Stat
The ISI's Journal for the Rapid
Dissemination of Statistics Research

(wileyonlinelibrary.com) DOI: 10.100X/sta.0000

Supporting information for "A test for second-order stationarity of time series based on unsystematic sub-samples"

Haeran Cho [a*]

Received 13 September 2016; Accepted 3 October 2016

A. Choices of parameters

Choice of $\mathcal{M}$. Adopting the arguments in Fryzlewicz (2014), we can derive a theoretical lower bound for M to ensure that the convergence rate of the probability in Theorem 1 is suitably bounded. Suppose that $(1-\delta_T^2/T^2)^M \le T^{-\varepsilon}$ in order to match the term $CT^{-\varepsilon}$ in the RHS of the inequality in Theorem 1. Since $\log(1-x) \approx -x$ for x close to 0, we obtain $M \ge \varepsilon\delta_T^{-2}T^2 \log T$. When δ_T is as large as $\delta_T \asymp T$, this leads to $M \asymp \log T$, i.e., only a logarithmic number of random intervals are necessary to ensure the consistency of our test.

As δ_T is mostly unknown in practice, we propose to use $M = 2000$ for data of moderately large $T \le 1024$. In simulation studies, this choice of M worked well on various stationary and non-stationary models. When several other values of M were tested on the same datasets, this did not alter the results noticeably.

To facilitate the identification of the pairs of intervals which lead to large $|\mathcal{C}_j(p,q)|$, we propose to perform the sub-sampling by (i) first locating where the top ϵ largest and smallest ordinates of $I_{j,k}$ are for all $j = -1, \ldots, -J_T^*$, and (ii) randomly drawing M intervals around such locations. Besides, when any two intervals overlap, we modify the intervals to produce two pairs of disjoint intervals, since otherwise too many pairs are "thrown away". In the simulation studies, we have set $\epsilon = 5$.

Choice of the wavelets for computing $I_{j,k}$. Haar wavelets, defined as

$$\psi^H_{j,k} = 2^{j/2}\mathbb{I}(0 \le k \le 2^{-j-1}-1) - 2^{j/2}\mathbb{I}(2^{-j-1} \le k \le 2^{-j}-1), \quad k = 0, \ldots, 2^{-j}-1,$$

($\mathbb{I}(\cdot)$ is an indicator function returning $\mathbb{I}(\mathcal{A}) = 1$ if the event $\mathcal{A}$ is true and $\mathbb{I}(\mathcal{A}) = 0$ otherwise), form one of the simplest non-decimated wavelet systems. As our interest does not lie in accurate estimation of the EWS, we use Haar wavelets to compute the wavelet periodograms instead of other wavelets of higher regularity.

[a] School of Mathematics, University of Bristol, Bristol, BS8 1TW, UK

[*]Email: haeran.cho@bristol.ac.uk

Copyright © 2016 John Wiley & Sons, Ltd.
Prepared using staauth.cls [Version: 2012/05/12 v1.00]

Choice of J_T^*. In practice, we have access to a limited number of scales $j = -1, \ldots, -\lfloor \log_2 T \rfloor$. At coarser scales (corresponding to large $|j|$), the supports of the wavelet vectors increase in their lengths and thus the corresponding wavelet periodogram sequences are strongly autocorrelated as well as providing little information about the local second-order structure of X_t. Therefore, while still letting J_T^* increase slowly with T as in (A3), we choose to disregard wavelet periodograms at coarser scales by setting $J_T^* = [\log_2 \log_2 T]$. For the range of T considered in our simulation studies, this choice corresponds to $J_T^* = 3$.

Choice of m_T. As with M, the theoretical choice of m_T is directly linked to δ_T and therefore it is often inapplicable in practice. In simulation studies, we set $m_T = [\sqrt{T}]$ which appears to perform reasonably well in practice.

B. Real data analysis

In this section, we illustrate the application of the proposed test to egg price series at a German agricultural market observed weekly between April 1967 and May 1990. Following Paparoditis (2010) and Preuß et al. (2013), we took the differenced series as our input time series $\{X_t\}_{t=1}^T$ with $T = 1200$, to see whether there is any non-stationarity in the second-order structure of the original time series besides that in its level.

Applying the test procedure with $M = 2000$, our test rejected the null hypothesis at $\alpha = 0.05$. This agrees with the results reported in Paparoditis (2010) as well as those from applying the PVD, PP and JWW to the data. When applied to the first 1024 observations of the data, the HWTOS does not reject the null hypothesis whereas the opposite is observed when it is applied to the latter 1024 observations (its implementation in the R package `locits` (Nason, 2013) is applicable to the data of length 2^J only for some integer J).

Denoting $(\widehat{p}_j, \widehat{q}_j) = \arg\max_{(p,q)\in\mathcal{D}} \{\widehat{\sigma}_j(p,q)\}^{-1}|\mathcal{C}_j(p,q)|$, we plot the pairs of the intervals chosen at the scales $j = -1$ and $j = -2$ along with the egg price series X_t in Figure B.1, which are also the two scales where $\max_{(p,q)\in\mathcal{D}} \{\widehat{\sigma}_j(p,q)\}^{-1}|\mathcal{C}_j(p,q)|$ was greater than the test criterion. We also present the pointwise average of the equal weights given to the pairs of intervals which returned $\{\widehat{\sigma}_j(p,q)\}^{-1}|\mathcal{C}_j(p,q)|$ exceeding the critical value at $j = -1$ and -2. It is clear that $(L_{\widehat{p}_j}, L_{\widehat{p}_j})$ at the two scales lie close to each other and the weights from the two scales behave similarly as well. The HWTOS returned Haar wavelet coefficients computed over entire span of X_t, $t = 177, \ldots, 1200$, as significant, which covers $(L_{\widehat{p}_j}, L_{\widehat{p}_j})$ for both $j = -1$ and -2.

C. Proofs

Throughout the proof, C is used to denote a fixed, positive constant which may vary from one occasion to another.

C.1. Proof of Lemma 1

Proof of (a).
From (3),

$$\frac{1}{n_p}\left|\int_{s_p}^{e_p} \beta_j\left(\frac{y}{T}\right)dy - \sum_{k=s_p}^{e_p} \mathbb{E}(I_{j,k})\right| = \frac{1}{n_p}\left|\sum_{k=s_p}^{e_p}\int_k^{k+1} \beta_j\left(\frac{y}{T}\right)dy - \sum_{k=s_p}^{e_p}\beta_j\left(\frac{k}{T}\right)\right| + O(n_p^{-1})$$
$$= \frac{1}{n_p}\left|\sum_{k=s_p}^{e_p}\int_k^{k+1}\left\{\beta_j\left(\frac{y}{T}\right) - \beta_j\left(\frac{k}{T}\right)\right\}dy\right| + O(n_p^{-1}) \leq T^{-1}TV_{[0,1]}(\beta_j) + O(n_p^{-1}).$$

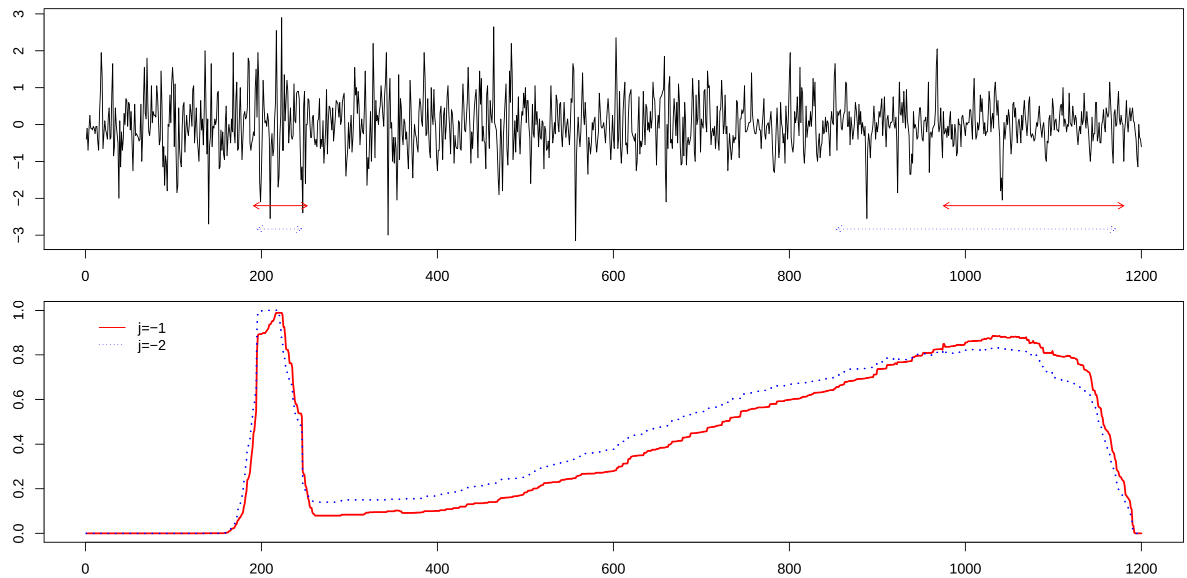

Figure B.1. Top: weekly egg price data with the intervals returned at scales $j = -1$ (upper arrows) and $j = -2$ (lower arrows); bottom: pointwise average of the weights given to L_p and L_q, which returned $\{\widehat{\sigma}_j(p,q)\}^{-1}|\mathcal{C}_j(p,q)|$ exceeding the test criterion, at $j = -1$ and $j = -2$.

Hence, $|\mathbb{E}\{\mathcal{C}_j(p,q)\} - \widetilde{\mathcal{C}}_j(p,q)| \le 2\sqrt{\frac{n_p n_q}{n_p+n_q}}(n_p^{-1} \vee n_q^{-1}) \asymp m_T^{-1/2}$. □

Proof of (b).

We can re-write $\mathcal{C}_j(p,q)$ *in a quadratic form involving* $\mathbf{X} = (X_0, \ldots, X_{T-1})^\top$, *as*

$$\begin{aligned}\mathcal{C}_j(p,q) &= \sum_k I_{j,k}\psi_k^{p,q} = \sum_k d_{j,k}^2 \psi_k^{p,q} = \sum_k \sum_t X_t \psi_{j,t-k} \sum_s X_s \psi_{j,s-k}\psi_k^{p,q} \\ &= \sum_t \sum_s X_t X_s \sum_k \psi_{j,t-k}\psi_{j,s-k}\psi_k^{p,q} =: \sum_t \sum_s X_t X_s r_{j,s,t}^{p,q} =: \mathbf{X}^\top \mathbf{R}_j^{p,q}\mathbf{X},\end{aligned}$$

where $\mathbf{R}_j^{p,q} = (r_{j,s,t}^{p,q})_{s,t}$ *is symmetric. Below, we omit* j, p, q *from the superscripts and subscripts, as* $r_{j,s,t}^{p,q} = r_{s,t}$ *and* $\mathbf{R}_j^{p,q} = \mathbf{R}$. *Then under (A2), Lemma 3.1 of Neumann & von Sachs (1997) shows that for* $\eta_T = \mathbf{Y}^\top \mathbf{R}\mathbf{Y}$ *with* $\mathbf{Y} = (Y_1, \ldots, T_T)^\top \sim \mathcal{N}_T(\mathbf{0}, \Sigma_T)$ *and* $\Sigma_T = cov(\mathbf{X})$,

$$\begin{aligned} cum_n(\mathbf{X}^\top \mathbf{R}\mathbf{X}) &= cum_n(\eta_T) + Q_n \text{ for } n = 2, 3, \ldots, \quad \text{where} \\ |cum_n(\eta_T)| &\le var(\eta_T) 2^{n-2}(n-1)!\{\lambda_{\max}(\mathbf{R})\lambda_{\max}(\Sigma_T)\}^{n-2} \\ Q_n &\le 2^{n-2}C_0^{2n}\{(2n)!\}^{1+\lambda}\max_{s,t}|r_{s,t}|\widetilde{\mathbf{R}}\|\mathbf{R}\|_\infty^{n-2}. \end{aligned} \tag{C.1}$$

Here, $\lambda_{\max}(\mathbf{R})$ *denotes the largest eigenvalue of* $\mathbf{R}$ *such that* $\lambda_{\max}(\mathbf{R}) \le \|\mathbf{R}\|_\infty = \max_s \sum_t |r_{s,t}|$, *and* $\widetilde{\mathbf{R}} = \sum_s \max_t |r_{s,t}|$. *In the definition of* $r_{s,t}$, *the summation of* $\psi_{j,t-k}\psi_{j,s-k}\psi_k^{p,q}$ *is taken over* $k \in (s - \mathcal{L}_j, s] \cap (t - \mathcal{L}_j, t] \cap (L_p \cup L_q)$ *due to the compactness of* $\boldsymbol{\psi}_j$, *and therefore*

$$\max_{s,t}|r_{s,t}| = \max_{s,t}|\sum_k \psi_{j,t-k}\psi_{j,s-k}\psi_k^{p,q}| \le \sqrt{\frac{n_p n_q}{n_p+n_q}}\left(\frac{1}{n_p} + \frac{1}{n_q}\right)\Psi_j(0) = O(m_T^{-1/2}). \tag{C.2}$$

Copyright © 2016 John Wiley & Sons, Ltd.

Similarly, since $|s-t| \le \mathcal{L}_j$ as well as $s \in [s_p, e_p + \mathcal{L}_j] \cup [s_q, e_q + \mathcal{L}_j]$ for non-zero $r_{s,t}$,

$$\|\mathbf{R}\|_\infty = \max_s \sum_t |r_{s,t}| \le \max_s \sum_{t:\,|s-t|\le\mathcal{L}_j} \sqrt{\frac{n_p n_q}{n_p+n_q}} \left(\frac{1}{n_p} + \frac{1}{n_q}\right) = O(\mathcal{L}_j m_T^{-1/2}), \tag{C.3}$$

$$\widetilde{\mathbf{R}} = \sum_s \max_t |r_{s,t}| \le \sqrt{\frac{n_p n_q}{n_p+n_q}} \left(\frac{n_p+\mathcal{L}_j}{n_p} + \frac{n_q+\mathcal{L}_j}{n_q}\right). \tag{C.4}$$

Further, $\lambda_{\max}(\Sigma_T) \le \max_s \sum_t |cov(X_s, X_t)| \le C_0^2(2!)^{1+\lambda} < \infty$ from (A1). According to (C.1)–(C.2) and (C.4), $var\{\mathcal{C}_j(p,q)\} = var(\eta_T) + Q_2$ where $\phi_j^{(3)}(p,q) = Q_2 = O(1)$. Following the practice taken in the proof of Lemma 3.2 in Neumann & von Sachs (1997), we focus on the case where $S_j(z)$ (and thus $c(z,\tau)$) smoothly vary over time with the regularity constant L_j ($L_j\mathcal{L}_j$). Extension to $S_j(z)$ of bounded total variation follows similarly.

Using the Gaussianity of $\mathbf{Y}$, *we have*

$$var(\eta_T) = 2tr(\mathbf{R}\Sigma_T\mathbf{R}\Sigma_T) = 2\sum_{s,t,v,w} r_{v,w} c_T\left(\frac{s+w}{2T}, s-w\right) r_{s,t} c_T\left(\frac{t+v}{2T}, t-v\right) := I, \quad \text{where}$$

$$c_T\left(\frac{t}{T},\tau\right) := cov(X_{t-\tau/2}, X_{t+\tau/2}) = \mathbb{E}\left(\sum_j\sum_k w_{j,k}\psi_{j,t-\tau/2-k}\xi_{j,k}\sum_{j'}\sum_{k'} w_{j',k'}\psi_{j',t+\tau/2-k'}\xi_{j',k'}\right)$$

$$= \sum_j\sum_k w_{j,k}^2\psi_{t-\tau/2-k}\psi_{t+\tau/2-k} = \sum_j\sum_k\left[S_j\left(\frac{t}{T}\right) + O\left\{\frac{L_j(t-k)}{T}\right\} + O\left(\frac{C_j}{T}\right)\right]\psi_{t-\tau/2-k}\psi_{t+\tau/2-k}$$

$$= \sum_j S_j\left(\frac{t}{T}\right)\Psi_j(\tau) + O\left(\frac{\sum_j L_j\mathcal{L}_j\Psi_j(\tau)}{T}\right) + O\left(\frac{\sum_j C_j\Psi_j(\tau)}{T}\right) = c\left(\frac{t}{T},\tau\right) + O\left(\frac{\log T}{T}\right),$$

with the last but one equality following from the fact that $|t-k| \le \mathcal{L}_j$ due to the compact support of $\psi_{j,k}$. Thus,

$$I = \sum_{s,t,v,w} r_{v,w} c\left(\frac{s+w}{2T}, s-w\right) r_{s,t} c\left(\frac{t+v}{2T}, t-v\right) + II =: III + II$$

where, as the support of the second argument of $c(z,\tau)$ is bounded as $|\tau| \le \mathcal{L}_j$ and $|v-w|, |s-t| \le \mathcal{L}_j$, $|II| \le T\mathcal{L}_j^3 m_T^{-1} O(\log T/T) = O(\mathcal{L}_j^3 m_T^{-1}\log T)$ using the bound obtained in (C.2).

Now we consider the effect of replacing $r_{s,t}$ by in III,

$$\Psi_j(s-t)\sqrt{\frac{n_p n_q}{n_p+n_q}}\left\{\frac{1}{n_p}\mathbb{I}(s\in[s_p,e_p],\ |s-t|\le\mathcal{L}_j) - \frac{1}{n_q}\mathbb{I}(s\in[s_q,e_q],\ |s-t|\le\mathcal{L}_j)\right\}$$

and similarly replacing $r_{v,w}$. Only those s,t,v,w that lie within the boundary of length $\mathcal{L}_j$ around $[s_p,e_p]$ or $[s_q,e_q]$, are affected by such replacement, and the size of such "boundary effect" is bounded by $\mathcal{L}_j^4 m_T^{-1}$. Further, the fact that $|s-v|, |s-w| \le C\mathcal{L}_j$ leads to

$$III = \sum_{s,t,v,w}\frac{n_p n_q}{n_p+n_q}\left\{\frac{1}{n_p^2}\mathbb{I}(s\in[s_p,e_p]) + \frac{1}{n_q^2}\mathbb{I}(s\in[s_q,e_q])\right\}\Psi_j(s-t)\Psi_j(v-w)$$
$$\times c\left(\frac{s+w}{2T}, s-w\right) c\left(\frac{t+v}{2T}, t-v\right) + O(\mathcal{L}_j^4 m_T^{-1}) =: IV + O(\mathcal{L}_j^4 m_T^{-1})$$

Copyright © 2016 John Wiley & Sons, Ltd.
Prepared using staauth.cls

where the constraints $\max\{|s-t|,|t-v|,|v-w|,|w-s|\} \le C\mathcal{L}_j$ *are omitted from the range of summations. Let* $s-w=\tau$ *and* $t-v=\tau'$*. Then,*

$$\begin{aligned} IV &= \sum_{s,v,\tau,\tau'} (\psi_s^{p,q})^2 \Psi_j(s-v-\tau)\Psi_j(s-v-\tau')\left\{c\left(\frac{s}{T},\tau\right)+O\left(\frac{L_j\mathcal{L}_j\tau}{2T}\right)\right\} \\ &\quad \times \left[c\left(\frac{s}{T},\tau'\right)+O\left\{\frac{L_j\mathcal{L}_j(v-s+\tau'/2)}{T}\right\}\right] \\ &= \sum_{s,v,\tau,\tau'} (\psi_s^{p,q})^2 \Psi_j(s-v-\tau)\Psi_j(s-v-\tau') c\left(\frac{s}{T},\tau\right) c\left(\frac{s}{T},\tau'\right) + O(\mathcal{L}_j^3 m_T^{-1}) \\ &=: V + O(\mathcal{L}_j^3 m_T^{-1}) \quad \text{using Definition 1 (b.iv), and} \\ V &= \sum_s (\psi_s^{p,q})^2 \sum_v \left\{\sum_\tau \Psi_j(s-v-\tau)c\left(\frac{s}{T},\tau\right)\right\}\left\{\sum_{\tau'} \Psi_j(s-v-\tau')c\left(\frac{s}{T},\tau'\right)\right\} \\ &= \sum_s (\psi_s^{p,q})^2 \sum_v \left\{\sum_\tau \Psi_j(s-v-\tau)c\left(\frac{s}{T},\tau\right)\right\}^2 =: \phi_j^{(1)}(p,q) + VI. \end{aligned}$$

It is easily shown that $VI = O(\mathcal{L}_j^3 m_T^{-1})$ *and and thus* $|\phi_j^{(2)}(p,q)| \le |II| + |III - IV| + |IV - V| + |VI| = O(\mathcal{L}_j^3 m_T^{-1} \log T)$*. In summary,*

$$var(\eta_T) = 2\sum_s (\psi_s^{p,q})^2 \sum_u \left\{\sum_\tau \Psi_j(\tau-u)c\left(\frac{s}{T},\tau\right)\right\}^2 + O(\mathcal{L}_j^3 m_T^{-1}\log T) + O(1).$$

□

Proof of (c).
From (C.2)–(C.4), $|cum_n(\mathcal{C}_j(p,q))| \le (n!)^{2\lambda+2} K^{n-2} (\mathcal{L}_j m_T^{-1/2})^{n-2}$ *for some* $K > 0$. □

C.2. Proof of Proposition 1

As noted above in Remark 1, $\sigma_j(p,q) \ge \sigma_*$ for some $\sigma_* > 0$ uniformly for all j and (p,q). Then,

$$\text{cum}\left(\frac{\mathcal{C}_j(p,q)}{\sigma_j(p,q)}\right) \le (n!)^{2\lambda+2}(K')^{n-2} m_T^{-(n-2)/2}$$

for some $K' > 0$. This implies that, using Lemma 1 of Rudzkis et al. (1978),

$$\mathbb{P}\left[\pm\frac{\mathcal{C}_j(p,q) - \mathbb{E}\{\mathcal{C}_j(p,q)\}}{\sigma_j(p,q)} \ge x\right] = (1-\Phi(x))(1+o(x))$$

uniformly for $0 \le x \le T^\nu$ for some $\nu > 0$, p and q. Following the same arguments adopted in the proof of Proposition 3.1 in Neumann & von Sachs (1997), we control the discrepancy between $\mathbb{E}\{\mathcal{C}_j(p,q)\}$ and $\widetilde{\mathcal{C}}_j(p,q)$ and conclude the proof. □

 Copyright © 2016 John Wiley & Sons, Ltd.

C.3. Proof of Proposition 2

From Proposition 1,

$$\mathbb{P}\left\{\max_{(p,q)\in\mathcal{D}}\max_{-J_T^*\le j\le -1}\left|\frac{\mathcal{C}_j(p,q)}{\sigma_j(p,q)}\right|\ge\Delta_T\right\}\le J_T^* M^2\mathbb{P}(|Z|\ge C_\delta\sqrt{\log T})\{1+o(\sqrt{\log T})\}$$
$$\le \frac{2J_T^* M^2}{C_\delta}\exp(-C_\delta^2\log T/2)\to 0$$

where Z denotes a standard normal variable, provided that $2\varpi < C_\delta^2/2$.

C.4. Proof of Theorem 1

Under (B2), the probability of randomly drawing M intervals so that none is contained in either $[a_1, b_1]$ or $[a_2, b_2]$, is bounded from above by $2\{1-c(1-c)\delta_T^2/T^2\}^M$. Denote such an event by $\mathcal{B}_T$. Then, on $\mathcal{B}_T^c$, there exists at least one pair of intervals in $\mathcal{M}$ such that one of which, say $L_{\widetilde{p}} = [s_{\widetilde{p}}, e_{\widetilde{p}}]$, is contained in $[a_1, b_1]$, and $L_{\widetilde{q}} = [s_{\widetilde{q}}, e_{\widetilde{q}}]$ in $[a_2, b_2]$. It trivially follows that $L_{\widetilde{p}}\cap L_{\widetilde{q}} = \emptyset$. From Proposition 1 and (B2)–(B3), we have

$$\begin{aligned}
\mathbb{P}(\mathcal{T}\le\Delta_T) &\le \mathbb{P}(\mathcal{T}\le\Delta_T|\mathcal{B}_T^c)+\mathbb{P}(\mathcal{B}_T)\\
&\le \mathbb{P}\left(\{\sigma_{j^*}(\widetilde{p},\widetilde{q})\}^{-1}|\mathcal{C}_{j^*}(\widetilde{p},\widetilde{q})|\le\Delta_T\right)+\mathbb{P}(\mathcal{B}_T)=: I+II,\quad\text{where}\\
I &\le \mathbb{P}\left\{\left|\frac{\mathcal{C}_{j^*}(\widetilde{p},\widetilde{q})-\widetilde{\mathcal{C}}_{j^*}(\widetilde{p},\widetilde{q})}{\sigma_{j^*}(\widetilde{p},\widetilde{q})}+\frac{\widetilde{\mathcal{C}}_{j^*}(\widetilde{p},\widetilde{q})}{\sigma_{j^*}(\widetilde{p},\widetilde{q})}\right|\le\Delta_T\right\}\\
&\le \mathbb{P}\left\{Z\ge\frac{1}{\sigma_{j^*}(\widetilde{p},\widetilde{q})}\sqrt{\frac{n_{\widetilde{p}}n_{\widetilde{q}}}{n_{\widetilde{p}}+n_{\widetilde{q}}}}\left|\frac{1}{n_{\widetilde{p}}}\int_{s_{\widetilde{p}}}^{e_{\widetilde{p}}}\beta_{j^*}(z)dz-\frac{1}{n_{\widetilde{q}}}\int_{s_{\widetilde{q}}}^{e_{\widetilde{q}}}\beta_{j^*}(z)dz\right|-\Delta_T\right\}\\
&\le \mathbb{P}\{Z\ge(\log T)^{-1/2}\nu_T m_T^{1/2}(\omega_1-\omega_2)/2\}\\
&\le \frac{2\sqrt{\log T}}{\sqrt{2\pi}\nu_T m_T^{1/2}(\omega_1-\omega_2)}\exp\left(-\frac{\nu_T^2 m_T(\omega_1-\omega_2)^2}{8\log T}\right)\le C\exp(-\varepsilon\log T)
\end{aligned}$$

with some $C, \varepsilon > 0$ and a standard normal variable Z, and $II\le 2\{1-c(1-c)\delta_T^2/T^2\}^M$. □

References

Fryzlewicz, P (2014), 'Wild binary segmentation for multiple change-point detection,' *Annals of Statistics*, **42**, pp. 2243–2281.

Nason, G (2013), *locits: Tests of stationarity and localized autocovariance*, R package version 1.4.

Neumann, MH & von Sachs, R (1997), 'Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra,' *Annals of Statistics*, **25**, pp. 38–76.

Paparoditis, E (2010), 'Validating stationarity assumptions in time series analysis by rolling local periodograms,' *Journal of the American Statistical Association*, **105**, pp. 839–851.

Preuß, P, Vetter, M & Dette, H (2013), 'A test for stationarity based on empirical processes,' *Bernoulli*, **19**, pp. 2715–2749.

Rudzkis, R, Saulis, L & Statulevičius, V (1978), 'A general lemma on probabilities of large deviations,' *Lithuanian Mathematical Journal*, **18**, pp. 226–238.

Copyright © 2016 John Wiley & Sons, Ltd.