Working Paper*, ():.

Wang, Q., Lin, Y., and Gulati, C. (2002). Asymptotics for general nonstationary fractionally integrated processes without prehistoric influence. *Journal of Applied Mathematics and Decision Sciences*, 6(4):255–269.

Wang, Q., Phillips, P. C., et al. (2009). Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory*, 25(3):710.

Ward, A. and Glynn, P. W. (2003). Properties of the reflected ornsteinuhlenbeck process. *Queueing Systems*, (44):109–123.

APPENDICES

A Proofs for Chapter 2

Proof of Lemma 1. Part i can be found in Theorem 1. of (Cavaliere and Taylor, 2007) and Remark 3.1. Part ii is from Proposition 3 of Cavaliere (2005b).

For part iii write the partial sum process for $\tilde{y}_t$ as following:

$$\tilde{y}_T(t) = T^{-d_1}\Delta_+^{-d_1} y_T(t) = T^{-1/2-d_1} \sum_{k=0}^{\lfloor Tt \rfloor -1} \pi_k(d_1) y_{\lfloor Tt \rfloor - k}$$
$$= T^{-1/2-d_1} \sum_{k=1}^{\lfloor Tt \rfloor} \pi_{\lfloor Tt \rfloor - k}(d_1) y_k$$

where $\pi_k(d_1) = \frac{\Gamma(k+d_1)}{\Gamma(d_1)\Gamma(k+1)}$, and Wang et al. (2002) claims that $\sum_{j=0}^{m} \pi_j(d_1) = \pi_m(d_1+1)$, thus we have:

$$\tilde{y}_T(t) = T^{-1/2-d_1} \sum_{k=1}^{\lfloor Tt \rfloor} \pi_{\lfloor Tt \rfloor - k}(d_1) y_k = T^{-1/2-d_1} \sum_{k=1}^{\lfloor Tt \rfloor} \pi_{\lfloor Tt \rfloor - k}(d_1) \sum_{j}^{k} u_k$$
$$= T^{-1/2-d_1} \sum_{k=1}^{\lfloor Tt \rfloor} \sum_{j=1}^{k} \pi_{\lfloor Tt \rfloor - k}(d_1) u_k = T^{-1/2-d_1} \sum_{k=1}^{\lfloor Tt \rfloor} \pi_{\lfloor Tt \rfloor - k}(d_1+1) u_k$$
$$= T^{-1/2-d_1} \sum_{k=1}^{\lfloor Tt \rfloor} \frac{(\lfloor Tt \rfloor - k)^{d_1}}{\Gamma(d_1+1)} u_k = T^{-1/2} \sum_{k=1}^{\lfloor Tt \rfloor} \frac{\left(\frac{\lfloor Tt \rfloor - k}{T}\right)^{d_1}}{\Gamma(d_1+1)} u_k$$
$$= T^{-1/2} \sum_{k=1}^{\lfloor Tt \rfloor} \frac{(t-k/T)^{d_1}}{\Gamma(d_1+1)} u_k = \sum_{k=1}^{\lfloor Tt \rfloor} \frac{(t-k/T)^{d_1}}{\Gamma(d_1+1)} T^{-1/2} \triangle y_k$$

here Δy_k can be written as $\int\limits_{(k-1)/T}^{k/T} dy_T(s)$ in the limit(see Phillips (1987b)). Then,

$$\tilde{y}_T(t) = \sum_{k=1}^{\lfloor Tt \rfloor} \frac{(t-k/T)^{d_1}}{\Gamma(d_1+1)} \int\limits_{(k-1)/T}^{k/T} dy_T(s) = \sum_{k=1}^{\lfloor Tt \rfloor} \int\limits_{(k-1)/T}^{k/T} \frac{(t-k/T)^{d_1}}{\Gamma(d_1+1)} dy_T(s) \tag{9}$$

$$\xrightarrow{w} \int\limits_0^t \frac{(t-s)^{d_1}}{\Gamma(d_1+1)} dy_T(s) \tag{10}$$

Note that last equality comes from the fact that $s \in [(k-1)/T, (k/T)]$ and as $T \to \infty$ $(k-1)/T$ and k/T will converge to the same limit, say s in this case. Then, since $y_T(r)$ and $(t-s)^{d_1}$ are continuous, we can apply FCLT and CMT to conclude that $\tilde{y}_T(t) \xrightarrow{w} \frac{C(1)}{\Gamma(d_1+1)} \int\limits_0^t (t-s)^{d_1} dJ_\omega^c(s)$.

Proof of Theorem 1. To prove part i, consider the residuals from the regression of δ_t on x_t for $t = 1 \dots T$, for $s \in [0,1]$:

$$\hat{x}_{\lfloor Ts \rfloor} = y_{\lfloor Ts \rfloor} - (\hat{\theta} - \theta)' \delta_{\lfloor Ts \rfloor} \tag{11}$$

$$T^{-1/2} \hat{x}_{\lfloor Ts \rfloor} = T^{-1/2} y_{\lfloor Ts \rfloor} - T^{-1/2} (\hat{\theta} - \theta)' \delta_{\lfloor Ts \rfloor} \tag{12}$$

We have already establish limiting distribution for first factor on the left hand side of equation (12). For second factor, define $N(T) = 1$ when $d_1 = 1$ and $N(T) = \left[\begin{smallmatrix} 1 & 0 \\ 0 & T^{-1} \end{smallmatrix}\right]$ when $d_t = [1, t]'$ and:

$$\begin{aligned}
&T^{-1/2} (\hat{\theta} - \theta)' \delta_{\lfloor Ts \rfloor} \\
&= \left(T^{-1} \sum_{t=1}^{T} T^{-1/2} y_s \delta_s N(T) \right) \left(T^{-1} \sum_{t=1}^{T} N(T) \delta_s \delta_s' N(T) \right)^{-1} N(T) \delta_{\lfloor Ts \rfloor} \\
&= \left(T^{-1} \sum_{t=1}^{T} T^{-1/2} y_s D_j(s/T) \right) \left(T^{-1} \sum_{t=1}^{T} D_j(s/T) D_j(s/T)' \right)^{-1} \\
&\times N(T) D_j(\lfloor Ts \rfloor / T)
\end{aligned}$$

by application of CMT and $D(\lfloor Ts \rfloor /T) \to D_j(s)$ we have

$$T^{-1/2}(\hat{\theta} - \theta)\delta_{\lfloor Ts \rfloor} \xrightarrow{w} \bar{\omega}C(1) \left(\int\limits_0^1 J_\eta^c(r) D_j(r) dr \right) \left(\int_0^1 D_j(r) D_j(r)' dr \right)^{-1} D(s)$$

Finally we have

$$T^{-1/2}\hat{x}_{\lfloor Ts \rfloor} \to \bar{\omega}C(1) J_\eta^c(s) - \bar{\omega}C(1) \left(\int\limits_0^1 J_\eta^c(r) D_j(r) dr \right) \left(\int_0^1 D_j(r) D_j(r)' dr \right)^{-1}$$
$$\times D(s) := J_{\eta,j}^c(s) \tag{13}$$

For part ii first consider

$$\begin{aligned} \tilde{\hat{x}}_t &= \Delta_+^{-d_1} y_t - T^{-d_1} \Delta_+^{-d_1} (\hat{\theta} - \theta) \delta_k \\ &= \sum_{k=0}^{T} \pi_k(d_1) y_{t-k} - (\hat{\theta} - \theta) \sum_{k=0}^{T} \pi_k(d_1) \delta_{t-k} \end{aligned}$$

We can write the partial sum process to find the limits

$$T^{-1/2-d_1} \tilde{\hat{x}}_t = T^{-1/2-d_1} \sum_{k=0}^{T} \pi_k(d_1) y_{t-k} - T^{-1/2-d_1} (\hat{\theta} - \theta) \sum_{k=0}^{T} \pi_k(d_1) \delta_{t-k}$$

First factor converges by Lemma 1 part iii. The second factor:

$$T^{-1/2}(\hat{\theta} - \theta) T^{-d_1} \sum_{k=0}^{T} \pi_k(d_1) \delta_{t-k}$$

here $T^{-1/2}(\hat{\theta} - \theta)' \to \bar{\omega}C(1) \left(\int\limits_0^1 J_\eta^c(s) D_j(s) ds \right) \left(\int_0^1 D_j(s) D_j(s)' ds \right)^{-1}$ by equation (13).

The convergence for $T^{-d_1} \sum\limits_{k=0}^{T} \pi_k(d_1) \delta_{t-k}$ is already proved by Nielsen (2009b), that is

$$T^{-d_1} \sum_{k=0}^{\lfloor Tt \rfloor} \pi_k(d_1) \delta_{\lfloor Tt \rfloor - k} \to \int_0^t \frac{(t-s)^{d-1}}{\Gamma(d)} D_j(s) ds$$

this establishes the proof.

Last part, iii, is derived by application of CMT using the objects we found in parts i-ii. □

Proof of Theorem 2. Part i directly follows from Theorem 3 of Cavaliere and Taylor (2007).

For part ii, define the partial sum $S_T(t) = T^{-1/2}\sum_{t=1}^{\lfloor Tt \rfloor} e_t$ and $\tilde{S}_T(t) = T^{-d_1}\Delta_+^{-d_1}S_T(t)$. Note that from equation (10) we have: $\tilde{S}_T(t) \xrightarrow{w} \int_0^t \frac{(t-s)^{d_1}}{\Gamma(d_1+1)}dS_T(s)$. Now, $B_{\hat{\eta}T}(s) := S_T(\hat{\eta}((\lfloor Ts \rfloor/T)T)$ and $B_{\hat{\eta},d_1,T}(s) = T^{-d_1}\Delta_+^{-d_1}S_T(\hat{\eta}((\lfloor Ts \rfloor/T)T)$, then we have $B_{\hat{\eta},d_1,T}(s) =\xrightarrow{w} \int_0^t \frac{(t-s)^{d_1}}{\Gamma(d_1+1)}dS_T(\hat{\eta}((\lfloor Ts \rfloor/T)T)$. But, from part i, Theorem 3 of (Cavaliere and Taylor, 2007) indicates that $S_T(\hat{\eta}((\lfloor Ts \rfloor/T)T) \xrightarrow{w} B_\eta(s)$. From CMT we obtain the result. □

Proof of Theorem 3. This theorem follows Nielsen (2009b) with slight modifications. Assume that y_t is stationary with finite second moment, in fact:

Size directly follows theorem 1, and test has size of α

$$T^{-1}\sum_{t=1}^{T} y_t^2 = \bar{\omega}^2$$

now for denominator Nielsen (2009b) showed that

$$var(\tilde{y}_t) = \sum_{m=0}^{t-1}\sum_{k=0}^{t-1}\pi_m(d_1)\pi_k(d_1)\gamma(m-k) \leq C\sum_{m=0}^{t-1}|\gamma(m)|\sum_{k=0}^{t-1-m}(k+m)^{d_1-1}k^{d_1-1}$$

Here, $C > 0$ is defined to be a generic constant by Nielsen (2009b). Now, if $d_1 < 1/2$

$$var(\tilde{y}_t) \leq \sum_{m=0}^{\infty}|\gamma(m)| < \infty$$

But, if $d_1 > 1/2$ from Nielsen (2009b)

$$T^{1-2d_1} var(\tilde{y}_t) \leq CT^{1-2d_1} \sum_{m=0}^{t-1} |\gamma(m)| \sum_{k=0}^{t-1-m} (k+m)^{d_1-1} k^{d_1-1}$$

$(k+m)^{d_1-1}k^{d_1-1} \leq k^{2d_1-2}$ if $d_1 < 1$ and $(k+m)^{d_1-1}k^{d_1-1} \leq T^{2d_1-2}$ if $d_1 \geq 1$, then this give us

$$T^{1-2d_1} var(\tilde{y}_t) \leq CT^{1-2d_1} \sum_{m=0}^{t-1} |\gamma(m)| \begin{cases} T^{-1} \sum_{k=0}^{T} (k/T)^{2d_1-2} \to \int_0^1 x^{2d_1-2} dx < \infty \\ T^{1-2d_1} \sum_{k=0}^{T} T^{2d_1-2} < \infty \end{cases}$$

Note that $var(\tilde{y}_t) \simeq T^{-1} \sum_{t=1}^{T} \tilde{y}_t^2$ Therefore, under the alternative hypothesis $\rho_\eta(d_1)$ is $O_p(max(T, T^{2d})$ which implies test statistic diver to ∞. Since our test rejects for large values of test statistic, we will achieve the asymptotic power against stationary alternatives.

B Proofs for Chapter 3

Lemma 4. Assume that $Z_t = \sum_{k=1}^{t} \psi(L)u_k$ where u_k is in equation 3.8 and further assume the assumptions $\mathcal{A}$ and $\mathcal{B}$ hold. Let $Z_T(t)$ be partial sum process for Z_t. If $\mathcal{D}[0,1]$ is endowed with the Skorohod topology, then as $T \to \infty$, $Z_T(t) \Rightarrow \sigma\psi(1)B^{\underline{c},\overline{c}}(t)$ where $t \in [0,1]$. In particular, we can define $\underline{c} = \underline{b}/\sqrt{\sigma^2 T}$ and $\overline{c} = \overline{b}/\sqrt{\sigma^2 T}$

Proof of Lemma 4. We first need to define the partial sum process

$$Z_T(t) = T^{-1/2} \sum_{i=1}^{\lfloor Tt \rfloor} v_t \tag{14}$$

where $v_t = \psi(L)(\epsilon_t + \underline{\xi}_{z,t} - \overline{\xi}_{z,t})$, $\underline{\xi}_{z,t}$ and $\overline{\xi}_{z,t}$ are defined in equations (3.10) and (3.11). And also define $u_t^* = \epsilon_t + \underline{\xi}_t - \overline{\xi}_t$.

Assume that the assumptions in Lemma 4 hold. We first apply Beveridge Nelson decomposition to the partial sum process defined in (14):

$$Z_T(t) = \psi(1)T^{-1/2}\sum_{i=1}^{\lfloor Tt\rfloor} u_t^* + T^{-1/2}(\tilde{u}_0 - \tilde{u}_t)$$

where $\tilde{u}_t = \sum_{j=0}^{\lfloor Tt\rfloor} \tilde{\psi}_j u^*_{\lfloor Tt\rfloor - j}$ and $\tilde{\psi}_j = \sum_{k=j+1}^{\infty} \psi_j$. Note that $\sup_t |u_t^*| = \sup_t |\epsilon_t + \underline{\xi}_t - \overline{\xi}_t| = O_p(T^{1/2})$ from assumptions $\mathcal{A}.1$ and $\mathcal{A}.4$. This implies that $T^{-1/2}\sum_{i=1}^{\lfloor Tt\rfloor} u_t - T^{-1/2}\sum_{i=1}^{\lfloor Tt\rfloor} u_t^* \xrightarrow{p} 0$. Additonally, from Cavaliere (2005a) theorem 1. $T^{-1/2}\sum_{i=1}^{\lfloor Tt\rfloor} u_t^* \xrightarrow{D} \sigma B^{\underline{c},\overline{c}}(t)$. Finally from CMT we have $Z_T(t) \xrightarrow{D} \sigma\psi(1)B^{\underline{c},\overline{c}}(t)$. □

Proof of Theorem 4. Consider following process $\{y_t\}$ defined by 3.6-3.9. Since $1 - \frac{c}{T} = e^{-c/T} + O(T^{-2})$, from Nielsen (2009b), we can write

$$y_t = \sum_{k=1}^{t} e^{-c_\rho(t-k)/T}\psi(L)u_k = \sum_{k=1}^{t}\psi(L)u_k + \sum_{k=1}^{t-1}\left(e^{-c_\rho(t-k)/T} - e^{-c_\rho(t-k-1)/T}\right) \times \sum_{m=1}^{k}\psi(L)u_m$$

(Nielsen, 2009a) shows that $e^{-c_\rho(t-k-1)/T} = e^{-c_\rho(t-k)/T} + \frac{c_\rho}{T}e^{-c_\rho(t-k)/T}\left(1 + O(T^{-1}\right)$. This implies

$$y_t = \sum_{k=1}^{t}\psi(L)u_k - \frac{c_\rho}{T}\sum_{k=1}^{t-1}\left(e^{-c_\rho(t-k)/T}\right)\sum_{m=1}^{k}\psi(L)u_m\left(1 + O(T^{-1}\right)$$

We can ignore $O(T^{-1})$ part in this representation since it will vanish in the limit. Then the partial sum process for y_t can be written as:

$$y_T(t) = \frac{1}{T^{1/2}}\sum_{k=1}^{\lfloor Tt\rfloor}\psi(L)u_k - \frac{1}{T^{1/2}}\frac{c_\rho}{T}\sum_{k=1}^{\lfloor Tt\rfloor - 1}\left(e^{-c_\rho(t-k)/T}\right)\sum_{m=1}^{k}\psi(L)u_m$$

Now define $Z_t(t) = \frac{1}{T^{1/2}}\sum_{k=1}^{\lfloor Tt\rfloor}\psi(L)u_k \xrightarrow{D} \sigma\psi(1)B^{\underline{c},\overline{c}}(t)$ from Lemma 4. Note that u_t is increments of bounded random walk process, thus from Cavaliere (2005a) the convergence is established. Now, we can write the partial sum of y_t in terms of partial sum process of Z_t:

$y_T(t) = Z_T(t) - \frac{c_\rho}{T} \sum_{k=1}^{\lfloor Tt \rfloor - 1} e^{-c_\rho (t-k)/T} Z_T(k)$

From CMT we obtain the result,

$$y_T(t) \xrightarrow{D} \sigma\psi(1)\left(B^{\underline{c},\bar{c}}(t) - c_\rho \int_0^t e^{-c_\rho(t-r)} B^{\underline{c},\bar{c}}(r)dr\right)$$

□

Proof of Theorem 5. First note that we can write the partial sum porcess for $\tilde{y}_t$ as follows:

$$\tilde{y}_T(t) = T^{-d_1}\Delta_+^{-d_1} y_T(t) = T^{-1/2-d_1} \sum_{k=0}^{\lfloor Tt \rfloor - 1} \pi_k(d_1) y_{\lfloor Tt \rfloor - k} = T^{-1/2-d_1}$$
$$\times \sum_{k=1}^{\lfloor Tt \rfloor} \pi_{\lfloor Tt \rfloor - k}(d_1) y_k$$

where $\pi_k(d_1) = \frac{\Gamma(k+d_1)}{\Gamma(d_1)\Gamma(k+1)}$, and $\sum_{j=0}^{m} \pi_j(d_1) = \pi_m(d_1+1)$ from Wang et al. (2002), we have:

Now,

$$
\begin{aligned}
\tilde{y}_T(t) &= T^{-1/2-d_1}\sum_{k=1}^{\lfloor Tt\rfloor}\pi_{\lfloor Tt\rfloor-k}(d_1)y_k = T^{-1/2-d_1}\sum_{k=1}^{\lfloor Tt\rfloor}\pi_{\lfloor Tt\rfloor-k}(d_1)\sum_{j}^{k}u_k\\
&= T^{-1/2-d_1}\sum_{k=1}^{\lfloor Tt\rfloor}\sum_{j=1}^{k}\pi_{\lfloor Tt\rfloor-k}(d_1)u_k = T^{-1/2-d_1}\sum_{k=1}^{\lfloor Tt\rfloor}\pi_{\lfloor Tt\rfloor-k}(d_1+1)u_k\\
&= T^{-1/2-d_1}\sum_{k=1}^{\lfloor Tt\rfloor}\frac{(\lfloor Tt\rfloor-k)^{d_1}}{\Gamma(d_1+1)}u_k = T^{-1/2}\sum_{k=1}^{\lfloor Tt\rfloor}\frac{\left(\frac{\lfloor Tt\rfloor-k}{T}\right)^{d_1}}{\Gamma(d_1+1)}u_k\\
&= T^{-1/2}\sum_{k=1}^{\lfloor Tt\rfloor}\frac{(t-k/T)^{d_1}}{\Gamma(d_1+1)}u_k = \sum_{k=1}^{\lfloor Tt\rfloor}\frac{(t-k/T)^{d_1}}{\Gamma(d_1+1)}T^{-1/2}\triangle y_k\\
&= \sum_{k=1}^{\lfloor Tt\rfloor}\frac{(t-k/T)^{d_1}}{\Gamma(d_1+1)}\int_{(k-1)/T}^{k/T}dy_T(s) = \sum_{k=1}^{\lfloor Tt\rfloor}\int_{(k-1)/T}^{k/T}\frac{(t-k/T)^{d_1}}{\Gamma(d_1+1)}dy_T(s)\\
&= \int_0^t\frac{(t-s)^{d_1}}{\Gamma(d_1+1)}dy_T(s)
\end{aligned}
$$

Note that last equality uses the fact that $s\in[(k-1)/T,k/T]$. Then, since $y_T(r)$ and $(t-s)^{d_1}$ are continuous, we can apply FCLT and CMT to conclude that $\tilde{y}_T(t)\Rightarrow\frac{\sigma\psi(1)}{\Gamma(d_1+1)}\int_0^t(t-k)^{d_1}dJ^{\underline{c},\bar{c}}(t)(k)$.

□

Proof of Theorem 6. The proof is straightforward application of CMT. We just divide the limiting distribution in theorem 4 with theorem 5. The long run variances in each object will vanish with division. □

Proof of Theorem 7. For part 1, first consider $\gamma=[0,0]$. Then for following process $\{\widetilde{\widetilde{y}}_t\}$ defined recursively by:

$$
\widetilde{\widetilde{y}}_t=\begin{cases}\left(1-\frac{c}{T}\right)\widetilde{\widetilde{y}}_{t-1}+u_t & \text{if}\quad\left(1-\frac{c}{T}\right)\widetilde{\widetilde{y}}_{t-1}+u_t\in[\underline{b},\bar{b}]\\ \bar{b} & \text{if}\quad\left(1-\frac{c}{T}\right)\widetilde{\widetilde{y}}_{t-1}+u_t>\bar{b}\\ \underline{b} & \text{if}\quad\left(1-\frac{c}{T}\right)\widetilde{\widetilde{y}}_{t-1}+u_t<\underline{b}\end{cases}
$$

where u_t is defined in (3.15) and set $v_t=\triangle_+^{-d}\epsilon_t$.

Since $1-\frac{c}{T}=e^{-c/T}+O(T^{-2})$, from **??**, we can write

$$\widetilde{\widetilde{y}}_t=\sum_{k=1}^{t}e^{-c(t-k)/T}u_k=\sum_{k=1}^{t}u_k+\sum_{k=1}^{t-1}\left(e^{-c(t-k)/T}-e^{-c(t-k-1)/T}\right)\sum_{m=1}^{k}u_m$$

(Nielsen, 2009a) shows that $e^{-c(t-k-1)/T}=e^{-c(t-k)/T}+\frac{c}{T}e^{-c(t-k)/T}\left(1+O(T^{-1})\right.$. This implies

$$\widetilde{\widetilde{y}}_t=\sum_{k=1}^{t}u_k-\frac{c}{T}\sum_{k=1}^{t-1}\left(e^{-c(t-k)/T}\right)\sum_{m=1}^{k}u_m\left(1+O(T^{-1})\right)$$

We can ignore $O(T^{-1})$ part in this representation. Now define following partial sum processes:

$$S_T(t)=\frac{1}{\kappa T^{1/2+d}}\sum_{k=1}^{\lfloor Tt\rfloor}v_k,\;\; L_T(t)=\frac{1}{\kappa T^{1/2+d}}\sum_{k=1}^{\lfloor Tt\rfloor}\underline{\xi}_k\;\; and\; U_T(t)=\frac{1}{\kappa T^{1/2+d}}\sum_{k=1}^{\lfloor Tt\rfloor}\overline{\xi}_k$$

Then,

$$\begin{aligned}\widetilde{\widetilde{y}}_T(t)&=\frac{1}{\kappa T^{1/2+d}}\sum_{k=1}^{t}u_k-\frac{1}{\kappa T^{1/2+d}}\frac{c}{T}\sum_{k=1}^{t-1}\left(e^{-c(t-k)/T}\right)\sum_{m=1}^{k}u_m\\&=S_T(t)+L_T(t)-U_T(t)-\frac{c}{T}\sum_{k=1}^{\lfloor Tt\rfloor-1}\left(e^{-c(t-k)/T}\right)\left(S_T(k)+L_T(k)-U_T(k)\right)\end{aligned}$$

Next, apply the broken line process to $\widetilde{\widetilde{y}}_t$ and create its continuous approximant on $\mathcal{C}[0,1]$ with $t\in[0,1]$ as follows:

$$S_T^*(t)=(\kappa T)^{-1/2-d}\sum_{i=1}^{[Tt]}v_i+v_{[Tt]+1}\left(\frac{Tt-[Tt]}{\kappa T^{1/2+d}}\right)$$

$$L_T^*(t) = \begin{cases} \left(\kappa^2(d)T^{2(d+1/2)}\right)^{-1/2} \sum_{i=1}^{[Tt]} \underline{\xi}_i & \text{if } \underline{\xi}_{\lfloor Tt \rfloor+1} = 0 \\ \left(\kappa^2(d)T^{2(d+1/2)}\right)^{-1/2} \sum_{i=1}^{[Tt]} \underline{\xi}_i & \text{if } \underline{\xi}_{\lfloor Tt \rfloor+1} > 0 \\ + \left(\kappa^2(d)T^{2(d+1/2)}\right)^{-1/2} \left(\frac{Tt-[Tt]-\frac{u_{[Tt]+1}-\underline{\xi}_{\lfloor Tt \rfloor+1}}{\underline{\xi}_{\lfloor Tt \rfloor+1}}}{1-\frac{u_{[Tt]+1}-\underline{\xi}_{\lfloor Tt \rfloor+1}}{\underline{\xi}_{\lfloor Tt \rfloor+1}}} \right) \underline{\xi}_{\lfloor Tt \rfloor+1} & \\ \times \mathbb{I}_{\left\{Tt \geq [Tt] + \frac{u_{[Tt]+1}-\underline{\xi}_{\lfloor Tt \rfloor+1}}{\underline{\xi}_{\lfloor Tt \rfloor+1}}\right\}} & \end{cases}$$

$$U_T^*(t) = \begin{cases} \left(\kappa^2(d)T^{2(d+1/2)}\right)^{-1/2} \sum_{i=1}^{[Tt]} \overline{\xi}_i & \text{if } \overline{\xi}_{\lfloor Tt \rfloor+1} = 0 \\ \left(\kappa^2(d)T^{2(d+1/2)}\right)^{-1/2} \sum_{i=1}^{[Tt]} \overline{\xi}_i & \text{if } \overline{\xi}_{\lfloor Tt \rfloor+1} > 0 \\ + \left(\kappa^2(d)T^{2(d+1/2)}\right)^{-1/2} \left(\frac{Tt-[Tt]-\frac{u_{[Tt]+1}-\overline{\xi}_{\lfloor Tt \rfloor+1}}{\overline{\xi}_{\lfloor Tt \rfloor+1}}}{1-\frac{u_{[Tt]+1}-\overline{\xi}_{\lfloor Tt \rfloor+1}}{\overline{\xi}_{\lfloor Tt \rfloor+1}}} \right) \underline{\xi}_{\lfloor Tt \rfloor+1} & \\ \times \mathbb{I}_{\left\{Tt \geq [Tt] + \frac{u_{[Tt]+1}-\overline{\xi}_{\lfloor Tt \rfloor+1}}{\overline{\xi}_{\lfloor Tt \rfloor+1}}\right\}} & \end{cases}$$

The above implies that

$$\widetilde{\widetilde{y}}_t^* = S_T^*(t) + L_T^*(t) - U_T^*(t) - \frac{c}{T} \sum_{k=1}^{\lfloor Tt \rfloor - 1} \left(e^{-c(t-k)/T}\right) \left(S_T^*(k) + L_T^*(k) - U_T^*(k)\right)$$

Notice that $S_T(t) \Rightarrow B_{d+1}(t)$, in other words, a Fractional Brownian motion. What needs to be demonstrated for the proof to follow however, is that $S_T^*(t) \Rightarrow B_{d+1}(t)$. This obtains by noting that

$$\begin{aligned}
\sup_{t\in[0,1]} |S_T^*(t) - S_T(t)| &= \sup_{t\in[0,1]} \left| v_{[Tt]+1} \left(\frac{Tt - [Tt]}{(\kappa(d)T^{(d+1/2)}} \right) \right| \\
&\leq \left(\kappa(d)T^{(d+1/2)}\right)^{-1} \max_{t=1,\ldots,T} |v_t| \\
&= o_P(1)
\end{aligned}$$

The above inequality follows from Cavaliere (2005a) with fractionally integrated innovations. Invoking Theorem 4.1 of Billingsley (1968) then implies that $S_T^*(t) \Rightarrow B_{d+1}(t)$.

The main part of the proof now follows by noting that $\tilde{\tilde{y}}_T^*(t)$ obtained above satisfies the construction for regulated Orstein-Uhlenbeck with fractional brownian motion increments. In particular, this construction is quite different than Harrison (1985)'s formula for regulated brownian motions. First, note that the function $\tilde{\tilde{y}}_T^*(t) = g_{n,\underline{c}}^{\overline{c}}(S_T^*(t), c)$ regulates $S_T^*(t)$ as a nearly integrated process which lies between $[\underline{c}, \overline{c}]$. Further, this process satisfy following three conditions:

1. The limits $L_T^*(s)$ and $U_T^*(s)$ are continuous and nondecreasing with $L_T^*(0) = U_T^*(0) = 0$
2. $\tilde{\tilde{y}}_T^*(t) = S_T^*(t) + L_T^*(t) - U_T^*(t) - \frac{c}{T} \sum_{k=1}^{\lfloor Tt \rfloor - 1} \left(e^{-c(t-k)/T}\right) (S_T^*(k) + L_T^*(k) - U_T^*(k) \in [\underline{c}, \overline{c}]$ for all t
3. $L_T^*(s)$ only increases when $\tilde{\tilde{y}}_T^*(s) = \underline{c}$ and $U_T^*(s)$ increases only when $\tilde{\tilde{y}}_T^*(s) = \overline{c}$.

However, from Reed et al. (2012), Ward and Glynn (2003), $\tilde{\tilde{y}}_T^*(t)$ is a unique solution to regulation problem for Orstein-Uhlenbeck process with locality to unit root parameter c_ρ. The continuing mapping theorem (CMT) and the fact that $S_T^*(t) \Rightarrow B_{d+1}(t)$ then entails that

$$\begin{aligned}\tilde{\tilde{y}}_T^*(t) &\Rightarrow g_{\underline{c}}^{\bar{c}}\left(B_{d+1}(t), c\right) \\ &= B_{d+1}(t) + L(t) - U(t) - c_\rho \int_0^t \left(e^{-c_\rho(t-k)}\right)\left((B_{d+1}(k) + L(k) - U(k)\right) dk\end{aligned}$$

What remains to be shown is that $\tilde{\tilde{y}}_T^*(t)$ converges to $\tilde{\tilde{y}}_T^*(t)$. A first step toward this result is to note the following:

$$\begin{aligned}&\left|\tilde{\tilde{y}}_T^*(t) - \tilde{\tilde{y}}_T(t)\right| \\ &= \left|\left(L_T^*(t) - \frac{L_{[Tt]}}{\kappa T^{1/2+d}}\right) - \left(U_T^*(t) - \frac{U_{[Tt]}}{\kappa(d)T^{(d+1/2)}}\right) + v_{[Tt]+1}\left(\frac{Tt - [Tt]}{\kappa(d)T^{(d+1/2)}}\right)\right| \\ &\leq \frac{2}{\kappa(d)T^{(d+1/2)}}\left|v_{[Tt]+1}\right|\end{aligned}$$

To show the inequality above holds, consider following objects:

$$\left|L_T^*(t) - \frac{L_{[Tt]}}{\kappa(d)T^{(d+1/2)}}\right| \;\; and \;\; \left|U_T^*(t) - \frac{U_{[Tt]}}{\kappa(d)T^{(d+1/2)}}\right|$$

These are both smaller than $\left(\kappa(d)T^{(d+1/2)}\right)^{-1}\left|v_{[Tt]+1}\right|$ and due to the fact that the nonzero set of $L_T^*(t)$ and $U_T^*(t)$ are clearly disjoint. Convergence follows by the same arguments used in the convergence of $S_T^*(t)$. That is,

$$\sup_{t\in[0,1]}\left|\tilde{\tilde{y}}_T^*(t) - \tilde{\tilde{y}}_T(t)\right| \leq \frac{2}{\kappa(d)T^{(d+1/2)}}\left|v_t\right| = o_P(1)$$

This completes the part of the proof which demonstrates that $\tilde{\tilde{y}}_T^*(t) \Rightarrow J_{d+1}^{c,\bar{c}}(t)$. To prove the postulate of Theorem 7 however, the convergence of $\tilde{\tilde{y}}_T(t)$ to $y_T(t)$ remains to be shown. This is completed in Lemma 5.

Lemma 5. Let $\{\tilde{\tilde{y}}_t\}$ and $\{y_t\}$ be defined as above. Then,

$$\max_{t=0,\ldots,T} \left|\tilde{\tilde{y}}_t - y_t\right| \leq \max\left(\max_{t=0,\ldots,T} \underline{\xi}_t, \max_{t=0,\ldots,T} \overline{\xi}_t\right)$$
$$\leq \max_{t=0,\ldots,T} \underline{\xi}_t + \max_{t=0,\ldots,T} \overline{\xi}_t$$

Proof of Lemma 5. This is in fact just another version of Lemma 7 in Cavaliere (2005a) and the proof there remains exactly valid here. This fact then implies that

$$\left|\tilde{\tilde{y}}_T(t) - y_T(t)\right| \Rightarrow 0$$

□

Then, we can conclude that $y_T(t) \Rightarrow J_{d+1}^{c,\bar{c}}(t)$. Now we are ready to introduce trend and drift terms. Define the partial sum process of x_t in (3.13) as

$$x_T(t) = c_\gamma t + y_T(t)$$

from assumption $\mathcal{C}$.1. But, we have already establish the limiting distribution of $y_T(t)$, thus from CMT we can say that:

$$x_T(t) \Longrightarrow J_{d+1}^{c,\bar{c}}(t) + c_\gamma t$$

□

Now for part 2, we begin with no deterministic terms. First note that we can write the partial sum process for $\tilde{x}_t = \tilde{y}_t$ as follows:

$$\tilde{y}_T(t) = T^{-d_1}\Delta_+^{-d_1} y_T(t) = T^{-d_1}T^{-1/2-d}\sum_{k=0}^{\lfloor Tt\rfloor-1}\pi_k(d_1)y_{\lfloor Tt\rfloor-k} = T^{-1/2-d_1-d}$$
$$\times\sum_{k=1}^{\lfloor Tt\rfloor}\pi_{\lfloor Tt\rfloor-k}(d_1)y_k$$

where $\pi_k(d_1) = \frac{\Gamma(k+d_1)}{\Gamma(d_1)\Gamma(k+1)}$.

Now, as in theorem 51

$$\tilde{y}_T(t) = T^{-1/2-d_1-d}\sum_{k=1}^{\lfloor Tt\rfloor}\frac{(\lfloor Tt\rfloor - k)^{d_1}}{\Gamma(d_1+1)}\triangle y_k = \sum_{k=1}^{\lfloor Tt\rfloor}\frac{(t-k/T)^{d_1-1}}{\Gamma(d_1)}T^{-1/2-d}\triangle y_k$$
$$= \sum_{k=1}^{\lfloor Tt\rfloor}\frac{(t-k/T)^{d_1-1}}{\Gamma(d_1)}\int_{(k-1)/T}^{k/T} dy_T(s)$$

Rest of the proof is the same as in theorem 5. From FCLT and CMT we can conclude that $\tilde{y}_T(t) \Rightarrow \frac{1}{\Gamma(d_1+1)}\int_0^t (t-s)^{d_1} dJ_{d+1}^{\underline{c},\bar{c}}(s)$. For trend and drift term

$$\tilde{x}_T(t) = \tilde{y}_T(t) + T^{-d_1-1/2}\Delta_+^{-d_1}\gamma\delta_t = \sum_{k=0}^{\lfloor Tt\rfloor-1}\pi_k(d_1)_{\lfloor Tt\rfloor-k}\gamma\delta_k \tag{15}$$
$$= \tilde{X}_T + T^{-1}\sum_{k=1}^{\lfloor Tt\rfloor}\frac{(t-k/T)^{d_1-1}}{\Gamma(d_1)}T^{-1/2-d}(\gamma_0+\gamma_1 k) \tag{16}$$
$$= \tilde{X}_T + T^{-1}\sum_{k=1}^{\lfloor Tt\rfloor}\frac{(t-k/T)^{d_1-1}}{\Gamma(d_1)}(T^{-1/2-d}\gamma_0 + c_\gamma(k/T)) \tag{17}$$
$$\Rightarrow \frac{1}{\Gamma(d_1+1)}\int_0^t (t-s)^{d_1} dJ_{d+1}^{\underline{c},\bar{c}}(s) + \gamma_c\frac{1}{\Gamma(d_1)}\int_0^t (t-s)^{d_1-1}s ds \tag{18}$$

Last line follows CMT and γ_0 vanish in the limit.

□

Proof of Lemma 2. The proof for this lemma directly follows CMT as in proof of Theorem 6. □

Proof of Theorem 8. First note that for $\gamma = [0,0]$ $\hat{x}_t(\underline{b},\bar{b}) = \hat{S}_t(\underline{b},\bar{b}) = y_t -$

$\hat{L}_y(t,\underline{b}) + \hat{U}_y(t,\overline{b})$ now define following partial sum processes

$$\hat{S}_T(t) = \frac{1}{T^{1/2}} \sum_{i=1}^{[Tt]} u_i$$
$$\hat{L}_y(t,\underline{b}) = \sum_{i=1}^{[Tt]} K(x_i - \underline{b})$$
$$\hat{U}_y(t,\overline{b}) = \sum_{i=1}^{[Tt]} K(x_i - \overline{b})$$

We can write $\hat{S}_t(\underline{b},\overline{b}) = y_t - L_t^* + U_t^* - (\hat{L}_t(\underline{b}) - L_t^*) + (\hat{U}_t(\overline{b}) - U_t^*)$, that is $\hat{S}_t(\underline{b},\overline{b}) = S_t - (\hat{L}_t(\underline{b}) - L_t^*) + (\hat{U}_t(\overline{b}) - U_t^*)$. Then, we can write: $\hat{S}_t(\underline{b},\overline{b}) = S_T(t) - (\hat{L}_T(t,\underline{b}) - L_T^*(t)) + (\hat{U}_T(t,\overline{b}) - U_T^*(t))$, where $L_t^*(t)$ and $U_T^*(t)$ is constructed as in Cavaliere (2005). First, $L_T^*(t) \xrightarrow{D} L_t$ from Cavaliere (2005), and $S_T(t) \xrightarrow{D} \sigma_s W(s)$, σ_s is long run variance of the process S_t. In Wang and Phillips (2009), it stated that:

$$\frac{c_T}{T} \sum_{i=0}^{[Tt]} K(c_T x_{i,T}) \xrightarrow{D} \tau L_X(t,0)$$

$L_X(t,0)$ is local time of process continous time process X_t at 0. Now set $c_T = \sqrt{T}$ and generalize $x_{i,T}(b) = T^{-1/2} \sum_{k=0}^{[Ti]} v_k - T^{-1/2} b$ where $v_k = \Delta S_k + \underline{\xi}_t - \overline{\xi}_t$. This implies that :

$$T^{-1/2} \sum_{i=0}^{[Tt]} K(c_T(y_T(t) - c)) \xrightarrow{D} \tau L_{y-c}(t,0) = L_y(t,c)$$

where $c = T^{-1/2} b$ and second equality comes from definition of local time which says that Local time of the process $y_t - c$ at 0 is same as local time of y_t at c, thus:

$$T^{-1/2} \sum_{i=0}^{[Tt]} K(y_{[Tt]} - \underline{b}) \xrightarrow{D} \tau L_y(t,\underline{c}) = L_t \xleftarrow{D} L_T^*(t)$$
$$T^{-1/2} \sum_{i=0}^{[Tt]} K(y_{[Tt]} - \overline{b}) \xrightarrow{D} \tau L_y(t,\underline{c}) = U_t \xleftarrow{D} U_T^*(t)$$

Note that $y_{[Tt]} = \sqrt{T} y_T(t)$. Since we are using gaussian kernel, we also have $\tau = \int_{-\infty}^{\infty} K_h(s)^2 ds = 1$. Here, we can see that after scaling with $T^{-1/2}$ the differences $(\hat{L}_t(\underline{b}) - L_t^*)$ and $(\hat{U}_t(\underline{b}) - U_t^*)$ will converge to zero in probability. Therefore, we can conclude that $S_T(\underline{b}, \overline{b})(t) \xrightarrow{D} \sigma_s J(t)$.

For part 2, since the differences $(\hat{L}_t(\underline{b}) - L_t^*)$ and $(\hat{U}_t(\underline{b}) - U_t^*)$ are $O_p(T^{1/2})$,

$$\begin{aligned}
\tilde{\hat{S}}_T(\underline{b}, \overline{b})(t) &= T^{-d_1} \triangle_+^{-d_1} \hat{S}_T(\underline{b}, \overline{b})(t) \\
&= T^{-1/2-d_1} \sum_{k=0}^{[Ts]} \pi_{[Ts]-k}(d_1) S_k + o_p(1)
\end{aligned}$$

Here first partial sum converges to $\widetilde{J}(t, d_1)$ as in Nielsen (2009a), thus we obtain the result.

for the series with trend and drift term, we can follow the same way to obtain limiting distributions as in the proof of theorem 7 since we only require CMT and assumption $\mathcal{C}$ with $d = 0$.

Proof of Lemma 3. From theorem 8 and CMT the nominator

$$T^{-2} \sum_{t=0}^{T} \widehat{X}_t^2 \longrightarrow \sigma_s^2 \int_0^1 J_i(s)^2 ds$$

and the denominator

$$T^{-2-2d_1} \sum_{t=0}^{T} \widehat{X}_t^2 \longrightarrow \sigma_s^2 \int_0^1 \widetilde{J}_i(s, d_1)^2 ds$$

Dividing these two terms and applying CMT, the proof is complete. □

C Proofs for Chapter 4

The proofs in this section mostly follow Nielsen (2010) and Trokic (2014). However, we will also have supporting lemmas.

Before these lemmas and proof of the theorem, we need to illustrate some basic identities. First recall assumption 6. First decompose u_t as:

$$u_t = \begin{bmatrix} u_t^{(1)} \\ u_t^{(2)} \end{bmatrix}$$

such that

$$R'Y_t = \begin{bmatrix} \triangle_+^{-d} u_t^{(1)} \\ \triangle_+^{-(d-b)} u_t^{(2)} \end{bmatrix}$$

Then we define $Z_t = \triangle_+^{-d} u_t^{(1)}$ and $w_t = \triangle_+^{-(d-b)} u_t^{(2)}$. Here $u_t^{(1)}$ is $p - r \times 1$ vector of error terms that correspond to non-stationary part of the decomposition and $u_t^{(2)}$ is $r \times 1$ vector of error term that correspond to stationary part of the decomposition. finally we can write $R'_{p-r} Y_t = Z_t$ and $R'_r Y_t = w_t$.

We can find some alternative forms of these fractional integrated processes.

$$Z_t = Z_{t-1} + \triangle_+^{-d+1} u_t^{(1)} \tag{19}$$

$$w_t = w_{t-1} + \triangle_+^{-d+b+1} u_t^{(2)} \tag{20}$$

For Haar wavelet, we can find these transformations as following:

$$R'V_{1,t} = \begin{bmatrix} Z_{2t} + Z_{2t-1} \\ w_{2t} + w_{2t-1} \end{bmatrix} \tag{21}$$

which can be written as:

$$R'V_{1,t} = \begin{bmatrix} 2Z_{2t-1} + \triangle_+^{-d+1} u_{2t}^{(1)} \\ 2w_{2t-1} + \triangle_+^{-(d-b)+1} u_{2t}^{(2)} \end{bmatrix}$$

or $R'_{p-r} V_{1,t} = 2Z_{2t-1} + \triangle_+^{-d+1} u_{2t}^{(1)}$ and $R'_r V_{1,t} = 2w_{2t-1} + \triangle_+^{-(d-b)+1} u_{2t}^{(2)}$ Now, we define following objects as in Nielsen (2010): $F_T^H = \sum_{t=1}^{t/2} R'V_{1,t}V'_{1,t}R$ and $G_T^H = \sum_{t=1}^{t/2} R'\tilde{V}_{1,t}\tilde{V}'_{1,t}R$ which can be decomposed as

$$F_T^H = \begin{bmatrix} \sum_{t=1}^{T/2} R'_{p-r}V_{1,t}V'_{1,t}R_{p-r} & \sum_{t=1}^{T/2} R'_{p-r}V_{1,t}V'_{1,t}R_r \\ \sum_{t=1}^{T/2} R'_rV_{1,t}V'_{1,t}R_{p-r} & \sum_{t=1}^{t/2} R'_rV_{1,t}V'_{1,t}R_r \end{bmatrix} = \begin{bmatrix} F_{11,T}^H & F_{12,T}^H \\ F_{12,T}^H & F_{22,T}^H \end{bmatrix}$$

$$G_T^H = \begin{bmatrix} \sum_{t=1}^{T/2} R'_{p-r}\tilde{V}_{1,t}\tilde{V}'_{1,t}R_{p-r} & \sum_{t=1}^{T/2} R'_{p-r}\tilde{V}_{1,t}\tilde{V}'_{1,t}R_r \\ \sum_{t=1}^{T/2} R'_r\tilde{V}_{1,t}\tilde{V}'_{1,t}R_{p-r} & \sum_{t=1}^{T/2} R'_r\tilde{V}_{1,t}\tilde{V}'_{1,t}R_r \end{bmatrix} = \begin{bmatrix} G_{11,T}^H & G_{12,T}^H \\ G_{12,T}^H & G_{22,T}^H \end{bmatrix}$$

Following Lemma will be useful for rest of the proofs:

Lemma 6. Assume that $u_t = \sum_{j=0}^{\infty} \tau_j \epsilon_{t-j}$ is a linear stationary process with the sequence $\{\epsilon_t\}$ are i.i.d white noise process and $\sum_{j=0}^{\infty} \|\tau_j\| < \infty$, then following statements are also true: Sequence $\{x_t\}$, where $x_t = u_{2t-k}$ for all $t = 1, 2, ...$ and for all $k = 0, 1, 2, ...$ is also linear stationary process with wold decomposition.

Proof. We can decompose x_t as

$$x_t = u_{2t-k} = \tau_0\epsilon_{2t-k} + \sum_{j=0}^{\infty} \tau_{2j+1}\epsilon_{2t-k-1-2j} + \sum_{j=0}^{\infty} \tau_{2j+2}\epsilon_{2t-k-2-2j}$$

here we can define the sequences $\epsilon_{t-j}^{(1)} = \epsilon_{2t-k-1-2j}$ and $\epsilon_{t-j}^{(2)} = \epsilon_{2t-k-2-2j}$ for all $t = 1, 2, ...,$ $k = 0, 1, ...$ and $j = 0, 1, ...$ Finally we can write:

$$x_t = u_{2t-k} = \tau_0\epsilon_{2t-k} + \sum_{j=0}^{\infty} \tau_j^{(1)}\epsilon_{t-j}^{(1)} + \sum_{j=0}^{\infty} \tau_j^{(2)}\epsilon_{t-j}^{(2)}. \tag{22}$$

To elaborate the notation let $k = 0$ then $\epsilon_t^{(1)} = ..., \epsilon_1, \epsilon_3, ...$ and $\epsilon_t^{(2)} = ..., \epsilon_2, \epsilon_4, ...$ Further define $\epsilon_t^{(3)} = \epsilon_{2t}$.

Notice that three components in equation 22 are independent and they do not have common ϵ_ts, they are also individually stationary. Stationarity of $\epsilon_t^{(3)}$ is obvious and other components exhibit wold decomposition since $\epsilon_t^{(1)}$ and $\epsilon_t^{(2)}$ are i.i.d white noise and the coefficient sequence are summable since $\sum_{j=0}^{\infty}\|\tau_j^{(1)}\|+\sum_{j=0}^{\infty}\|\tau_j^{(2)} = \|\sum_{j=0}^{\infty}\|\tau_j\| < \infty$. Then we can conclude that x_t is also a linear stationary process which has a Wold decomposition of form $x_t = \sum_{j=0}^{\infty}\tau_j^*\epsilon_{t-j}^*$ with $\sum_{j=0}^{\infty}\|\tau_j^*\| < \infty$ and ϵ_t^* is i.i.d white noise sequence with 0 mean and finite variance. □ □

Remark: We can simply say that long run variance of x_t is Φ^*.

Following Lemma is from Nielsen (2010) and will be very crucial for our analysis.

Lemma 7. Nielsen (2010) Lemma 5. Let $X_{j,t} = \sum_{k=0}^{\infty}\tau_{j,k}\epsilon_{j,t-k}$ for $j = 1,2$ be p_j-dimensional stationary linear processes with $\sum_{k=0}^{\infty}\|\tau_{j,k}\| < \infty$ for $p_j \times P_j$ matrices $\tau_{j,k}$ and p_j-dimensional i.i.d random variables $\epsilon_{j,t}$ with zero mean and finite fourth moments. Define the product moment $Q_{ijT}(u,v) = T^{-1}\sum_{t=1}^{T}\triangle_+^{-u}X_{i,t}\triangle_+^{-v}X_{j,t}'$,

(a) For $1/2 < v \leq u$ it holds that

$$T^{-u-v+1}Q_{ijT}(u,v) = O_p(1)$$

(b) For $v < 1/2 < u$

$$T^{-u+1/2}Q_{ijT}(u,v) \xrightarrow{p} 0$$

(c) For $v = 1/2 \leq u$

$$T^{-u-v+1}(logT)^{-\mathbb{1}_{\{u=1/2\}}}Q_{ijT}(u,v) = O_p(1)$$

The proof can be found in Nielsen (2010). Following Lemma generalize the Lemma 6 of Nielsen (2010). Now we define following limits:

$$\begin{aligned}
F_{11}^H &= 2^{2d}\Phi\int_0^1 B_{j,d}^{p-r}(s)B_{j,d}^{p-r}(s)'ds\Phi \\
G_{11}^H &= 2^{2d+2d_1}\Phi\int_0^1 B_{j,d+d1}^{p-r}(s)B_{j,d+d1}^{p-r}(s)'ds\Phi \\
F_{22}^H &= var(\triangle_+^{-(d-b)}u_{2t}^{(2)})
\end{aligned}$$

where, Φ^* is long run variance of $u_{2t}^{(1)}$. We also set $\phi = \max(d - b + d_1, 1/2)$

Lemma 8. Under the assumptions of Theorem 3, as $T \to \infty$

(a) $T_1^{-2d} F_{11T}^H = T_1^{-2d} \sum_{t=1}^{T/2} R'_{p-r} V_{1,t} V'_{1,t} R_{p-r} \xrightarrow{D} F_{11}^H$

(b) $T_1^{-1} F_{22T}^H = T_1^{-2d} \sum_{t=1}^{T/2} R'_r V_{1,t} V'_{1,t} R_r \xrightarrow{D} F_{22}^H$

(c) $T_1^{-d-1/2} F_{12T}^H = T_1^{-d-1/2} \sum_{t=1}^{T/2} R'_{p-r} V_{1,t} V'_{1,t} R_r \xrightarrow{p} 0$

(d) $T_1^{-2d-2d_1} G_{11T}^H = T_1^{-2d} \sum_{t=1}^{T/2} R'_{p-r} \tilde{V}_{1,t} \tilde{V}'_{1,t} R_{p-r} \xrightarrow{D} G_{11}^H$

(e) $T_1^{-2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} G_{22T}^H = T_1^{-2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} R'_r \tilde{V}_{1,t} \tilde{V}'_{1,t} R_r$

$$\begin{cases} \xrightarrow{D} var(\triangle_+^{-(d-b+d_1)} u_{2t}^{(2)}) & \textit{if } d - b + d_1 < 1/2 \\ = O_p(1) & \textit{if } d - b + d_1 \geq 1/2 \end{cases}$$

(f) $T_1^{-d-d_1-\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} G_{22T}^H = T_1^{-d-d_1-\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}}$

$$\times \sum_{t=1}^{T/2} R'_r \tilde{V}_{1,t} \tilde{V}'_{1,t} R_r \begin{cases} \xrightarrow{p} 0 & \textit{if } d - b + d_1 \leq 1/2 \\ = O_p(1) & \textit{if } d - b + d_1 > 1/2 \end{cases}$$

Proof. **of Lemma 8**

(a) First assume $\delta_t = 0$, we can define transformed wavelet coefficient as: $R_{p-r} V_{1,t} = \frac{1}{\sqrt{2}}(Z_{2t} + Z_{2t-1}) = \frac{1}{\sqrt{2}}(2Z_{2t-1} + \triangle_+^{-d+1} u_{2t}^{(1)})$ from equation 19. First, define $X_t = Z_{2t-1} = \sum_{i=1}^{2t} \Psi_{2t-i}(d) u_{i-1}^{(1)}$ and setting $v_t = u_{2t}^{(1)}$ which is linear stationary process with long run variance Φ^* from Lemma 6, we can also define $V_t = \triangle_+^{-d+1} v_t$. The partial sum processes for these objects can be written as:

$$\begin{aligned} T_1^{1/2-d} X_{T_1}(s) &= T_1^{1/2-d} \sum_{i=1}^{2[T_1 s]-1} \Psi_{2[T_1 s]-i}(d) u_{i-1} \\ &= 2^{-1/2+d} T^{1/2-d} \sum_{i=1}^{[Ts]-1} \Psi_{[Ts]-i}(d) u_{i-1} \\ T^{3/2-d} V_T(s) &= T^{3/2-d} \sum_{i=1}^{[Ts]-1} \Psi_{[Ts]-i}(d) v_i \end{aligned}$$

From Marinucci and Robinson (2002), $T_1^{1/2-d} X_{T_1}(s) \xrightarrow{D} 2^{-1/2+d} \Phi B_d^{p-r}(s)$ and $T^{3/2-d} V_T(s) \xrightarrow{D} \Phi^* W_{d-1}^{p-r}(s)$. But also note that $T_1^{1/2-d} V_T(s) \xrightarrow{p} 0$ as $T \to \infty$. Thus, we can say that $T_1^{1/2-d} R_{p-r} V_{1T}(s) \xrightarrow{p} 2^d \Phi W_d^{p-r}(s)$, since the $V_T(s)$ component vanishes. Now we can write:

$$T_1^{-2d} F_{11T}^H = T_1^{-2d} \sum_{t=1}^{T/2} \frac{1}{2} R'_{p-r} V_{1T} V'_{1T} R_{p-r}$$

After establishing convergence of $T_1^{1/2-d} R'_{p-r} V_{1,T_1}(r)$, note that

$$T_1^{-1} \sum_{t=1}^{T/2} T_1^{1-2d} R'_{p-r} V_{1,t} V'_{1,t} R_{p-r} = \int_0^1 T_1^{1-2d} R'_{p-r} V_{1,T_1}(s) V_{1,T_1}(s)' R_{p-r} ds$$

from continuous mapping theorem we can deduce that

$$T_1^{-1} \sum_{t=1}^{T/2} T_1^{1-2d} R'_{p-r} V_{1,t} V'_{1,t} R_{p-r} \xrightarrow{D} 2^{2d} \Phi \int_0^1 W_d^{p-r}(s) W_d^{p-r}(s)' ds \Phi = F_{11}^H$$

Also note that according to lemma 7.a one can see that this product moment is $O_p(T^{2d})$. Note that when $d = 1$ and $p = 1$ the same convergence rate is obtained by Fan and Gencay (2010).

Now assume $j \neq 0$, then we need to use $\hat{Y}_t$ instead of Y_t. Remark that $\hat{V}_{1,t} = \frac{1}{\sqrt{2}} \left(\hat{Y}_{2t} + \hat{Y}_{2t-1} \right)$. We can write $\hat{V}_{1,t} = V_{1,t} - \frac{1}{\sqrt{2}} (\hat{\alpha} - \alpha)' \tilde{\delta}_t$ where $\tilde{\delta}_t = 2$ if $j = 1$ and $\tilde{\delta}_t = [2, 4t-1]'$ if $j = 2$. As in previous part we need to find the distribution of $T_1^{1/2-d} R'_{p-r} \tilde{V}_{1,T_1}(s) = T_1^{1/2-d} R'_{p-r} V_{1,T_1}(s) - T_1^{1/2-d} 2^{-1/2} R'_{p-r} (\hat{\alpha} - \alpha)' \tilde{\delta}_{[T_1 s]}$. We have already prove the first part now consider:

$$\begin{aligned} T_1^{1/2-d} 2^{-1/2} R'_{p-r} (\hat{\alpha} - \alpha)' \tilde{\delta}_{[Ts]} &= T^{1/2-d} 2^{-1+d} R'_{p-r} \\ &\quad \times (\hat{\alpha} - \alpha)' N_j(T)^{-1} N_j(T) \tilde{\delta}_{[T_1 s]} \end{aligned}$$

where $N_j(T) = 1$ if $j = 1$ and $N_j(T) = diag([1, T^{-1}])$. We will divide this object into two segment. For first segment:

$$T^{1/2-d}2^{-1+d}R'_{p-r}(\hat{\alpha} - \alpha)' =$$

$$2^{-1+d}\left(T^{-1}\sum_{t=1}^{T} T^{1/2-d}R_{p-r}Y_t\delta_t' N_j(T)\right)\left(T^{-1}\sum_{t=1}^{T} N_j(T)\delta_t\delta_t' N_j(T)\right)^{-1} \xrightarrow{D}$$

$$2^{-1+d}\left(\Phi\int_0^1 W_d^{p-r}(s)D_j(s)'ds\right)\left(\int_0^1 D_j(s)D_j(s)'ds\right)^{-1} \tag{23}$$

where $D_j(s) = 1$ for $j = 1$ and $D_j(s) = [1, s]'$ for $j = 2$. The second part $N_j(T)\tilde{\delta}_{[Ts]} \to D_j(s)$ for $0 \leq s \leq 1$. Notice that $\tilde{\delta}_{[T_1 s]} = [2\ 4\lfloor T_1 s\rfloor] = [2\ 2\lfloor Ts\rfloor]$. Therefore we can conclude:

$$T_1^{1/2-d}R'_{p-r}\hat{V}_{1,T_1}(s) \xrightarrow{D} 2^d\Phi B_{j,d}^{p-r} = 2^d\Bigg(\Phi W_d^{p-r}(s)$$
$$-\left(\Phi\int_0^1 W_d^{p-r}(s)D_j(s)'ds\right)\times\left(\int_0^1 D_j(s)D_j(s)'ds\right)^{-1} D_j(s)\Bigg)$$

Finally, plugging this into test statistic we obtain:

$$T_1^{-1}\sum_{t=1}^{T/2} T_1^{1-2d}R'_{p-r}\hat{V}_{1,t}\hat{V}'_{1,t}R_{p-r} \xrightarrow{D} 2^{2d}\Phi\int_0^1 B_{j,d}^{p-r}(s)B_{j,d}^{p-r}(s)'ds\Phi = F_{11}^H$$

□

(b) Assume $\delta_t = 0$. We decompose this object to 3 components

$$
\begin{aligned}
T_1^{-1}F_{22T} &= T_1^{-2d}\sum_{t=1}^{T/2} R_r' \frac{1}{2}(Y_{2t}+Y_{2t-1})(Y_{2t}+Y_{2t-1})'R_r \qquad (24)\\
&= T_1^{-1}\sum_{t=1}^{T/2}\frac{1}{2}(w_{2t}+w_{2t-1})(w_{2t}+w_{2t-1})'\\
&= T_1^{-1}\sum_{t=1}^{T/2}\frac{1}{2}(w_{2t}w_{2t}'+2w_{2t}w_{2t-1}'+w_{2t-1}w_{2t-1}')\\
&= T_1^{-1}\sum_{t=1}^{T/2}\frac{1}{2}w_{2t}w_{2t}'+T_1^{-1}\sum_{t=1}^{T/2}w_{2t}w_{2t-1}'+T_1^{-1}\sum_{t=1}^{T/2}\frac{1}{2}w_{2t-1}w_{2t-1}'
\end{aligned}
$$

First and third summations will converge to $\frac{1}{2}var(\triangle_+^{-(d-b)}u_{2t}^{(2)})$ from lemma 6 and Nielsen (2010) lemma 6.b. But second summation is $O_p(T^{2(d-b)})$ which has slower convergence rate then T because of lemma 7. Thus, this term will vanish after scaling with T_1^{-1}. As a result $T_1^{-1}F_{22T}^H \xrightarrow{D} F_{22}^H$.

Now assume $\delta_t \neq 0$ we can rewrite equation 24 as:

$$
\begin{aligned}
T_1^{-1}F_{22T} &= T_1^{-1}\sum_{t=1}^{T/2} R_r'(V_{1,t}-\frac{1}{\sqrt{2}}(\hat{\alpha}-\alpha)'\tilde{\delta}_t)(V_{1,t}-\frac{1}{\sqrt{2}}(\hat{\alpha}-\alpha)'\tilde{\delta}_t)'R_r\\
&= T_1^{-1}\sum_{t=1}^{T/2} R_r'V_{1,t}V_{1,t}'R_r-\frac{2}{\sqrt{2}}T_1^{-1}\sum_{t=1}^{T/2}R_r'V_{1,t}\tilde{\delta}_t'(\hat{\alpha}-\alpha)R_r\\
&+ T_1^{-1}\sum_{t=1}^{T/2}R_r'(\hat{\alpha}-\alpha)'\tilde{\delta}_t\tilde{\delta}_t'(\hat{\alpha}-\alpha)R_r \qquad (25)
\end{aligned}
$$

In equation 25, we have found the distribution of first summation. For second object in this equation, consider

$$
T_1^{-1}\sum_{t=1}^{T/2} R_r'V_{1,t}\tilde{\delta}_t N_j(T) = O_p(T^{d-b-1/2}) \qquad (26)
$$

since $R_r'V_{1,t}$ part should be scaled by $T^{1/2-(d-b)}$ for proper convergence from

FCLT. Further, next part will be

$$\begin{aligned} N_j(T)^{-1}(\hat{\alpha}-\alpha)R_r &= \left(T^{-1}\sum_{t=1}^{T} N_j(T)\delta_t\delta_t' N_j(T)\right)^{-1} \\ &\times \left(T^{-1}\sum_{t=1}^{T} N_j(T)\delta_t R_r Y_t'\right) \\ &= O_p(^{d-b-1/2}) \end{aligned}$$

with similar argument as above. Then we can conclude that this component is $O_p(T^{2(d-b)-1})$. Since $2(d-b)-1<0$ this part converges to zero in probability.

Similarly, final component

$$T_1^{-1}\sum_{t=1}^{T/2} R_r'(\hat{\alpha}-\alpha)'\tilde{\delta}_t\tilde{\delta}_t'(\hat{\alpha}-\alpha)R_r = O_p(T^{2(d-b)-1}) \tag{27}$$

from above arguments, then this part also vanishes as $T\to\infty$ and the result is same as in $j=0$. □

(c) For $j=0$ we have

$$\begin{aligned} T_1^{-d-1/2}F_{12T}^H &= T_1^{-d-1/2}\sum_{t=1}^{T/2} R_{p-r}'\frac{1}{2}(Y_{2t}+Y_{2t-1})(Y_{2t}+Y_{2t-1})'R_r \\ &= T_1^{-d-1/2}\sum_{t=1}^{T/2}\frac{1}{2}(Z_{2t}+Z_{2t-1})(w_{2t}+w_{2t-1})' \\ &= T_1^{-d-1/2}\sum_{t=1}^{T/2}\frac{1}{2}(2Z_{2t-1}+u_{2t}^{(1)})(w_{2t}+w_{2t-1})' \\ &= T_1^{-d-1/2}\sum_{t=1}^{T/2}\frac{1}{2}(2Z_{2t-1}w_{2t}'+u_{2t}^{(1)}w_{2t}' \\ &+ 2Z_{2t-1}w_{2t-1}'+u_{2t}^{(1)}w_{2t-1}') \\ &= T_1^{-d-1/2}\sum_{t=1}^{T/2}\frac{1}{2}2Z_{2t-1}w_{2t}' + T_1^{-d-1/2}\sum_{t=1}^{T/2}\frac{1}{2}u_{2t}^{(1)}w_{2t}' \\ &+ T_1^{-d-1/2}\sum_{t=1}^{T/2}\frac{1}{2}2Z_{2t-1}w_{2t-1}' + T_1^{-d-1/2}\sum_{t=1}^{T/2}\frac{1}{2}u_{2t}^{(1)}w_{2t-1}' \end{aligned}$$

From lemma 7.b first and third summations converges to 0 in probability. Further, since $u_{2t}^{(1)}$ and w_{2t-1} stationary ergodic processes, from LLN second and fourth summations converges to a constant with rate T_1. However, since scale $T_1^{-d-1/2}$ is faster than this rate, both summation converges to zero in probability. This result give us $T_1^{-d-1/2}F_{12T}^H \xrightarrow{p} 0$.

Now assume $j = 1, 2$. Then

$$\begin{aligned}
T_1^{-d-1/2}F_{12T}^H &= T_1^{-d-1/2}\sum_{t=1}^{T/2} R'_{p-r}(V_{1,t} - \frac{1}{\sqrt{2}}(\hat{\alpha}-\alpha)'\tilde{\delta}_t) \\
&\times (V_{1,t} - \frac{1}{\sqrt{2}}(\hat{\alpha}-\alpha)'\tilde{\delta}_t)'R_r \\
&= T_1^{-d-1/2}\sum_{t=1}^{T/2} R'_{p-r}V_{1,t}V_{1,t}R_r \qquad (28) \\
&- \frac{1}{\sqrt{2}}T_1^{-d-1/2}\sum_{t=1}^{T/2} R'_{p-r}V_{1,t}\tilde{\delta}'_t(\hat{\alpha}-\alpha)R_r \qquad (29) \\
&- \frac{1}{\sqrt{2}}T_1^{-d-1/2}\sum_{t=1}^{T/2} R_{p-r}(\hat{\alpha}-\alpha)'\tilde{\delta}_t V'_{1,t}R_r \qquad (30) \\
&+ \frac{1}{2}T_1^{-d-1/2}\sum_{t=1}^{T/2} R_{p-r}(\hat{\alpha}-\alpha)'\tilde{\delta}_t\tilde{\delta}'_t\hat{\alpha}-\alpha)R_r \qquad (31)
\end{aligned}$$

We have already shown the convergence of element in equation 28 above. We will show remaining terms all converges to zero also.

First consider the object in equation 29; by Cauchy-Schwarz inequality we have it is

$$O_p\left(\left(T_1^{2d}\sum_{t=1}^{T_1}\mathbb{E}\,\|R'_{p-r}V_{1,t}\|^2\right)^{1/2} \times \left(T_1^{-1}\sum_{t=1}^{T_1}\mathbb{E}\,\|\tilde{\delta}'_t(\hat{\alpha}-\alpha)R_r\|^2\right)^{1/2}\right)$$

First factor is $O_p(1)$ from part a, and second factor is $o_p(1)$ from 27. Hence multiplication is $o_p(1)$ which implies that object in 29 converges to zero in probability. The object in 30 can be written as

$$T_1^{-d+1/2} R_{p-r}(\hat{\alpha}-\alpha) N_j(T) T_1^{-1} \sum_{t=1}^{T/2} N_j(T)^{-1} \tilde{\delta}_t V_{1,t}' R_r$$

From equation 26 $T_1^{-1} \sum_{t=1}^{T/2} N_j(T)^{-1} \tilde{\delta}_t V_{1,t}' R_r = O_p(T^{d-b-1/2})$, and further $T_1^{-d+1/2} R_{p-r}(\hat{\alpha}-\alpha) N_j(T)$ is $O_p(1)$ from part (a). The multiplication will be $O_p(T^{d-b-1/2})$ which indicates this object will converge to 0 also in probability.

Finally last object is $O_p(T^{d-b-1/2})$ following same arguments as 23 and 27. Since $d-b-1/2<0$ this object also converges to zero in probability. □

(d) For $j=0$ first consider $\tilde{V}_{1,T}(t)$.

$$\tilde{V}_{1,T}(t) = T_1^{-d_1} \Delta_+^{-d_1} V_{1,T}(t) = T_1^{-d-d_1+1/2} \sum_{k=0}^{\lfloor T_1 t \rfloor -1} \pi_k(d_1) V_{1,\lfloor T_1 t \rfloor - k}$$

$$= T_1^{-d-d_1+1/2} \sum_{k=1}^{\lfloor T_1 t \rfloor} \pi_{\lfloor T_1 t \rfloor - k}(d_1) V_{1,k}$$

Now define $V_{1,k} = \sum_{j=1}^{k} v_j$

$$
\begin{aligned}
\tilde{V}_{1,T}(t) &= T_1^{-d-d_1+1/2} \sum_{k=1}^{\lfloor T_1 t \rfloor} \pi_{\lfloor T_1 t \rfloor - k}(d_1) V_{1,k} \\
&= T_1^{-d-d_1+1/2} \sum_{k=1}^{\lfloor T_1 t \rfloor} \pi_{\lfloor T_1 t \rfloor - k}(d_1) \sum_{j=1}^{k} v_j \qquad (32) \\
&= T_1^{-d-d_1+1/2} \sum_{k=1}^{\lfloor T_1 t \rfloor} \sum_{j=1}^{k} \pi_{\lfloor T_1 t \rfloor - k}(d_1) v_j \\
&= T_1^{-1/2-d} \sum_{k=1}^{\lfloor T_1 t \rfloor} \pi_{\lfloor T_1 t \rfloor - k}(d_1+1) v_k \qquad (33) \\
&= T_1^{-d-d_1+1/2} \sum_{k=1}^{\lfloor T_1 t \rfloor} \frac{(\lfloor T_1 t \rfloor - k)^{d_1}}{\Gamma(d_1+1)} v_k \\
&= T_1^{-d+1/2} \sum_{k=1}^{\lfloor T_1 t \rfloor} \frac{\left(\frac{\lfloor T_1 t \rfloor - k}{T_1}\right)^{d_1}}{\Gamma(d_1+1)} v_k \qquad (34) \\
&= T_1^{-d+1/2} \sum_{k=1}^{\lfloor T_1 t \rfloor} \frac{(t - k/T_1)^{d_1}}{\Gamma(d_1+1)} v_k \\
&= \sum_{k=1}^{\lfloor T_1 t \rfloor} \frac{(t - k/T_1)^{d_1}}{\Gamma(d_1+1)} T_1^{-d+1/2} \triangle V_{1,k} \qquad (35)
\end{aligned}
$$

Here $T_1^{-d+1/2} \Delta V_{1,k}$ can be written as $\int_{(k-1)/T_1}^{k/T_1} dV_{1,T}(s)$ in the limit(see Phillips (1987b)), where $V_{1,T}(s)$ is partial sum process for $V_{1,t}$. Then,

$$
\begin{aligned}
\tilde{V}_{1,T}(t) &= \sum_{k=1}^{\lfloor T_1 t \rfloor} \frac{(t - k/T_1)^{d_1}}{\Gamma(d_1+1)} \int_{(k-1)/T_1}^{k/T_1} dV_{1,T}(s) \\
&= \sum_{k=1}^{\lfloor T_1 t \rfloor} \int_{(k-1)/T_1}^{k/T_1} \frac{(t - k/T_1)^{d_1}}{\Gamma(d_1+1)} dV_{1,T}(s) \qquad (36) \\
&\xrightarrow{w} \int_0^t \frac{(t-s)^{d_1}}{\Gamma(d_1+1)} dV_{1,T}(s) \qquad (37)
\end{aligned}
$$

But, $R'_{p-r} V_{1,T}(s) \to 2^d \Phi W_d^{p-r}(s)$ from part (a), then following convergence hold $R'_{p-r}\tilde{V}_{1,T}(s) \to 2^d \Phi W^{p-r}_{d+d_1}(s)$ following CMT.

Assume $j = 1, 2$. Then we need to consider

$$\begin{aligned} T_1^{-d-d_1+1/2}\tilde{\hat{V}}_{1,T_1}(s) &= T_1^{-d-d_1+1/2}\triangle_+^{-d_1}(R'_{p-r}V_{1,T_1}(s) - \tfrac{1}{\sqrt{2}}R'_{p-r}(\hat{\alpha}-\alpha)\tilde{\delta}_t) \\ &= T_1^{-d-d_1+1/2}\triangle_+^{-d_1}R'_{p-r}V_{1,T_1}(s) \\ &\quad -\tfrac{1}{\sqrt{2}}T_1^{-d-d_1+1/2}\triangle_+^{-d_1}R'_{p-r}(\hat{\alpha}-\alpha)\tilde{\delta}_t \\ &= T_1^{-d-d_1+1/2}R'_{p-r}\tilde{V}_{1,T_1}(s) \\ &\quad -\tfrac{1}{\sqrt{2}}T_1^{-d-d_1+1/2}\triangle_+^{-d_1}R'_{p-r}(\hat{\alpha}-\alpha)\tilde{\delta}_t \end{aligned}$$

As we mentioned first component will have convergence:

$$T_1^{1/2-d-d_1}R_{p-r}\tilde{V}_{1,T_1}(s) \xrightarrow{p} 2^d\Phi W^{p-r}_{d+d_1}(s)$$

We will focus on second object.

Let $P_T = T_1^{-d-d_1+1/2}\triangle_+^{-d_1}R'_{p-r}(\hat{\alpha}-\alpha)'\tilde{\delta}_t$, then

$$\begin{aligned} P_T &= T_1^{-d-d_1+1/2}2^{-1/2}\sum_{k=1}^{t}\Psi_{t-k}(d_1)R'_{p-r}(\hat{\alpha}-\alpha)\tilde{\delta}_k \\ &= 2^{-1/2}\left(T_1^{1/2-d}R'_{p-r}(\hat{\alpha}-\alpha)'N_j(T)^{-1}\right) \\ &\times \left(T_1^{-d_1}\right)\sum_{k=1}^{t}\Psi_{t-k}(d_1)N_j(T)\tilde{\delta}_k \end{aligned} \quad (38)$$

First factor converges $2^{-1+d}\left(\Phi\int_0^1 W_d^{p-r}(s)D_j(s)'ds\right)\left(\int_0^1 D_j(s)D_j(s)'ds\right)^{-1}$ from 23, second factor is deterministic and converges to $\int_0^s \frac{(s-r)^{d_1-1}}{\Gamma(d_1)}D_j(s)ds$ as in Nielsen (2010). Finally we obtain

$$\begin{aligned} T_1^{-d-d_1+1/2}\tilde{\hat{V}}_{1,T_1}(s) &\xrightarrow{D} 2^d\Bigg[\phi W^{p-r}_{d+d_1}(s) - \left(\Phi\int_0^1 W_d^{p-r}(s)D_j(s)'ds\right) \\ &\times \left(\int_0^1 D_j(s)D_j(s)'ds\right)^{-1}\int_0^s \frac{(s-r)^{d_1-1}}{\Gamma(d_1)}D_j(s)ds\Bigg] \end{aligned}$$

where

$$B_{j,d,d_1}^{p-r}(s) = \left[\phi W_{d+d_1}^{p-r}(s) - \left(\Phi \int_0^1 W_d^{p-r}(s) D_j(s)' ds\right)\right.$$
$$\left. \times \left(\int_0^1 D_j(s) D_j(s)' ds\right)^{-1} \int_0^s \frac{(s-r)^{d_1-1}}{\Gamma(d_1)} D_j(s) ds\right]$$

□

(e) First assume $j = 0$. When $d - b + d_1 < 1/2$ proof exactly follows as in b. When $d - b + d_1 > 1/2$ write

$$\begin{aligned} T_1^{2\phi} \quad (\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} G_{22T} &= T_1^{\phi} (\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} R_r' \tfrac{1}{2} \tilde{V}_{1,t} \\ &\quad \times (\tilde{V}_{1,t})' R_r \\ &= T_1^{2\phi} (\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} \tfrac{1}{2} \tilde{w}_{2t} \tilde{w}_{2t}' \\ &\quad + T_1^{2\phi} (\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} \tilde{w}_{2t} \tilde{w}_{2t-1}' \\ &\quad + T_1^{2\phi} (\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} \tfrac{1}{2} \tilde{w}_{2t-1} \tilde{w}_{2t-1}' \end{aligned} \tag{39}$$

We can apply lemma 7.(a) to all summations above and see all of these summations are $O_p(1)$.

When $d - b + d_1 = 1/2$ we can apply lemma 7(c) to all summation. Then all of these summations are $O_p(1)$.

Now assume $j = 1, 2$ and define $J_t = \triangle_+^{-d_1} = \tilde{\delta}_t'(\hat{\alpha} - \alpha)R_r$

$$\begin{aligned} T_1^{2\phi} \quad (\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} G_{22T} = T_1^{2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} R_r'(\tilde{V}_{1,t} - \tfrac{1}{\sqrt{2}} J_t) \\ \times(\tilde{V}_{1,t} - \tfrac{1}{\sqrt{2}} J_t)' R_r \\ = T_1^{2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} R_r' \tilde{V}_{1,t} \tilde{V}_{1,t}' R_r \qquad (40) \\ -\tfrac{1}{\sqrt{2}} T_1^{2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} R_r' \tilde{V}_{1,t} J_t' R_r \qquad (41) \\ -\tfrac{1}{\sqrt{2}} T_1^{2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} R_r'(J_t \tilde{V}_{1,t}' R_r \qquad (42) \\ +\tfrac{1}{2} T_1^{2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} J_t J_t \qquad (43) \end{aligned}$$

First note that $J_t = O_p(T^{d-b+d_1-1/2})$ from 38 if we plug R_r instead of R_{p-r}. Then object in 43 is

$$O_p\left(T_1^{2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \sum_{t=1}^{T/2} \mathbb{E}\left\| J_t \right\|^2 \right) = O_p(1)$$

Next, for 41 and 42, we apply Cauchy and Schwarz inequality then these objects become

$$O_p\left(T_1^{2\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \big(\sum_{t=1}^{T/2} \mathbb{E}\left\| J_t \right\|^2 \big)^{1/2} \big(\sum_{t=1}^{T/2} \mathbb{E}\left\| R_r' \tilde{V}_{1,t} \right\|^2 \big)^{1/2} \right) = O_p(1)$$

Since these three extra object is $O_p(1)$ the result does not change. □

(f) If $d - b + d_1 < 1/2$ proof is same as part (c). When $d - b + d_1 \geq 1/2$.

$$\begin{aligned} T_1^{-d-d_1-\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} G_{22T}^H = O_p(T_1^{-d-d_1-\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}} \\ \times \sum_{t=1}^{T/2} \mathbb{E}[R_{p-r}' \hat{\tilde{V}}_{1,t} \hat{\tilde{V}}_{1,t}' R_r]) \end{aligned}$$

now applying Cauchy-Schwarz inequality we have:

$$O_p\left(T_1^{-d-d_1-\phi}(\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}}G_{22T}^H\right) \leq O_p\left(\left(T_1^{-2d-2d_1}\sum_{t=1}^{T/2}\mathbb{E}\,\|R'_{p-r}\tilde{V}_{1,t}\|^2\right)^{1/2} \times\left(T_1^{-2\phi}(\log(T_1))^{-2\mathbb{1}_{\{\phi=1/2\}}}\right)\sum_{t=1}^{T/2}\mathbb{E}\,\|R'_r\tilde{V}_{1,t}\|^2\right)^{1/2}\right)$$

Now first part is $O_p(1)$ from part (d) and second part is $O_p((\log(T_1))^{-\mathbb{1}_{\{\phi=1/2\}}})$ which becomes $O_p(1)$ if $d-b+d_1 > 1/2$ and if $d-b+d_1 = 1/2$ this term will be $O_p(\log(T_1)^{-1})$ and it converges to 0 in probability.

□

Proof. **of Theorem 4**: Above lemma will be main engine of the proof of Theorem 10. The proof of theorem 4 will follow almost same as the proof of Theorem 1 of Nielsen (2010). The difference is that instead of scaling with some functions of T we scale with $T_1 = T/2$ and we will replace F and G matrices with corresponding F^H and G^H versions. Since we have same convergence properties for newly defined wavelet based moments, we will obtain same result as in Nielsen (2010). The proof below then mostly follows Nielsen (2010):

First note that our test statistic $\Lambda_{p,r}^H(d_1)$ is function of ordered eigenvalues of $A_T^H(B_T^H)^{-1}$. These are also eigen values of the eigen problem

$$|\lambda G_T^H - F_T^H| = 0 \tag{44}$$

since pre and post multiplication of A_T^H and B_T^H with orthonormal matrix R does not change the eigenvalues problem. Let,

$$M_T^H = \begin{bmatrix} T_1^{-d-d_1-1/2+\phi} I_{p-r} & 0_{p-r\times r} \\ 0_{r\times p-r} & T_1^{-1/2} I_r \end{bmatrix} \quad and$$
$$\overline{F}^H = \begin{bmatrix} 0_{p-r\times p-r} & 0_{p-r\times r} \\ 0_{r\times p-r} & F_{22}^H \end{bmatrix} \quad and \tag{45}$$

If we pre and post multiply F_T^H with M_T^H and apply lemma 6 we get:

$$M_T^H F_T^H M_T^H = \begin{bmatrix} T_1^{-2d-2d_1-1+2\phi} F_{11T}^H & T_1^{-d-d_1-1+\phi} F_{12T}^H \\ T_1^{-d-d_1-1+\phi} F_{12T}^{\prime w} & T_1^{-1} F_{22T}^H \end{bmatrix} \xrightarrow{p} \overline{F}^H \quad and \tag{46}$$
$$T_1^{1-2\phi} M_T^H G_T^H M_T^H = \begin{bmatrix} T_1^{-2d-2d_1} G_{11T}^H & T_1^{-d-d_1-\phi} G_{12T}^H \\ T_1^{-d-d_1-\phi} G_{12T}^{\prime w} & T_1^{-2\phi} G_{22T}^H \end{bmatrix},$$

where $T_1^{-2d-2d_1} G_{11T}^H \xrightarrow{D} G_{11}^H$ which is symmetric and positive definite a.s (From lemma 8(d)), $T_1^{-d-d_1-\phi} G_{12T}^H = O_p(1)$ (from lemma 8(f), $T_1^{-2\phi} G_{22T}^H = O_p(1)$ (from lemma 8(e)) and F_{22}^H is deterministic, symmetric, and positive definite.

Let $\rho = \lambda T_1^{2\phi-1}$. The ordered eigenvalues of (44) are the same as:

$$|\rho T_1^{1-2\phi} M_T^H G_T^H M_T^H - M_T^H F_T^H M_T^H| = 0$$

Now we define $\Upsilon(A,B)$ as the function that associates with a pair of matrices (A,B) and the ordered eigenvalues of the problem $|\rho A - B| = 0$. Note that this function is continous in its two matrix arguments and additionally we also have $(T_1^{1-2\phi} M_T^H G_T^H M_T^H, M_T^H F_T^H M_T^H)$ is tight, it follows from (46) that

$$\|\Upsilon(T_1^{1-2\phi} M_T^H G_T^H M_T^H, M_T^H F_T^H M_T^H) - \Upsilon((T_1^{1-2\phi} M_T^H G_T^H, \overline{F}^H)\| \xrightarrow{p} 0$$

The solutions to right hand side of convergence $\Upsilon((T_1^{1-2\phi}M_T^H G_T^H, \overline{F}^H)$ satisfy

$$
\begin{aligned}
0 &= \left| \rho \begin{bmatrix} T_1^{-2d-2d_1} G_{11T}^H & T_1^{-d-d_1-\phi} G_{12T}^H \\ T_1^{-d-d_1-\phi} G_{12T}^{\prime w} & T_1^{-2\phi} G_{22T}^H \end{bmatrix} \right| - \begin{bmatrix} 0_{p-r\times p-r} & 0_{p-r\times r} \\ 0_{r\times p-r} & F_{22}^H \end{bmatrix} \\
&= |\rho T_1^{-2d-2d_1} G_{11T}^H || \rho \Big(T_1^{-2\phi} G_{22T}^H - (T_1^{-d-d_1-\phi} G_{12T}^H)' \\
&\times (T_1^{-2d-2d_1} G_{11T}^H)^{-1} (G_{22T}^H - (T_1^{-d-d_1-\phi} G_{12T}^H)) \Big) - F_{22}^H | \qquad (47)
\end{aligned}
$$

We can use this decomposition since $T_1^{-2d-2d_1} G_{11T}^H$ is symmetric and positive definite a.s.

From the first absolute value in equation (47), $p_j \xrightarrow{p} 0 for j = 1, ..., p-r$ since $T_1^{-2d-2d_1} G_{11T}^H \xrightarrow{D} G_{11}^H$ which is symmetric and positive definite a.s. For the second absolute value in (47) $\nu = \rho^{-1}$ which solves

$$
\begin{aligned}
0 &= |\Big(T_1^{-2\phi} G_{22T}^H - (T_1^{-d-d_1-\phi} G_{12T}^H)' \\
&\times (T_1^{-2d-2d_1} G_{11T}^H)^{-1} (G_{22T}^H - (T_1^{-d-d_1-\phi} G_{12T}^H)) \Big) - \nu F_{22}^H | \\
&= |(F_{22}^H)^{-1/2} \Big(T_1^{-2\phi} G_{22T}^H - (T_1^{-d-d_1-\phi} G_{12T}^H)' \\
&\times (T_1^{-2d-2d_1} G_{11T}^H)^{-1} (G_{22T}^H - (T_1^{-d-d_1-\phi} G_{12T}^H)) \Big) F_{22}^H)^{-1/2} - \nu I_r | \qquad (48)
\end{aligned}
$$

where second equality holds because F_{22}^H is symmetric and positive definite. We utilize CMT and have $(T_1^{-2d-2d_1} G_{11T}^H)^{-1} \xrightarrow{D} (G_{11}^H)^{-1}$ since G_{11}^H is positive definite a.s. As a result, $(T_1^{-2d-2d_1} G_{11T}^H)^{-1} = O_p(1)$, further $T_1^{-d-d_1-\phi} G_{12T}^H = O_p(1)$ and $T_1^{-2\phi} G_{22T}^H = O_p(1)$. Then the solution of (48) satisfy $\nu_j = O_p(1)$ for j=1,...,r. This implies that $\rho_j = O_p(1)$ for j=n-r+1,...,n and consequently the r largest roots of (4.8) satisfy $\lambda_j^{-1} = O_p(T^{2\phi-1})$ for $j = n-r+1, ..., n$.

Let

$$
K_T^H = \begin{bmatrix} T_1^{-d} I_{p-r} & 0_{p-r\times r} \\ 0_{r\times p-r} & T_1^{-1/2} I_r \end{bmatrix}
$$

such that, by lemma 8(a-c)

$$K_T^H F_T^H K_T^H = \begin{bmatrix} T_1^{-2d} F_{11T}^H & T_1^{d-1/2} F_{12T}^H \\ T_1^{d-1/2} F_{12T}^{w'} & T_1^{-1} F_{22T}^H \end{bmatrix} \xrightarrow{D} \begin{bmatrix} F_{11}^H & 0_{p-r\times r} \\ 0_{r\times p-r} & F_{22}^H \end{bmatrix} \tag{49}$$

and

$$\begin{aligned} T_1^{-2d_1} K_T^H G_T^H K_T^H &= \begin{bmatrix} T_1^{-2d-2d_1} G_{11T}^H & T_1^{d-1/2-2d_1} G_{12T}^H \\ T_1^{d-1/2-2d_1} G_{12T}^{w'} & T_1^{-1-2d_1} G_{22T}^H \end{bmatrix} \\ &\xrightarrow{D} \begin{bmatrix} G_{11}^H & 0_{p-r\times r} \\ 0_{r\times p-r} & 0_{r\times r} \end{bmatrix} \end{aligned} \tag{50}$$

The ordered eigenvalues of (44) are the same as the eigenvalues of: $|\lambda K_T^H G_T^H K_T^H - K_T^H F_T^H K_T^H| = 0$ and setting $\mu = (T_1^{2d_1}\lambda)^{-1}$, these eigen values are same as those of

$$|T_1^{-2d_1} K_T^H G_T^H K_T^H - \mu K_T^H F_T^H K_T^H| = 0 \tag{51}$$

Since the eigenvalues are continous functions of the argument matrices, (49) and (50) indicates that the ordered eigenvalues of (51) converges in distribution to those of

$$\left| \begin{bmatrix} G_{11}^H & 0_{p-r\times r} \\ 0_{r\times p-r} & 0_{r\times r} \end{bmatrix} - \mu \begin{bmatrix} F_{11}^H & 0_{p-r\times r} \\ 0_{r\times p-r} & F_{22}^H \end{bmatrix} \right| = 0$$

This equation has r zero roots (since F_{22}^2 is deterministic and positive definite) and $p-r$ a.s. positive roots (because both G_{11}^H and F_{11}^H are symmetric and positive definite a.s) given by the solutions of $|G_{11}^H - \mu F_{11}^H| = 0$. Finally, $T_1^{2d_1}\lambda_1, ..., T_1 2d_1 \lambda_{p-r}$ converge in distribution to the solutions of $|T_1^{2d_1}\lambda G_{11}^H - F_{11}^H| = 0$, then $T_1^{2d_1} \sum_{j=1}^{p-r} \lambda_j \xrightarrow{D} tr\left\{F_{11}^H (G_{11}^w)^{-1}\right\}$. Finally note that the trace operator will remove the Ψ terms from $F_{11}^H (G_{11}^w)^{-1}$. □ □

The proof of theorem 5 directly follows proof of theorem 2 of Nielsen (2010).

Proof. **of Theorem 5** In proof of Theorem 10, it is shown that $\lambda_{p-r+1}^{-1}, ..., \lambda_{p-r_0}^{-1}$

are all $O_p(T^{2\phi-1})$. Hence, under the alternative, the test statistic $\Lambda^H_{p,r}(d_1) = T_1^{2d_1}\sum_{j=1}^{p-r_0}\lambda_j \geq T_1^{2d_1}\lambda_{p-r_0}$ and $(T_1^{2d_1}\lambda_{p-r_0})^{-1} = O_p(T^{2\phi-1-2d_1}$, nothing that $\lambda_j \geq 0$ for all j. We also have $2d_1 + 1 - 2\phi > 0$ when $d - b < 1/2$, it follows that $\Lambda^H_{p,r}(d_1)$ diverges as $T \to \infty$. Asymptotic size of the test follows from definition of $CV_{\xi,p-r_0}(d, d_1)$. □ □

Proof. **of Theorem 6** Now we only need to show convergence properties of new partial sum process for $\hat{V}^L_{1,t}$ is same as in Haar wavelet part. Now, with simple algebra and using equation 19 we can write:

$$\begin{aligned}
\hat{V}^L_{1,t} &= \sum_{l=0}^{L-1} g_l Y_{2t-l} - (\hat{\alpha}-\alpha)'\sum_{l=0}^{L-1} g_l \delta_{2t-l} \\
&= \sum_{l=0}^{L-1} g_l Y_{2t-L+1} + \sum_{l-0}^{L-1} g_l \sum_{j=0}^{L-2-l} u^{(1)}_{2t-l-j} - (\hat{\alpha}-\alpha)'\tilde{\delta}^L_t \\
&= \sqrt{2}Y_{2t-L+1} + \sum_{l-0}^{L-1} g_l \sum_{j=0}^{L-2-l} \triangle_+^{-d+1} u^{(1)}_{2t-l-j} - (\hat{\alpha}-\alpha)'\tilde{\delta}^L_t \\
&= V^L_{1,t} - (\hat{\alpha}-\alpha)'\tilde{\delta}^L_t
\end{aligned}$$

Define $S_t = R_{p-r}Y_{2t-L+1}$ and $U_{j,t,l} = R_{p-r}\triangle_+^{-d+1}u_{2t-l-j}$. Thus we have again two components: stochastic part and deterministic part. We can write the partial sum processes for stochastic part as following:

$$\begin{aligned}
T_1^{1/2-d} S_{T_1} &= T_1^{1/2-d}\sum_{i=1}^{2[T_1 r]} \Psi_{[2T_1 r]-i}(d) u_i^{(1)*} \\
&= 2^{d-1/2} T^{1/2-d} \sum_{i=1}^{[Tr]} \Psi_{[Tr]-i}(d) u_i^{(1)*} \\
T_1^{3/2-d} U_{j,T_1,l} &= T_1^{3/2-d}\sum_{i=1}^{2[T_1 r]} \Psi_{[2T_1 r]-i}(d-1) u_{i,j,l}^{(1)**} \\
&= 2^{d-3/2} T^{3/2-d} \sum_{i=1}^{[Tr]} \Psi_{[Tr]-i}(d-1) u_{i,j,l}^{(1)**}
\end{aligned}$$

where $u_i^{(1)*} = u^{(1)}_{2i-L+1}$ and $u_{i,j,l}^{(1)**} = u_{2t-l-j}$ for all $j = 0, ..., L-2-l$ and $j = 0, ..., L-1$. These sequences are shifted versions of original noise term. Note that these sequence converges by Marinucchi and Rabinson (2000) to limits:

$$T_1^{1/2-d}S_{T_1} = T_1^{1/2-d}\sum_{i=1}^{2[T_1r]}\Psi_{[2T_1r]-i}(d)u_i^{(1)*} \xrightarrow{D} 2^{d-1/2}\Phi B_d^{p-r} \tag{52}$$

$$T_1^{3/2-d}U_{j,T_1,l} = T_1^{3/2-d}\sum_{i=1}^{2[T_1r]}\Psi_{[2T_1r]-i}(d-1)u_{i,j,l}^{(1)**} = 2^{d-3/2}\Phi B_{d-1}^{p-r} \tag{53}$$

But the second part will vanish if we scale with $T_1^{1/2-d}$ as before. Then we can say that $T_1^{1/2-d}V_{1,T_1}^L(s) \xrightarrow{D} 2^{d-1/2}\Phi B_d^{p-r}$.

For deterministic part convergences can be derived almost same as in proof of lemma 8. But we need to change the definition of $\tilde{\delta}_t$ when $j=2$ as:

$$\begin{aligned}
\tilde{\delta}_t &= \sum_{l=0}^{L-1} g_l\delta_{2t-l} \\
&= \begin{bmatrix} \sum_{l=0}^{L-1} g_l & 2t\sum_{l=0}^{L-1} g_l - \sum_{l=0}^{L-1} g_l \end{bmatrix}' \\
&= \begin{bmatrix} \sqrt{2} & 2\sqrt{2}t - \sum_{l=0}^{L-1} g_l \end{bmatrix}' \\
&= \sqrt{2}\begin{bmatrix} 1 & 2t - \frac{1}{\sqrt{2}}\sum_{l=0}^{L-1} g_l \end{bmatrix}'
\end{aligned}$$

and if $j=1$ Then we can rewrite: $\tilde{\delta}_t = \sqrt{2}$. Finally, the limit process of this deterministic term $N_j(T)\tilde{\delta}_{\lfloor T_1 s\rfloor}$ can be written as

$$\sqrt{2}D_j(s) = \begin{bmatrix} 1 & s \end{bmatrix}'$$

But with this the distribution of $R'_{p-r}(\hat{\alpha}-\alpha)'\tilde{\delta}_t^L$ is identical to Haar wavelet case in the limit. Furthermore, it is also easy to show all other limits in lemma 8 are exactly identical. Then we can apply lemma 8 to this case and the proof of the theorem 11 is identical with proof of theorem 10. □