Change-Point Analysis of Time Series with Evolutionary Spectra*

Alessandro Casini† — University of Rome Tor Vergata

Pierre Perron‡ — Boston University

5th March 2021

Abstract

This paper develops change-point methods for the time-varying spectrum of a time series. We focus on time series with a bounded spectral density that change smoothly under the null hypotheses but exhibits change-points or becomes rougher under the alternative. We provide a general theory for inference about the degree of smoothness of the spectral density over time. We address two local problems. The first is the detection of discontinuities (or breaks) in the spectrum at an unknown time and frequency. The second involves abrupt yet continuous changes in the spectrum over a short time period at an unknown frequency without signifying a break. We consider minimax-optimal testing and estimation. We determine the optimal rate for the minimax distinguishable boundary, i.e., the minimum break magnitude such that we are still able to uniformly control type I and type II errors. We propose a novel procedure for estimation of the change-points based on a wild sequential top-down algorithm and show its consistency under shrinking shifts and possibly growing number of change-points. Our method can be used across many fields and a companion program is made available in popular software packages.

AMS 2020 Classification: 62G05, 62G10, 62M15.
Keywords: Change-point, Locally stationary, Frequency domain, minimax-optimal test.

*We thank Federico Belotti, Leopoldo Catania and Stefano Grassi for generous help with computer programming.

†Corresponding author at: Department of Economics and Finance, University of Rome Tor Vergata, Via Columbia 2, Rome, 00133, IT. Email: alessandro.casini@uniroma2.it.

‡Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, US. Email: perron@bu.edu.

1 Introduction

Classical change-point theory focuses on detecting and estimating structural breaks in the mean or regression coefficients of time series models. Early contributions include, among others, Hinkley (1971), Yao (1987), Andrews (1993), Horváth (1993) and Bai and Perron (1998). These works assumed the presence of a single or multiple change-points in a parameter of an otherwise stationary time series model, see the reviews of Aue and Hórvath (2013) and Casini and Perron (2019) for more details. More recently there has been a growing interest on functional and time-varying parameter models where the latter are characterized by infinite-dimensional parameters which change continuously over time [see, e.g., Dahlhaus (1997), Hörmann and Kokoszka (2010), Zhang and Wu (2012), Panaretos and Tavakoli (2013), Aue, Dubart Nourinho, and Hormann (2015) and van Delft and Eichler (2018) and Aue and van Delft (2020)].

Change-point problems have been also studied in the frequency domain in several fields. Adak (1998) investigated the detection of change-points in piecewise stationary time series by looking at the distance in the power spectral density for two adjacent regimes. He compared several distance metrics such as a Kolmogorov-Smirnov distance, a Crámer-Von Mises distance and the CUSUM-type distance proposed by Coates and Diggle (1986). Last and Shumway (2008) focused on detecting change-points in piecewise locally stationary series. They exploited some of the results in Kakizawa, Shumway, and Taniguchi (1998) and Huang, Ombao, and Stoffer (2004) to propose a Kullback-Liebler discrimination information but did not derive the null distribution of the test statistic. Locally stationary series [cf. Dahlhaus (1997)] are defined by a time-varying spectral density that changes smoothly in time. In this setup change-points are discontinuities in the time-varying spectrum. A related yet distinct problem is the one recently considered by Aue and van Delft (2020) who provided tests for stationarity for functional time series using frequency domain methods.

In this paper we consider a more general change-point problem and propose inference methods about the changes in the degree of smoothness of the spectrum of a locally stationary series, and hence, also about change-points in the spectrum as a special case. The key parameter is the regularity exponent that governs how smooth or rough the path of the spectral density can be over time. We address two local problems. The first is the detection of discontinuities (or breaks) in the spectrum at an unknown time and frequency. The second involves the detection of abrupt yet continuous changes in the spectrum over a short time period at an unknown frequency without signifying a break (i.e., the spectrum becomes rougher over a short time period). We consider minimax-optimal testing and estimation for both problems, following the notion developed by Ingster (1993). We determine the optimal rate for the minimax distinguishable boundary, i.e., the minimum break magnitude such that we are still able to uniformly control type I and type II errors.

The problem of discriminating discontinuities from a continuous evolution in a nonparametric framework has received relatively less attention than the classical change-point problem with a few important exceptions [Müller (1992), Müller and Stadtmuller (1999), Spokoiny (1998), Wu and Zhao (2007) and Bibinger, Jirak, and Vetter (2017)]. These works focused on a time domain setting while we focus on a frequency domain setting. This adds a difficulty in that e.g., the search for a break has to run over two dimensions, the time and frequency point at which the change-point occurs. Our test statistics are maximum of local two-sample t-tests based on the local periodogram and local smoothed periodogram. We construct a family of tests that allows the researcher to test for a change-point in the spectrum at a prespecified frequency and a family of tests that allows to detect a break in the spectrum without prior knowledge about the frequency at which the change occurs. The asymptotic null distribution follows an extreme value distribution. In order to derive the null distribution of the tests, we establish several asymptotic results, including bounds, for higher-order cumulants and spectra of locally stationary processes. These results are complementary to some results in Dahlhaus (1997), Panaretos and Tavakoli (2013), Aue and van Delft (2020) and Casini (2021), and extend some of the classical frequency domain results for stationary processes in the time series literature [e.g., Brillinger (1975)] to locally stationary processes.

We also address the problem of estimating the change-points. We allow for the number of change-points to increase with the sample size and for the distance between change-points to shrink to zero. We propose a novel procedure based on a wild sequential top-down algorithm that exploits the idea of bisection and we combine it with the recent wild resampling approach in Fryzlewicz (2014). We establish the consistency of the procedure for the number of change-points and their locations. We compare the rate of convergence with that of standard change-point estimators under the classical setting [e.g., Yao (1987), Bai (1994), Casini and Perron (2021a), Casini and Perron (2020)]. We verify the performance of our methods via simulations which show the benefits from using our approach. The advantage of using frequency domain methods to detect change-points is that it does not require to make assumptions about the data-generating process under the null hypotheses beyond the fact that the spectrum is bounded. Furthermore, the method allows for a broader range of alternative hypotheses than time domain methods which usually have power against some specific alternatives but not against others. Our methods are readily available for use in many fields such as speech processing, biomedical signal processing, seismology, and failure detection, economics and finance. It is also used as a pre-test before constrcuting the recent double kernel long-run variance estimators to account more flexibly for nonstationarity [cf. Casini (2021), Casini and Perron (2021b) and Casini, Deng, and Perron (2021)]. The rest of the paper is organized as follows. Section 2 introduces the statistical setting and the hypotheses testing problem. Section 3 develops asymptotic results for higher-order cumulants and spectra of locally stationary processes. Section 4 presents the test statistics and derive their null distributions. Section 5 establishes the consistency of the tests and their minimax optimality. Section 6 discusses the estimation of the change-points while Section 7 provides details for the implementation of the methods. Monte Carlo simulations are conducted in Section 8. Section 9 reports conclusive comments. The supplemental material [cf. Casini and Perron (2021c)] contains all mathematical proofs. The code to implement our method is provided in Matlab, R and Stata languages through a Github repository.

2 Statistical Environment and the Hypotheses Testing Problem

Section 2.1 introduces the statistical setting and Section 2.2 presents the hypotheses testing problem. We work in the frequency domain under the locally stationary framework introduced by Dahlhaus (1997). Casini (2021) extended the latter framework to allow for discontinuities in the spectrum which then results in segmented locally stationary processes. This corresponds to the alternative hypotheses of change-points in the spectrum. Since local stationarity is a special case of segmented local stationarity we begin with the latter.

2.1 Segmented Locally Stationary Processes

Suppose $\{X_t\}_{t=1}^{T}$ is defined on an abstract probability space $(\Omega,\, \mathscr{F},\, \mathbb{P})$, where Ω is the sample space, $\mathscr{F}$ is the σ-algebra and $\mathbb{P}$ is a probability measure. We use an infill asymptotic setting whereby we rescale the original discrete time horizon $[1,\, T]$ by dividing each t by T. Let $i \triangleq \sqrt{-1}$.

Definition 2.1. A sequence of stochastic processes $X_{t,T}$ $(t = 1, \ldots,\, T)$ is called segmented locally stationarity (SLS) with $m_0 + 1$ regimes, transfer function A^0 and trend $\mu_\cdot$ if there exists a representation

$$X_{t,T} = \mu_j\,(t/T) + \int_{-\pi}^{\pi} \exp\,(i\omega t)\, A^0_{j,t,T}\,(\omega)\, d\xi\,(\omega)\,, \qquad \left(t = T^0_{j-1} + 1, \ldots,\, T^0_j\right), \tag{2.1}$$

for $j = 1, \ldots,\, m_0 + 1$, where by convention $T^0_0 = 0$ and $T^0_{m_0+1} = T$ and the following holds:

(i) $\xi(\omega)$ is a stochastic process on $[-\pi, \pi]$ with $\overline{\xi(\omega)} = \xi(-\omega)$ and

$$\text{cum}\{d\xi(\omega_1), \dots, d\xi(\omega_r)\} = \varphi\left(\sum_{j=1}^{r} \omega_j\right) g_r(\omega_1, \dots, \omega_{r-1}) \, d\omega_1 \dots d\omega_r,$$

where $\text{cum}\{\cdot\}$ is the cumulant of rth order, $g_1 = 0$, $g_2(\omega) = 1$, $|g_r(\omega_1, \dots, \omega_{r-1})| \leq M_r < \infty$ and $\varphi(\omega) = \sum_{j=-\infty}^{\infty} \delta(\omega + 2\pi j)$ is the period 2π extension of the Dirac delta function $\delta(\cdot)$.

(ii) There exists a constant $K > 0$ (which depends on j) and a piecewise continuous function $A : [0, 1] \times \mathbb{R} \to \mathbb{C}$ such that, for each $j = 1, \dots, m_0 + 1$, there exists a 2π-periodic function $A_j : (\lambda_{j-1}, \lambda_j] \times \mathbb{R} \to \mathbb{C}$ with $A_j(u, -\omega) = \overline{A_j(u, \omega)}$, $\lambda_j^0 \triangleq T_j^0/T$ and for all T,

$$A(u, \omega) = A_j(u, \omega) \text{ for } \lambda_{j-1}^0 < u \leq \lambda_j^0, \tag{2.2}$$

$$\sup_{1 \leq j \leq m_0+1} \sup_{T_{j-1}^0 < t \leq T_j^0, \omega} \left| A_{j,t,T}^0(\omega) - A_j(t/T, \omega) \right| \leq K T^{-1}. \tag{2.3}$$

(iii) $\mu_j(t/T)$ is piecewise continuous.

The smoothness properties of A in u guarantees that $V_{t,T}$ has a piecewise locally stationary behavior. We refer to Casini (2021) for several theoretical properties of SLS processes. Zhou (2013) considered piecewise locally stationary processes in a time domain setting but his notion is less general than the framework considered by Casini (2021). We collect the break dates in $\mathcal{T} \triangleq \{T_1^0, \dots, T_{m_0}^0\}$.

Assumption 2.1. *(i) $\{X_{t,T}\}$ is a mean-zero segmented locally stationary process; (ii) $A(u, \omega)$ is twice continuously differentiable in u at all $u \neq \lambda_j^0$, $j = 1, \dots, m_0 + 1$ with uniformly bounded derivatives $(\partial/\partial u) A(u, \cdot)$ and $(\partial^2/\partial u^2) A(u, \cdot)$; (iii) $A(u, \omega)$ is twice left-differentiable in u with uniformly bounded derivatives $(\partial/\partial_- u) A(u, \cdot)$ and $(\partial^2/\partial_- u^2) A(u, \cdot)$.*

Assumption 2.2. *(i) $A(u, \omega)$ is twice differentiable in ω with uniformly bounded derivatives $(\partial/\partial \omega) A(\cdot, \omega)$ and $(\partial^2/\partial \omega^2) A(\cdot, \omega)$; (ii) $g_4(\omega_1, \omega_2, \omega_3)$ is continuous in its arguments.*

2.2 The Hypotheses Testing Problem

We focus on time-varying spectra that are bounded, thereby excluding unit root and long memory processes. We consider the following classes of time-varying spectra,

$$\boldsymbol{F}(\theta, D) = \left\{ \{f(u, \omega)\}_{u \in [0,1], \omega \in [-\pi, \pi]} \mid \sup_{\omega \in [-\pi, \pi],} \sup_{u, v \in [0,1], |v-u| < h} |f(u, \omega) - f(v, \omega)| \leq D h^{\theta} \right\}, \tag{2.4}$$

for $D < \infty$. The parameter $\theta > 0$ is the regularity exponent of f in the time dimension. For $\theta > 1$ f is constant in u and reduces to the spectral density of a stationary process. For $\theta = 1$ f is Lipschitz continuous in u. For $\theta < 1$ f is θ-Hölder continuous. Local stationarity corresponds to the case $\theta \in (1/2, 1]$ [see, e.g., Dahlhaus (1996b)]. However, most of the theoretical results concerning locally stationary processes require differentiability which here corresponds to $\theta = 1$. The regularity exponent θ is the parameter that describes the null hypotheses $\mathcal{H}_0$. Our focus is on (i) discontinuities of f in u and (ii) changes in the smoothness of the trajectory $u \mapsto f(u, \omega)$ for each ω. Case (i) involves a break in the spectrum, i.e., there exits $\lambda_b^0 \in (0, 1)$ such that $\Delta f(\lambda_b^0, \omega) \triangleq (f(\lambda_b^0, \omega) - \lim_{u \downarrow \lambda_b^0} f(u, \omega)) \neq 0$ for some $\omega \in [-\pi, \pi]$. Case (ii) involves a fall in the regularity exponent from θ to θ' after λ_b^0 for some period of time for some $\omega \in [-\pi, \pi]$ (i.e., the spectrum becomes rougher after some $\lambda_b^0 \in (0, 1)$ for some time period before returning to θ-smoothness). Case (i) is the one that has received most attention so far in the time series literature although under much stronger assumptions [e.g., $f(u, \omega) = f(\omega)$]. Case (ii) is a new testing problem in spectral analysis and is of considerable interest in several fields such as finance since it requires larger sample sizes than problem (i). We show below that our tests are consistent and have minimax optimality properties for (i) and, under some conditions, (ii). The latter problem constitutes a local problem. In this paper we do not consider global problems where for example the spectrum exhibits a fall in θ from θ to $\theta' \in (0, \theta)$ on $(\lambda_b^0, 1]$. This represents a continuous change in the smoothness of the spectrum that persists permanently until the end of the time interval. Different test statistics are needed for this case.

As discussed by Last and Shumway (2008), an important question is which magnitude of the discontinuity in the time-varying spectrum can be detected. Or equivalently, how much the time-varying spectrum can change over a short time period without signifying a discontinuity. We introduce the quantity b_T, called the detection boundary or simply "rate", which is defined as the minimum break magnitude $\Delta f(\lambda_b^0, \omega)$ such that we are still able to uniformly control the type I and type II errors. To present an introduction to minimax-optimal testing [cf. Ingster (1993)], we first restrict our attention to alternative hypotheses described by a break in the spectrum [i.e., case (i) above] and defer a more general treatment to Section 5. For some fractional break point $\lambda_b^0 \in (0, 1)$ and frequency ω_0, and a decreasing sequence b_T, we consider the following alternative hypotheses:

$$\begin{aligned}\boldsymbol{F}_{1,\lambda_b^0,\omega_0}(\theta, b_T, D) = \Big\{ & \{f(u, \omega)\}_{u\in[0,1],\,\omega\in[-\pi,\pi]} \mid \\ & (f(u, \omega) - \Delta f(u, \omega))_{u\in[0,1]} \in \boldsymbol{F}(\theta, D); \\ & \left|\Delta f\left(\lambda_b^0, \omega_0\right)\right| \geq b_T,\ \omega_0 \in [-\pi, \pi]\Big\}. \end{aligned} \tag{2.5}$$

We can then present the hypotheses testing problem that we wish to address:

$$\begin{aligned} &\mathcal{H}_0 : \{f(u,\omega)\}_{u\in[0,1]} \in \boldsymbol{F}(\theta, D)\ \forall\omega\in[-\pi,\pi] \\ &\mathcal{H}_1 : \exists\lambda_b^0 \in (0,1) \text{ and } \omega_0\in[-\pi,\pi] \text{ with } \left\{\{f(u,\omega)\}_{u\in[0,1],\,\omega\in[-\pi,\pi]}\right\} \in \boldsymbol{F}_{1,\lambda_b^0,\omega_0}(\theta, b_T, D). \end{aligned} \tag{2.6}$$

Observe that $\mathcal{H}_1$ demands at least one break but it allows for multiple breaks even across different ω. For the testing problem (2.6), we establish the minimax-optimal rate of convergence [see Ch. 2 in Ingster and Suslina (2003) for an introduction]. For nonrandomized tests ψ that map a sample $\{X_t\}_{t\geq 0}$ to zero or one, we consider the maximal type I error $\alpha_\psi(\theta) = \sup_{\{f(u,\omega)\}_{u\in[0,1],\,\omega\in[-\pi,\pi]}\in\boldsymbol{F}(\theta,D)} \mathbb{P}_f(\psi=1)$, and the maximal type II error

$$\beta_\psi(\theta, b_T) = \sup_{\lambda_b^0\in(0,1),\,\omega_0\in[-\pi,\pi]}\ \sup_{\left\{\{f(u,\omega)\}_{u\in[0,1],\,\omega\in[-\pi,\pi]}\right\}\in\boldsymbol{F}_{1,\lambda_b^0,\omega_0}(\theta,b_T,D)} \mathbb{P}_f(\psi=0),$$

and define the total testing error as $\gamma_\psi(\theta, b_T) = \alpha_\psi(\theta) + \beta_\psi(\theta, b_T)$. The notion of asymptotic minimax-optimality is as follows. We want to find sequences of tests and rates b_T such that $\gamma_\psi(\theta, b_T) \to 0$ as $T\to\infty$. The larger is b_T the easier is to distinguish between $\mathcal{H}_0$ from $\mathcal{H}_1$ but we may incur at the same time into a larger type II error $\beta_\psi(\theta, b_T)$. The optimal b_T^{opt}, named the minimax distinguishable rate, is the minimum value of $b_T > 0$ such that $\lim_{T\to\infty}\inf_\psi \gamma_\psi(\theta, b_T) = 0$. A sequence of tests ψ_T that satisfies the latter relation for all $b_T \geq b_T^{\text{opt}}$ is called minimax-optimal.

Minimax-optimality has been considered in other change-point problems that are very different from our setup. Loader (1996) and Spokoiny (1998) considered the nonparametric estimation of a regression function with breaks where the break size is fixed. Bibinger, Jirak, and Vetter (2017) considered breaks in the volatility of semimartingales under high-frequency asymptotics while we focus on breaks in the spectral density and thus we work in the frequency dimension in addition to the time dimension. Another difference from previous work is that we do not deal with i.i.d. data; we cannot use the same approach to derive the minimax lower bound as in Bibinger, Jirak, and Vetter (2017) because their information-theoretic reductions exploit independence while we need to rely on approximation theorems [cf. Berkes and Philipp (1979)] to establish that our statistical experiment is asymptotically equivalent in strong Le Cam sense to a high dimensional signal detection problem. This allows us to derive the minimax bound using the classical arguments based on the results in Ingster and Suslina (2003), Ch. 8.

3 Finite-Sample and Asymptotic Results on High-Order Cumulants and Spectra of Locally Stationary Series

This section establishes asymptotic results on second and high-order cumulants and spectra for locally stationary series. These results are used to derived the limiting distributions of the test statistics for the change-point problem introduced in the prevision section. Additionally, these results are of independent interest for the literature on locally stationary and nonstationarity processes more generally. We consider the tapered finite Fourier transform, the local periodogram and the smoothed local periodogram. Let $h:\,\mathbb{R}\rightarrow\mathbb{R}$ be a data taper with $h\left(x\right)=0$ for $x\notin\left[0,\,1\right)$, $H_{k,T}\left(\omega\right)=\sum_{s=0}^{T-1}h\left(s/T\right)^{k}\exp\left(-i\omega s\right)$ and (for n_{T} even)

$$d_{h,T}\left(u,\,\omega\right)\triangleq\sum_{s=0}^{n_{T}-1}h\left(\frac{s}{n_{T}}\right)X_{\left\lfloor Tu\right\rfloor -n_{T}/2+s+1,T}\exp\left(-i\omega s\right),$$

$$I_{h,T}\left(u,\,\omega\right)\triangleq\frac{1}{2\pi H_{2,n_{T}}\left(0\right)}\left|d_{h,T}\left(u,\,\omega\right)\right|^{2},$$

where $I_{h,T}\left(u,\,\omega\right)$ is the periodogram over a segment of length n_{T} with midpoint $\left\lfloor Tu\right\rfloor$. Let $\mathbf{X}_{t,T}=\left(X_{t,T}^{(a_{1})},\ldots,\,X_{t,T}^{(a_{p})}\right)$ with finite $p\geq1$. Denote by $\kappa_{\mathbf{X},t}^{(a_{1},\ldots,a_{r})}\left(k_{1},\ldots,\,k_{r-1}\right)$ the time-t r-order cumulant of $(X_{t+k_{1}}^{(a_{1})},\ldots,X_{t+k_{r-1}}^{(a_{r-1})},\,X_{t}^{(a_{r})})$ with $r\leq p$.

Assumption 3.1. *(i) $\{\mathbf{X}_{t,T}\}$ is a mean-zero locally stationary process (i.e., $m_{0}=0$); (ii) for all $1\leq j\leq p$ $A^{(a_{j})}\left(u,\,\omega\right)$ is 2π-periodic in ω and the periodic extensions are differentiable in u and ω with uniformly bounded derivative $\left(\partial/\partial u\right)\left(\partial/\partial\omega\right)A\left(u,\,\omega\right)$; (iii) g_{4} is continuous.*

Assumption 3.2. *There is an $l\geq0$ such that*

$$\sum_{k_{1},\ldots,\,k_{r-1}=-\infty}^{\infty}\left(1+\left|k_{r}\right|^{l}\right)\sup_{1\leq t\leq T}\left|\kappa_{\mathbf{X},t}^{(a_{1},\ldots,a_{r})}\left(k_{1},\ldots,\,k_{r-1}\right)\right|<\infty,\qquad(3.1)$$

for $j=1,\ldots,\,r-1$ and any r tuple $a_{1},\ldots,a_{r}$ when $r=2,\,3,\ldots$

Assumption 3.3. *The data taper $h:\,\mathbb{R}\rightarrow\mathbb{R}$ with $h\left(x\right)=0$ for $x\notin\left[0,\,1\right]$ is bounded and of bounded variation.*

Assumption 3.4. *The sequence $\{n_{T}\}$ satisfies $n_{T}\rightarrow\infty$ as $T\rightarrow\infty$ with $n_{T}/T\rightarrow0$.*

3.1 Local Finite Fourier Transform

The above theorem presents the asymptotic expression for the joint cumulants of the finite Fourier transform $d_{T}^{(\cdot)}\left(u,\,\cdot\right)$. Next, we use this result to obtain the limit distribution of the transform.

Let $\mathbf{d}_{h,T}(u,\,\omega) = [d_{h,T}^{(a_j)}(u,\,\omega)]$, $j = 1 \leq j \leq r$. Corresponding results for stationary series are reviewed in e.g. Brillinger (1975). Let $H_{n_T}^{(a_1,\ldots,a_r)}(\omega) = \sum_{s=0}^{n_T-1}\left(\prod_{j=1}^{r} h_{a_j}(s/n_T)\right)\exp(-i\omega s)$ and $H^{(a_1,\ldots,a_r)}(\omega) = \int\left(\prod_{j=1}^{r} h_{a_j}(t)\right)\exp(-i\omega t)\,dt$.

Theorem 3.1. *Suppose that Assumption 3.1-3.2 with $l = 0$ hold. Assume that $h^{(a_j)}(x)$ satisfies Assumption 3.3 for all $j = 1, \ldots, p$. Under Assumption 3.4,*

$$\begin{aligned}&\mathrm{cum}\left(d_{h,T}^{(a_1)}(u,\,\omega_1),\ldots,\,d_{h,T}^{(a_r)}(u,\,\omega_r)\right)\\&\quad= (2\pi)^{r-1} H_{n_T}^{(a_1,\ldots,a_r)}\left(\sum_{j=1}^{r}\omega_j\right) f(u,\,\omega_1,\ldots,\,\omega_{r-1}) + \varepsilon_T,\end{aligned}$$

where $\varepsilon_T = o(n_T)$ is uniform in ω_j, $(j = 1, \ldots, r)$. If Assumption 3.2 holds with $l = 1$, then $\varepsilon_T = O(n_T/T)$ uniformly in ω_j, $(j = 1, \ldots, r)$.

Theorem 3.2. *Suppose that Assumption 3.1-3.2 with $l = 0$ hold. Assume that $h^{(a_j)}(x)$ satisfies Assumption 3.3 for all $j = 1, \ldots, p$. Suppose $2\omega_j,\, \omega_j \pm \omega_k \not\equiv 0\,(\mathrm{mod}\,2\pi)$ for $1 \leq j < k \leq J_\omega$. Under Assumption 3.4, (i) $\mathbf{d}_{h,T}(u,\,\omega_j)$, $j = 1, \ldots, J_\omega$ are approximately independent $\mathscr{N}_p^{\mathrm{C}}\left(0, 2\pi n_T\left[H^{(a_l,a_r)}(0)\, f^{(a_l,a_r)}(u,\,\omega_j)\right]\right)$ $(l,\,r = 1, \ldots, p)$ variables; (ii) If $\omega = 0, \pm\pi, \pm 2\pi, \pm 3\pi$, $\ldots$, $\mathbf{d}_{h,T}(u,\,\omega)$ is asymptotically $\mathscr{N}_p\left(0, 2\pi n_T\left[H^{(a_l,a_r)}(0)\, f^{(a_l,a_r)}(u,\,\omega)\right]\right)$ $(l,\,r = 1, \ldots, p)$ independently from the previous variates.*

3.2 Local Periodogram

We now study several properties of the tapered local periodogram. We begin with the finite-sample bias and variance. We then present results on the asymptotic distribution which allow us to conclude that the local periodogram evaluated at distinct ordinates results in estimates that are asymptotically independent thereby mirroring the stationary case. These results are useful in order to develop hypothesis testing on the time-varying properties of the spectrum of a time series.

Theorem 3.3. *Suppose that Assumption 3.1 holds and that*

$$\sum_{k=-\infty}^{\infty}\sup_{u\in[0,\,1]}|c(u,\,k)| < \infty. \tag{3.2}$$

Assume that $h(x)$ satisfies Assumption 3.3. Under Assumption 3.4, we have for $-\infty < \omega < \infty$,

$$\mathbb{E}(I_{h,T}(u,\,\omega)) = \left(\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2\,d\alpha\right)^{-1}\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2 f(u,\,\omega-\alpha)\,d\alpha + O\left(\frac{\log(n_T)}{n_T}\right) \tag{3.3}$$

$$= f(u,\,\omega) + \frac{1}{2}\left(\frac{n_T}{T}\right)^2 \left(\int_0^1 h^2(x)\,dx\right)^{-1} \int_0^1 x^2 h^2(x+1/2)\,dx \frac{\partial^2}{\partial u} f(u,\,\omega) + o\left(\left(\frac{n_T}{T}\right)^2\right) + O\left(\frac{\log(n_T)}{n_T}\right).$$

The first equality shows that the expected value of $I_{h,T}(u,\,\omega)$ is a weighted average of the local spectral density at rescaled time u with weights concentrated in the neighborhood of ω and relative weights determined by the taper. The second equality shows that $I_{h,T}(u,\,\omega)$ is asymptotically unbiased for $f(u,\,\omega)$ and provides a bound on the asymptotic bias.

Theorem 3.4. *Suppose that Assumption 3.1 holds and that $h(\cdot)$ satisfies Assumption 3.3. Under Assumption 3.4 we have (i) for $-\infty < \omega_j,\,\omega_k < \infty$,*

$$\begin{aligned}&\mathrm{Cov}\left\{I_{h,T}(u,\,\omega_j),\, I_{h,T}(u,\,\omega_k)\right\} \\ &\quad= |H_{2,n_T}(0)|^{-2}\left(|H_{2,n_T}(\omega_j-\omega_k)|^2 + |H_{2,n_T}(\omega_j+\omega_k)|^2\right) f(u,\,\omega_j)^2 + O\left(n_T^{-1}\right),\end{aligned} \tag{3.4}$$

where the error term is uniform in ω_j and ω_k; (ii) for $2\omega_j,\,\omega_j \pm \omega_k \not\equiv 0\,(\mathrm{mod}\,2\pi)$ with $1 \leq j < k \leq J_\omega$, the variables $I_{h,T}(u,\,\omega_j)$, $j = 1, \ldots, J_\omega$ are asymptotically independent $f(u,\,\omega_j)\,\chi_2^2/2$ variates. Also, if $\omega = \pm\pi,\,\pm 3\pi, \ldots,$ $I_{h,T}(u,\,\omega)$ is asymptotically $f(u,\,\omega)\,\chi_1^2$ independent of the previous variates.

3.3 Smoothed Local Periodogram

The smoothed local periodogram is defined as

$$f_{h,T}(u,\,\omega) = \frac{2\pi}{n_T}\sum_{s=1}^{n_T-1} W_T\left(\omega - \frac{2\pi s}{n_T}\right) I_{h,T}\left(u,\,\frac{2\pi s}{n_T}\right) \qquad -\infty < \omega < \infty,$$

where $W_T(\omega)$, $-\infty < \omega < \infty$, is a family of weight functions of period 2π,

$$W_T(\omega) = \sum_{j=-\infty}^{\infty} b_{W,T}^{-1} W\left(b_{W,T}^{-1}(\omega + 2\pi j)\right) \qquad -\infty < \omega < \infty$$

where $b_{W,T}$ is a bandwidth and $W(\beta)$, $-\infty < \beta < \infty$, is a fixed function satisfying the following assumption.

Assumption 3.5. *$W(\beta)$, $-\infty < \beta < \infty$, is real-valued, even, of bounded variation, and satisfies $\int_{-\infty}^{\infty} W(\beta)\,d\beta = 1$ and $\int_{-\infty}^{\infty} |W(\beta)|\,d\beta < \infty$.*

Theorem 3.5. *Suppose that Assumption 3.1 holds, that $h(\cdot)$ satisfies Assumption 3.3, Assumption 3.4 holds with $l=1$, and $W(\cdot)$ satisfies Assumption 3.5. Then,*

$$\begin{aligned}\mathbb{E}\left(f_{h,T}(u,\,\omega)\right) &= \int_{-\infty}^{\infty} W(\beta)\, f(u,\,\omega - b_{W,T}\beta)\, d\beta + O\left((n_T b_{W,T})^{-1}\right) + O\left(\log(n_T)\, n_T^{-1}\right) \qquad (3.5)\\ &= f(u,\,\omega) + \frac{1}{2}\left(\frac{n_T}{T}\right)^2 \left(\int_0^1 h^2(x)\,dx\right)^{-1} \int_0^1 x^2 h^2(x+1/2)\,dx \frac{\partial^2}{\partial u} f(u,\,\omega)\\ &\quad + \frac{1}{2} b_{W,T}^2 \int_0^1 x^2 W(x)\,dx \frac{\partial^2}{\partial \omega} f(u,\,\omega) + O\left(\frac{\log(n_T)}{n_T}\right) + o\left(\left(\frac{n_T}{T}\right)^2\right) + o\left(b_{W,T}^2\right).\end{aligned}$$

The error terms are uniform in ω.

Theorem 3.6. *Suppose that Assumption 3.1 holds, that $h(\cdot)$ satisfies Assumption 3.3, Assumption 3.4 holds with $l=1$, and $W(\cdot)$ satisfies Assumption 3.5. Let $b_{W,T} \to 0$, $b_{W,T} n_T \to \infty$, as $T \to \infty$. Then $f_{h,T}(u,\,\omega_1), \ldots, f_{h,T}(u,\,\omega_{J_\omega})$ are asymptotically jointly normal with $\lim_{T\to\infty} \mathbb{E}\left(f_{h,T}(u,\,\omega)\right) = f(u,\,\omega)$ and*

$$\begin{aligned}&\lim_{T\to\infty} n_T b_{W,T} \mathrm{Cov}\left(f_{h,T}(u,\,\omega_j),\, f_{h,T}(u,\,\omega_k)\right) \qquad (3.6)\\ &\qquad = 2\pi\left[\eta\{\omega_j - \omega_k\} + \eta\{\omega_j + \omega_k\}\right] \int h(t)^4\,dt \left[\int h(t)^2\,dt\right]^{-2} \int W(\alpha)^2\,d\alpha\, f(u,\,\omega_j)^2.\end{aligned}$$

Consistency of the spectral density estimates of a stationary time series were obtained by Grenander and Rosenblatt (1957) and Parzen (1957). Asymptotic normality was considered by Rosenblatt (1959), Brillinger and Rosenblatt (1967), Hannan (1970) and Anderson (1971). Theorem 3.6 presented corresponding results for the locally stationary case which nightlight the nature of the smoothing over time in addition to over the frequency domain. Panaretos and Tavakoli (2013) established similar results for functional stationary processes while Aue and van Delft (2020) established some results for functional locally stationary processes using a different notion of local stationarity.

4 Change-Point Tests for Discontinuity in the Spectrum

We construct tests for the hypothesis testing problem in (2.6). The test statistics involve the local periodogram. It follows from the results of Section 3 that under local stationarity $I_{h,T}(t/T,\,\omega)$ is asymptotically unbiased for $f(t/T,\,\omega)$. Let $\widetilde{I}_{r,T}(\omega) = M_{S,T}^{-1} \sum_{j\in\mathbf{S}_r} I_{h,T}(j/T,\,\omega)$ where $\mathbf{S}_r = \{rm_T - m_T/2+1,\, rm_T - m_T/2+1+m_{S,T}, \ldots,\, rm_T + M_{S,T} m_{S,T}/2\}$, $m_{S,T} = \left\lfloor m_T^{1/2}\right\rfloor$ and $M_{S,T} = \lfloor m_T/m_{S,T}\rfloor$. $\widetilde{I}_{r,T}(\omega)$ denotes the average local periodogram around time rm_T where $r = 1, \ldots, M_T = \lfloor T/m_T\rfloor$

where we do not use all the m_T local periodograms $I_{h,T}(j/T,\,\omega)$ in the block r but only those separated by $m_{S,T}$ points. Thus, $\mathbf{S}_r$ is the set of indices of a sub-sample of the local periodograms in the r block. We need to consider the sub-sample because otherwise there is too much dependence among the adjacent terms e.g., $I_{h,T}(j/T,\,\omega)$ and $I_{h,T}((j+1)/T,\,\omega)$. $\widetilde{I}_{r,T}(\omega)$ is an asymptotically unbiased estimate for $f(rm_T/T,\,\omega)$. Thus, a large distance between $\widetilde{I}_{r,T}(\omega)$ and $\widetilde{I}_{r+1,T}(\omega)$ suggests the presence of a jump or unsmooth break in the spectrum close to time $(r+1)\,m_T$ and at frequency ω. We first present a test statistic for the detection of a change-point in the spectrum $f(\cdot,\,\omega)$ for a given frequency ω. A second test statistic that we consider is one that can detect change-points in $u\in[0,\,1]$ occurring at any frequency $\omega\in[-\pi,\,\pi]$. The latter is arguably more useful in practice because often the practitioner does not know a priori whether the spectrum is discontinuous at a low or high frequency.

We consider the following test statistic,

$$\mathrm{I}_{\max,T}(\omega)\triangleq\max_{r=1,\ldots,\,M_T-2}\left|\frac{\widetilde{I}_{r+1,T}(\omega)-\widetilde{I}_{r,T}(\omega)}{\sigma_r(\omega)}\right|,\qquad\omega\in[-\pi,\,\pi]\,,\tag{4.1}$$

where $\sigma_r^2(\omega)\triangleq\mathrm{Var}(\sqrt{M_{S,T}}\widetilde{I}^*_{r,T}(\omega))$ and $\widetilde{I}^*_{r,T}(\omega)$ is equal to $\widetilde{I}_{r,T}(\omega)$ with $I_{h,T}(j/T,\,\omega)$ replaced by $I^*_T(j/T,\,\omega)=I_{h,T}(j/T,\,\omega)-\mathbb{E}(I_{h,T}(j/T,\,\omega))$. Test statistics of the form of (4.1) were also used in the context of nonparametric change-point analysis [cf. Wu and Zhao (2007) and Bibinger, Jirak, and Vetter (2017)] and forecasting [cf. Casini (2018)]. For the test statistic (4.1), the derivation of the null distribution uses the (strong) invariance principle for nonstationary processes [see e.g., Wu (2007a) and Wu and Zhou (2011)]. We need to impose some conditions on the serial dependence.

Let $\{e_t\}$, $t\in\mathbb{Z}$, be a sequence of i.i.d. random variables. Assume $I^*_{h,T}(j/T,\,\omega)=H_h(j/T,\,\mathscr{F}_{j+n_T/2})$, where $\mathscr{F}_t\triangleq\{\ldots,\,e_{t-1},\,e_t\}$ and $H_h:\,[0,\,1]\times\mathbb{R}^\infty\mapsto\mathbb{R}$ is a measurable function. We use the physical dependence measure introduced by Wu (2005, 2007) for stationary processes and extended to nonstationary processes by Wu and Zhou (2011). Let $\{e'_t\}_{t\in\mathbb{Z}}$ be an independent copy of $\{e'_t\}_{t\in\mathbb{Z}}$. Let $\mathscr{L}^q$ denote the space generated by the q-norm, $q>0$. For all j assume $I^*_{h,T}(j/T,\,\omega)\in\mathscr{L}^q$. For $w\geq 0$ define the physical dependence measure,

$$\begin{aligned}\phi_{w,q}&=\sup_{j\in\{S_r;\,r=1,\ldots,\,M_T-2\}}\left\|I^*_{h,T}(j/T,\,\omega)-I^*_{h,T,\{w\}}(j/T,\,\omega)\right\|_q\\&=\sup_{j\in\{S_r;\,r=1,\ldots,\,M_T-2\}}\left\|H_h\left(j/T,\,\mathscr{F}_{j+n_T/2}\right)-H_h\left(j/T,\,\mathscr{F}_{j+n_T/2,\{w\}}\right)\right\|_q,\end{aligned}\tag{4.2}$$

where $\mathscr{F}_{j+n_T/2,\{w\}}$ is a coupled version of $\mathscr{F}_{j+n_T/2}$ with $e_{j+n_T/2-w}$ replaced by an i.i.d. copy $e'_{j+n_T/2-w}$. Assume $\Upsilon_{n,q}=\sum_{j=n}^{\infty}\phi_{j,q}<\infty$. Let $\tau_T=T^{\vartheta_1}(\log(T))^{\vartheta_2}$ where $\vartheta_1=(1/2-1/q+\gamma/q)\,/\,(1/2-1/q+\gamma)$ and $\vartheta_2=(\gamma+\gamma/q)\,/\,(1/2-1/q+\gamma)$ for some $\gamma>0$.

Assumption 4.1. *Let $2<q\leq 4$ and assume either (i) $\Upsilon_{n,q}=O\left(n^{-\gamma}\right)$ or (ii) $\phi_{w,q}=O\left(\rho^{w}\right)$ for some $\rho\in(0,\,1)$.*

Condition 1. (i) The sequence $\{m_T\}$ satisfies $m_T\rightarrow\infty$ as $T\rightarrow\infty$, and

$$M_{S,T}^{1/2}m_T^{\theta}T^{-\theta}\left(\log\left(M_T\right)\right)^{1/2}+\tau_T^2\log\left(M_T\right)M_{S,T}^{-1}+M_{S,T}n_T^4\log\left(M_T\right)T^{-4}+M_{S,T}\left(\log\left(n_T\right)\right)^2\log\left(M_T\right)n_T^{-2}$$
$$+M_{S,T}\left(\log\left(n_T\right)\right)^2\log\left(M_T\right)n_T^{-2}\rightarrow 0; \tag{4.3}$$

(ii) For $\{b_{W,T}\}$ defined in Section 3 assume that $\log\left(M_T\right)m_Tb_{W,T}^4\rightarrow 0$.

Part (i) imposes a lower bound and an upper bound on the growth condition of the sequence $\{m_T\}$. The upper bound relates to the smoothness of $A\left(u,\,\omega\right)$ under the null hypotheses and to n_T. Part (ii) is used for the additional test statistics proposed below. Let $\gamma_{M_T}=\left[4\log\left(M_T\right)-2\log\left(\log\left(M_T\right)\right)\right]^{1/2}$ and $\mathscr{V}$ denote a random variable defined by $\mathbb{P}\left(\mathscr{V}\leq v\right)=\exp(-\pi^{-1/2}\exp\left(-v\right))$.

Theorem 4.1. *Assume Assumption 3.1, 3.3-3.4, 4.1, eq. (3.2) and Condition 1-(i) hold. Under $\mathcal{H}_0$, we have $\sqrt{\log\left(M_T\right)}\left(M_{S,T}^{1/2}\mathrm{I}_{\max,T}\left(\omega\right)-\gamma_{M_T}\right)\Rightarrow\mathscr{V}$ for any $\omega\in[-\pi,\,\pi]$.*

Theorem 4.1 shows that the asymptotic null distribution follows an extreme vale distribution. We discuss the consistent estimation of $\sigma_r^2\left(\omega\right)$ below.

The test statistic $\mathrm{I}_{\max,T}\left(\omega\right)$ detects breaks in the spectrum for a given frequency ω. A first alternative would be to consider a double-sup statistic which takes the maximum also over $\omega\in[-\pi,\,\pi]$. This changes only slightly the distribution because the $\mathrm{I}_{\max,T}\left(\omega\right)$ are independent across ω. This follows from Theorem 3.4 which shows that $I_{h,T}\left(u,\,\omega_j\right)$ and $I_{h,T}\left(u,\,\omega_k\right)$ are asymptotically independent if $2\omega_j,\,\omega_k\pm\omega_k\not\equiv 0\,(\mathrm{mod}\,2\pi)$, and the maximum of independent extreme value distributed random variables still follows an extreme value distribution. We need to introduce the notation for such double-sup setting. We specify a framework based on an infill procedure over the frequency domain $[-\pi,\,\pi]$ by assuming that there are n_ω frequencies $\omega_1,\ldots,\,\omega_{n_\omega}$, with $\omega_1=-\pi$ and $\omega_{n_\omega}=\pi-\epsilon$, $\epsilon>0$, and $|\omega_j-\omega_{j+1}|=O\left(n_\omega^{-1}\right)$ for $j=1,\ldots,\,n_\omega-2$. Assume that $n_\omega\rightarrow\infty$ as $T\rightarrow\infty$. Let $\Pi\triangleq\{\omega_1,\,\omega_2,\ldots,,\,\omega_{n_\omega-1},\,\omega_{n_\omega}\}$.

Assumption 4.2. *Assume that $2\omega_j,\,\omega_j\pm\omega_k\not\equiv 0\,(\mathrm{mod}\,2\pi)$ for $\omega_j,\,\omega_k\in\Pi$.*

Theorem 4.2. *Suppose that Assumption 3.1, 3.3-3.4, 4.1-4.2, eq. (3.2) and Condition 1-(i) hold. Under $\mathcal{H}_0$, we have $\mathrm{I}_{\mathrm{Dmax},T}=\max_{\omega\in\Pi}\sqrt{\log\left(M_T\right)}\left(M_{S,T}^{1/2}\mathrm{I}_{\max,T}\left(\omega\right)-\gamma_{M_T}\right)-\log\left(n_\omega\right)\Rightarrow\mathscr{V}$.*

The double-sup specification results in the extra factor $\log(n_\omega)$ affecting the null distribution of the test. The last test statistic that we consider is based on the smoothed local periodogram $f_{h,T}(u, \omega)$. Under $\mathcal{H}_0$ the unsmoothed version is only asymptotically unbiased for $f(u, \omega)$ while $f_{h,T}(u, \omega)$ is consistent for $f(u, \omega)$. We define $\widetilde{f}_{r,T}(\omega) = M_{S,T}^{-1} \sum_{j \in \mathbf{S}_r} f_{h,T}(j/T, \omega)$ and

$$\mathrm{S}_{\max,T}(\omega) \triangleq \max_{r=1,\ldots,M_T-2} \left| \frac{\widetilde{f}_{r,T}(\omega) - \widetilde{f}_{r+1,T}(\omega)}{\sigma_{f,r}(\omega)} \right|, \tag{4.4}$$

where $\omega \in [-\pi, \pi]$, $\sigma_{f,r}^2(\omega) \triangleq \mathrm{Var}(\sqrt{M_{S,T}} \widetilde{f}_{r,T}(\omega))$, and $f_{h,T}^*(j/T, \omega) = f_{h,T}(j/T, \omega) - \mathbb{E}(f_{h,T}(j/T, \omega))$. Smoothing over frequencies introduces short-range dependence over ω. The maximum is taken over the following set of frequencies $\Pi' \triangleq \{\omega_1, \omega_{2+\lfloor n_T b_T \rfloor}, \ldots, \omega_{n_\omega - \lfloor n_T b_T \rfloor - 1}, \omega_{n_\omega}\}$. Let $n'_\omega = \lfloor n_\omega / (\lfloor n_T b_T \rfloor + 1) \rfloor$. Note that $\Pi' \subset \Pi$. The reason is that, due to the short-range dependence introduced by the smoothed periodogram, we cannot consider the maximum over all frequencies in Π because the statistics would not be independent. Let $\mathrm{S}_{\mathrm{Dmax},T} \triangleq \max_{\omega_k \in \Pi'} \sqrt{\log(M_T)} (M_{S,T}^{1/2} \mathrm{S}_{\max,T}(\omega_k) - \gamma_{M_T}) - \log(n'_\omega)$.

Theorem 4.3. *Assumption 3.1, 3.3-3.4, 4.1-4.2, eq. (3.2) and Condition 1 hold. Under* $\mathcal{H}_0$ *we have* $\sqrt{\log(M_T)} (M_{S,T}^{1/2} \mathrm{S}_{\max,T}(\omega_k) - \gamma_{M_T} \Rightarrow \mathscr{V}$ *and* $\mathrm{S}_{\mathrm{Dmax},T} \Rightarrow \mathscr{V}$.

5 Consistency and Minimax Optimal Rate of Convergence

In this section we discuss the consistency and minimax-optimal lower bound under a more general setting than the one introduced in Section 2. We consider alternative hypotheses where f is less smooth than under $\mathcal{H}_0$, thereby including the case of breaks as a special case. Suppose that under $\mathcal{H}_0$ the spectrum of X_t behaves until time $T\lambda_b^0$ as a spectrum in $\boldsymbol{F}(\theta, D)$ where $\theta > 0$ and $D < \infty$. After $T\lambda_b^0$, the regularity exponent θ drops to some θ' with $0 < \theta' < \theta$ for some non-trivial period of time. That is, since $\boldsymbol{F}(\theta, D) \subset \boldsymbol{F}(\theta', D)$, we need that f behaves as θ'-regular for some period of time such that there exists a ω with $\{f(u, \omega)\}_{u \in [0,1]} \notin \boldsymbol{F}(\theta, D)$. This guarantees that $\mathcal{H}_0$ and $\mathcal{H}'_1$ (to be defined below) are well-separated. To this end, define for some function g_u with $u \in [0, 1]$, $\Delta_h^{\theta'} g_u = (g_{u+h} - g_u) / |h|^{\theta'}$, $h \in [-u, 1-u]$. The set of possible alternatives is then defined as

$$\boldsymbol{F}'_{1,\lambda_b^0,\omega_0}(\theta, \theta', b_T, D) = \{\{f(u, \omega)\}_{u \in [0,1], \omega \in [-\pi, \pi]} \in \boldsymbol{F}(\theta', D) \mid \inf_{|h| \leq 2m_T/T} \Delta_h^{\theta'} f\left(\lambda_b^0, \omega_0\right) \geq b_T \quad \text{or} \quad \Delta_h^{\theta'} f\left(\lambda_b^0, \omega_0\right) \leq -b_T\}.$$

Note that $\boldsymbol{F}'_{1,\lambda_b^0,\omega_0}$ depends on m_T but since m_T depends on θ we can omit m_T from the argument of $\boldsymbol{F}'_{1,\lambda_b^0,\omega_0}$. This leads to the following testing problem

$$\begin{aligned} &\mathcal{H}_0 : \{f(u,\omega)\}_{u\in[0,1]} \in \boldsymbol{F}(\theta, D) \\ &\mathcal{H}'_1 : \exists\lambda_b^0 \in (0,1) \text{ and } \omega_0 \in [-\pi,\pi] \text{ with } \left\{\{f(u,\omega)\}_{u\in[0,1],\,\omega\in[-\pi,\pi]}\right\} \in \boldsymbol{F}'_{1,\lambda_b^0,\omega_0}(\theta, \theta', b_T, D). \end{aligned} \tag{5.1}$$

In the context of infinite-dimensional parameter problems one faces the issue of distinguishability between the null and the alternative hypotheses. It is evident that one cannot test $f \in \boldsymbol{F}(\theta, D)$ versus $f \in \boldsymbol{F}(\theta', D)$ for $\theta > \theta'$. First, since $\boldsymbol{F}(\theta, D) \subset \boldsymbol{F}(\theta', D)$, one has to at least remove the set of functions in $\boldsymbol{F}(\theta, D)$ from those in $\boldsymbol{F}(\theta', D)$. Yet, as discussed by Ingster and Suslina (2003), this would not be enough. The reason is that the two hypotheses are still too close. That explains why we focus on spectral densities f that belong to $\boldsymbol{F}'_{1,\lambda_b^0,\omega_0}(\theta, \theta', b_T, D)$ under $\mathcal{H}_1$. Those spectral densities are rough enough so as not to be close to functions in $\boldsymbol{F}(\theta, D)$. This is captured by the requirement that the different quotient $\Delta_h^{\theta'} f$ exceeds the so-called rate b_T. As $T \to \infty$ the requirement becomes less stringent since $b_T \to 0$. See Hoffmann and Nickl (2011) and Bibinger, Jirak, and Vetter (2017) for similar discussions in different contexts.

We now move to the derivation of the minimax lower bound. For a purely technical reason inherent to the proof we need to restrict attention to a strictly positive spectral density in the frequency dimension at which the null hypotheses is violated. That is, let $f_-(\omega_0) = \inf_{u\in[0,1]} f(u,\omega_0) > 0$. Such restriction is not imposed on $f(u,\omega)$ for $\omega \neq \omega_0$.

Theorem 5.1. *Assume Assumption 2.1-2.2, 3.3-3.4, eq. (3.2), $0 < \theta' < \theta$ and $f_-(\omega_0) > 0$. Consider either set of hypotheses $\{\mathcal{H}_0, \mathcal{H}_1\}$ or $\{\mathcal{H}_0, \mathcal{H}'_1\}$. Then, for*

$$b_T \leq (T/\log(M_T))^{-\frac{\theta-\theta'}{2\theta+1}} D^{-\frac{2\theta'+1}{2\theta+1}} f_-(\omega_0),$$

where $\theta' = 0$ for $\mathcal{H}_1$, we have $\lim_{T\to\infty} \inf_\psi \gamma_\psi(\theta, b_T) = 1$.

The theorem implies that b_T^{opt} is such that $b_T^{\text{opt}} > (T/\log(M_T))^{-\frac{\theta-\theta'}{2\theta+1}} D^{-\frac{2\theta'+1}{2\theta+1}} f_-(\omega_0)$. Note that the lower bound does not depend on ω. Thus, we can derive tests based on b_T^{opt}. For example, for the test statistic (4.1) we obtain the following test $\psi^* : \psi^*(\{X_t\}_{1\leq t\leq T}) = 1$ if $\mathrm{I}_{\max,T}(\omega) \geq 2D^*\sqrt{\log(M_T^*)/m_T^*}$ for $\omega \in [-\pi,\pi]$, where $D^* > 2$, $m_T^* = (\sqrt{\log(M_T^*)} T^\theta/D)^{\frac{2}{2\theta+1}}$ and $M_T^* = \lfloor T/m_T^* \rfloor$.

Next, we establish the optimal rate for minimax distinguishability. Note that either alternatives $\mathcal{H}_1$ or $\mathcal{H}'_1$ allows for multiple breaks which may occur close to each other. For technical reasons one has to either assume that the breaks do not cancel each other or assume that they cannot be too close. Here we assume the latter thereby following the definition of segmented locally stationary.

Theorem 5.2. *Assume Assumption 2.1-2.2, 3.3-3.4 and eq.* (3.2) *hold. Consider either alternative hypothesis* $\mathcal{H}_1$ *or* $\mathcal{H}_1'$, *with* $\lambda_j^0 < \lambda_{j+1}^0$ *for* $j = 1, \ldots, m_0 - 1$. *If* $(n_T/T)^2 + \log(n_T)/n_T \leq 2D^* \sqrt{\log(M_T^*)/m_T^*}$ *and*

$$b_T^* > \left(4D^* \sup_{u \in [0,\,1]} f(u,\, \omega_0) + 2\right)^{-\frac{\theta - \theta'}{2\theta + 1}} (T/\log(M_T^*))^{-\frac{\theta - \theta'}{2\theta + 1}} D^{-\frac{2\theta' + 1}{2\theta + 1}}, \tag{5.2}$$

then $\lim_{T \to \infty} \gamma_{\psi^*}(\theta,\, b_T^*) = 0$ *and* $b_T^{\mathrm{opt}} \propto (T/\log(M_T))^{-\frac{\theta - \theta'}{2\theta + 1}}$.

The theorem shows that a smooth change in the regularity exponent θ cannot be distinguish from a break of magnitude smaller than b_T^{opt} or in other words $\boldsymbol{F}_{1, \lambda_b^0, \omega_0} \nsubseteq \boldsymbol{F}'_{1, \lambda_b^0, \omega_0}$ because the change in θ to θ' has to persist for some time. This is also indicated by the restriction $\theta' > 0$. The minimax bound is similar to the one established by Bibinger, Jirak, and Vetter (2017) for the volatility of a Itô semimartingale. The theorem suggests that knowledge of the frequency ω_0 for which the spectrum changes regularity is irrelevant for the determination of the bound. However, we conjecture that if the spectrum exhibits a break or smooth change of the form discussed above simultaneously across multiple frequencies then the lower bound may be further decreased as one can pool additional information from inspection of the spectrum for the set of frequencies subject to the change. The key assumption would be that the change occurs at the same time λ_b^0 for a given set of frequencies ω. This may be of some interest in economics and finance since it has been documented for some time series that a break in the mean or trend (or other low frequency components) can be associated to a break in the volatility at the same date. Thus, $f(u,\, \omega)$ may as well exhibit a break simultaneously at high and low frequencies. We leave this to future research. The theorem also requires further restrictions on the relation between n_T and m_T.

6 Estimation of the Change-Points

We now discuss estimation of the break locations. We first discuss the case of a single break and then present the results for the case of multiple breaks.

6.1 Single Break Alternatives

Let $I_T(j/T,\, \omega) = (2\pi n_T)^{-1} |d_T(j/T,\, \omega)|^2$ with $d_T(j/T,\, \omega) = \sum_{s=0}^{n_T - 1} X_{j - n_T + s + 1, T} \exp(-i\omega s)$. The estimation of the change-point is based on the following statistic:

$$\mathrm{D}_{r,T}(\omega) \triangleq m_T^{-1/2} \left| \sum_{j=r-m_T+1}^{r} I_T(j/T,\,\omega) - \sum_{j=r+1}^{r+m_T} I_T(j/T,\,\omega) \right|, \qquad \omega \in [-\pi,\,\pi].$$

where $r = m_T + n_T - 1, \ldots, T - m_T$. We use the untapered periodogram, $I_T(\cdot,\,\cdot)$, instead of tapered one, $I_{h,T}(\cdot,\,\cdot)$, to simplify the notation in the proofs. Note that the maximum of the statistics $\mathrm{D}_{r,T}(\omega)$'s is a version of $\mathrm{I}_{\max,T}$ that uses overlapping blocks and that does not involve the normalization $\sigma_r(\omega)$. The change-point estimator is defined as $T\widehat{\lambda}_{b,T} = \arg\max_{r=m_T+n_T-1,\ldots,T-m_T} \mathrm{D}_{r,T}(\omega)$. We consider the following alternative hypotheses:

$$\mathcal{H}_{1,\mathrm{S}}: \quad \left\{ f\left(T_b^0/T,\,\omega_0\right) - f\left(T_{b,+}^0/T,\,\omega_0\right) = \delta_T \neq 0, \quad \omega_0 \in [-\pi,\,\pi] \right\},$$

where $T_{b,+}^0 = \lim_{s \downarrow T_b^0,\, s > T_b^0} s$. The break magnitude can be either fixed or converge to zero as specified by the following assumption.

Assumption 6.1. *$\delta_T \to 0$ and $\delta_T m_T^{1/2}/\sqrt{\log(T)} \to (0,\,\infty]$.*

Proposition 6.1. *Assume Assumption 3.1, 3.3-3.4, 4.1, eq. (3.2) and Condition 1-(i) hold. Under $\mathcal{H}_{1,\mathrm{S}}$, if δ_T is fixed or satisfies Assumption 6.1, we have (i) $\widehat{\lambda}_{b,T} - \lambda_b^0 = O_{\mathbb{P}}(\sqrt{m_T \log(T)}/(T\delta_T))$.*

It is useful to compare the rate of convergence in Proposition 6.1 with that of classical change-point estimators of a break in a constant mean. For fixed shifts, the latter rate of convergence is $O_{\mathbb{P}}(T^{-1})$ while for shrinking shifts is $O_{\mathbb{P}}((T\delta_T^2)^{-1})$ where $\delta_T \to 0$ with $\delta_T T^{1/2-\vartheta}$ for some $\vartheta \in (0,\,1/2)$ [cf. Yao (1987)]. Unlike the classical change-point problem where the mean is constant except for the break, our problem involves a time-varying target also under the null. For fixed shifts the rate of convergence in our problem is slower. The smallest break magnitude allowed by Proposition 6.1 is $\delta_T = O(\sqrt{\log(T)}/m_T^{1/2})$. Under this condition the convergence rate for the classical change-point estimator is $O_{\mathbb{P}}(m_T(T\log(T))^{-1})$ which is the same as the one suggested by Proposition 6.1, saved for a logarithmic term. However, in classical change-point setting $\delta_T \to 0$ is allowed at a faster rate. This is obvious since in our setting a small break can be confounded with a smooth local change in the target.

Under the alternative $\mathcal{H}_1'$ the estimator is consistent only when θ-regularity is violated in a small interval around λ_b^0. If the length of this interval exceeds $O(\sqrt{m_T \log(T)}/T\delta_T)$, then consistency does not hold because this becomes a global problem which cannot be addressed by the estimation method considered in this section.

6.2 Multiple Breaks Alternatives

Let us assume that there are m_0 break points in $f(u, \omega_0)$. Let $0 < \lambda_1^0 < \ldots < \lambda_{m_0}^0 < 1$. We consider the following alternative hypotheses:

$$\mathcal{H}_{1,\mathrm{M}}: \quad \left\{ f\left(T_l^0/T, \omega_l\right) - f\left(T_{l,+}^0/T, \omega_l\right) = \delta_{l,T} \neq 0, \quad \omega_l \in [-\pi, \pi] \quad \text{for } 1 \leq l \leq m_0 \right\}.$$

We provide a consistency result for both m_0 and the actual location of the breaks λ_l^0, $1 \leq l \leq m_0$. Let $\mathcal{I} \subseteq \{2m_T, \ldots, \lfloor T/m_T \rfloor m_T - m_T\}$ denote a generic index set. One possibility would be to test for a break in $\mathcal{I}$ by using the test $\psi^*(\{X_r\}_{r\in\mathcal{I}})$, and if the test rejects one would estimate the break using

$$T\widehat{\lambda}_T(\mathcal{I}) = \underset{r\in\mathcal{I}}{\operatorname{argmax}} \max_{\omega\in[-\pi, \pi]} \mathrm{D}_{r,T}(\omega). \tag{6.1}$$

One would then update the set $\mathcal{I}$ by excluding a v_T-neighborhood of $T\widehat{\lambda}_T$ and repeat the above steps. This is a sequential top-down algorithm exploiting the classical idea of bisection. However, this procedure does not maximize power. For example, consider the first step of the algorithm in which we test for the first break. If the true break T_1^0 falls in between two indices in $\mathcal{I}$, r_1 and $r_2 = r_1 + m_T$ say, then $\min\{\mathrm{D}_{r_1,T}(\omega), \mathrm{D}_{r_2,T}(\omega)\} > \mathrm{D}_{r,T}(\omega)$ for $r \neq r_1, r_2$ with large probability and $\mathrm{D}_{r_1,T}(\omega)$ would take the difference between two adjacent blocks where the one to the left of r_1 contains observations prior to the change-point while the block to the right of r_1 contains observations from both regimes [a similar argument applies to $\mathrm{D}_{r_2,T}(\omega)$]. Clearly, this does not maximize either power or precision of the location estimate. One would need to compare two adjacent blocks exactly separated at $T_1^0 \in (r_1, r_2)$. This is a common problem in the change-point literature whose importance is often underestimated. Here we introduce a wild sequential top-down algorithm. Continuing with the above example, we draw randomly without replacement $K \geq 1$ separation points $r^\diamond$ from the interval (r_1, r_2) and compute for each separation point $\mathrm{D}_{r^\diamond,T}(\omega)$ where $r^\diamond \in (r_1, r_2)$. We take the maximum of them. Then, we update $\mathcal{I}$ by removing r_1 and adding $r^\diamond$. We repeat this for all indices in $\mathcal{I}$. Because the K separation points are drawn randomly, there is always some probability to pick up the separation point that guarantees the highest power. The careful reader may wonder why we do not take all integers between r_1 and r_2 and compute $\mathrm{D}_{r^\diamond,T}(\omega)$ for each of them. The reason is that in applications involving high frequency data (e.g., weakly, daily, and so on) that would be computationally more intensive if there are multiple breaks and one wants m_T to change when searching for an additional break since the sample where there are undetected breaks becomes smaller after a break is found. We are now ready to present the algorithm. Guidance as to suitable choice of K will be given below.

Let $v_T \to \infty$ with $v_T/T \to 0$ and $m_T/v_T \to 0$.

Algorithm 1. *Set* $\widehat{\mathcal{I}} = \{2m_T, 3m_T, \ldots, \lfloor T/m_T \rfloor m_T - m_T\}$ *and* $\widehat{\mathcal{T}} = \emptyset$.
(1) For $r \in \widehat{\mathcal{I}} \backslash \{2m_T\}$, *uniformly draw (without replacement)* K *points* $r_\cdot^\diamond$ *from* $\mathbf{I}(r) = \{r - m_T + 1, \ldots, r\}$ *and compute* $\overline{r}^\diamond = \arg\max_{k=1,\ldots,K} \max_{\omega \in [-\pi, \pi]} \mathrm{D}_{r_k^\diamond, T}(\omega)$; *set* $\widehat{\mathcal{I}} = (\widehat{\mathcal{I}} \backslash \{r\}) \cup \{\overline{r}^\diamond\}$.
(2) If $\psi^*(\{X_r\}_{r\in\widehat{\mathcal{I}}}) = 0$ *where* $\psi^*(\cdot)$ *is defined in Section 5, return* $\widehat{\mathcal{T}} = \emptyset$. *Otherwise proceed with step (3);*
(3) Estimate the change-point via (6.1) by using $\widehat{\mathcal{I}}$. *Name it* $T\widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right)$;
(4) Set $\widehat{\mathcal{I}} = \widehat{\mathcal{I}} \backslash \left\{T\widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right) - v_T, \ldots, T\widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right) + v_T\right\}$ *and* $\widehat{\mathcal{T}} = \widehat{\mathcal{T}} \cup \left\{T\widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right)\right\}$. *Repeat step (1).*

Finally, arrange the estimated change-points in $\widehat{\mathcal{T}}$ in chronological order and use the symbol $|\mathcal{S}|$ for the cardinality of a set $\mathcal{S}$. Name these change-points $\widehat{\lambda}_{l,T}$, $l = 1, \ldots, m_0$. To each $\widehat{\lambda}_{l,T}$ the procedure can return the frequency $\widehat{\omega}_l$

Assumption 6.2. *We have* $\delta_{l,T} \to 0$ *with* $\inf_{1\leq l\leq m_0} \delta_{l,T} \geq 2D^* m_T^{-1/2} (\log(T))^{2/3}$. *For* $\nu_T = o(T/v_T)$ *it holds that* $\inf_{1\leq l\leq m_0-1} \left|\lambda_{l+1}^0 - \lambda_l^0\right| \geq \nu_T^{-1}$.

Proposition 6.2. *Assume Assumption 3.1, 3.3-3.4, 4.1, eq. (3.2) and Condition 1-(i) hold. Then, under* $\mathcal{H}_{1,\mathrm{M}}$ *we have (i)* $\mathbb{P}\left(\left|\widehat{\mathcal{T}}\right| = m_0\right) \to 1$ *and* $\sup_{1\leq l\leq m_0} \left|\widehat{\lambda}_{l,T} - \lambda_l^0\right| = o_{\mathbb{P}}(1)$, *and (ii)* $\sup_{1\leq l\leq m_0} \left|\widehat{\lambda}_{l,T} - \lambda_l^0\right| = O_{\mathbb{P}}\left(\sqrt{m_T \log(T)}/\left(T \inf_{1\leq l\leq m_0} \delta_{l,T}\right)\right)$. *Furthermore, if* $K = O(a_T m_T)$ *with* $a_T \in (0, 1]$ *such that* $a_T \to 1$, *then the breaks are detected in descending order of magnitude.*

The number of draws K maybe fixed or increase with the sample size. However, the algorithm can return break dates according to the descending order of break magnitudes only if $K = O(a_T m_T)$ with a_T as above. The number of breaks can go to infinity as long as $m_0/\nu_T \to 0$. Note that at each loop of the algorithm it is not possible to know to which λ_l^0 $(l = 1, \ldots, m_b)$ the estimate $\widehat{\lambda}_T$ is consistent for. Only after all breaks are detected and we rearrange the estimated change-points in $\widehat{\mathcal{T}}$ in chronological order, we can learn such information.

7 Implementation

In this section we explain how to consistently estimate $\sigma_r(\omega)$ and $\sigma_{f,r}(\omega)$, and how to choose the tuning parameters. Let $\widehat{I}_{h,T}(j/T, \omega) = I_{h,T}(j/T, \omega) - \widetilde{I}_{r,T}(\omega)$ for $j \in S_r$. Define $\widehat{\sigma}_r^2(\omega) = \sum_{j=-M_{S,T}+1}^{M_{S,T}-1} K_1(b_{1,T} j m_{S,T}) \widehat{\Gamma}_r(j)$ where

$$\widehat{\Gamma}_r(j) = \begin{cases} M_{S,T}^{-1} \sum_{t \in \left\{\mathbf{s}_r / \left\{\ldots, r m_T - m_T/2 + 1 + m_{S,T}(j-1), j>0\right\}\right\}} \widehat{I}_{h,T}(t/T, \omega) \widehat{I}_{h,T}((t - j m_{S,T})/T, \omega), & j \geq 0 \\ M_{S,T}^{-1} \sum_{t \in \left\{\mathbf{s}_r / \left\{\ldots, r m_T - m_T/2 + 1 + m_{S,T}(j-1), j>0\right\}\right\}} \widehat{I}_{h,T}((t)/T, \omega) \widehat{I}_{h,T}((t + j m_{S,T})/T, \omega) & j < 0 \end{cases}$$

The estimate $\widehat{\sigma}_r^2(\omega)$ is a local long run variance estimator where K_1 is a kernel and $b_{1,T}$ is the associated bandwidth.

Let $\widehat{f}_{h,T}(j/T,\,\omega) = f_{h,T}(j/T,\,\omega) - \widetilde{f}_{r,T}(\omega)$ for $j \in S_r$. Define $\widehat{\sigma}_{f,s}^2(\omega)$ similarly to $\widehat{\sigma}_r^2(\omega)$ but $\widehat{f}_{h,T}$ in place of $\widehat{I}_{h,T}$. Consistency of either estimates follows from the results in Vogt (2012) and Casini (2021).

The choice of the sequences can be based on mean-squares error (MSE) criterion or cross-validation where one can exploit derived results for locally stationary series. For example, data-dependent methods for bandwidths in the context of locally stationary processes were investigated by Casini (2021), Dahlhaus (2012), Dahlhaus and Giraitis (1998) and Richter and Dahlhaus (2019). The optimal amount of smoothing depends on θ, q and on the amount of dependence in the data. Further, the optimal values have to satisfy Condition 1. In this work we focus on the optimal order for the bandwidths, neglecting the constants. We relegate to future work a more detailed analysis of data-dependent methods for this problem where multiple smoothing directions are present. For spectral densities satisfying Lipschitz continuity $\theta = 1$ so that $m_T \propto T^{2/3-\epsilon}$ while for $\theta = 1/2$ we have $m_T \propto T^{1/2-\epsilon}$, $\epsilon > 0$. Often Lipschitz continuity is assumed and so we use optimal bandwidths for $\theta = 1$. Assuming $q = 4$ and γ large enough we have $\tau_T \propto T^{1/4}$ and so the optimal values that satisfy Condition 1 are $m_T \propto T^{2/3-\epsilon}$, $n_T \propto T^{5/8-\epsilon}$ and $b_{W,T} = n_T^{-1/6}$. We employ two data tapers $h(t/T) = 1$ for all $0 \leq t \leq T$ and $h(t/T) = \exp\left(-2^{-1}\left((t - T/2)/T\right)^2\right)$ for all $0 \leq t \leq T$. Also we set $v_T = T^{2/3}$so as to guarantee the condition $m_T/v_T \to 0$. We set $b_{1,T} = M_{S,T}^{-1/3}$.

8 Small-Sample Evaluations

In this section, we conduct a Monte Carlo analysis to evaluate the properties of the proposed methods. Our simulation exercise is similar to Last and Shumway (2008). We first discuss the detection of the breaks and then the localization of the change-points. We investigate different types of change-point and compare the statistics $\mathrm{I}_{\max,T}(\omega)$, $\mathrm{I}_{\mathrm{Dmax},T}$, $\mathrm{S}_{\mathrm{Dmax},T}$, the statistic in Adak (1998) and in Last and Shumway (2008). We consider the following data-generating processes where in all models the innovation e_t is a Gaussian white noise $e_t \sim i.i.d.\,\mathscr{N}(0,\,1)$. Model M1 involves a stationary AR(1) process $X_t = \rho X_{t-1} + e_t$ with $\rho = 0.3$ and 0.6, while M2 involves a locally stationary AR(1) $X_t = \rho(t/T)\,X_{t-1} + e_t$ where $\rho(t/T) = 0.4\cos(0.8 - \cos(2t/T))$. Note that $\rho(t/T)$ varies smoothly from 0.1389 to 0.4. Model M1 and M2 are used to verify the finite-sample size of the tests. We verify the power in models M3 and M4 using the model in M1 and M2 for the first regime of M3 and M4, respectively. That is, in model M3 we specify

$$X_t = \begin{cases} 0.3X_{t-1} + e_t, & 1 \leq t \leq \lfloor T\lambda_1^0 \rfloor \\ 0.6X_{t-1} + 0.7e_t, & \lfloor T\lambda_1^0 \rfloor + 1 \leq t \leq \lfloor T\lambda_2^0 \rfloor \\ 0.6X_{t-1} + e_t & \lfloor T\lambda_2^0 \rfloor + 1 \leq t \leq T. \end{cases},$$

while in model M4 we specify

$$X_t = \begin{cases} \rho\left(t/T\right) X_{t-1} + 0.7e_t, & 1 \leq t \leq \lfloor T\lambda_1^0 \rfloor \\ 0.8X_{t-1} + e_t, & \lfloor T\lambda_1^0 \rfloor + 1 \leq t \leq \lfloor T\lambda_2^0 \rfloor \\ \rho\left(t/T\right) X_{t-1} + 0.7e_t & \lfloor T\lambda_2^0 \rfloor + 1 \leq t \leq T, \end{cases},$$

where $\rho\left(t/T\right)$ is as in model M2. In model M3 there are three regimes. The model in the first regime is as in M1, the second regime involves higher serial dependence while in the third regime the variance doubles relative to the second regime. In model M4 the data-generating process in the first regime is the same as in model M2. In the second regime the model is a stationary AR process with strong serial dependence while in the third regime X_t assumes the same dynamics as in the first regime. Model M3-M4 feature alternative hypotheses in the forms of breaks in the spectrum. We consider the alternative hypothesis in of more rough variation without signifying a break in the following model, named M5, $X_t = \sigma\left(t/T\right) e_t$, where $\sigma^2\left(t/T\right) = \overline{\sigma}^2 + \cos\left(1 + \cos\left(10t/T\right)\right)$ with $\overline{\sigma}^2 = 1$. Note that even though $\sigma^2\left(\cdot\right)$ is locally stationary, it alternates its degree of smoothness throughout the sample. It starts from $\sigma^2\left(\cdot\right) = 1$ and it maintains this value for some time, then it increases slowly to $\sigma^2\left(\cdot\right) = 2$ and keeps this value for some time. Finally, it decreases slowly back $\sigma^2\left(\cdot\right) = 1$. Thus, $\sigma^2\left(\cdot\right)$ alternates periods where it is constant and periods where it becomes rougher (i.e., it varies smoothly). Importantly, no break occurs; only a change in the smoothness as specified in $\mathcal{H}_1'$ in Section 5.

Next, we consider the estimation of the number of change-points (m_0) and of the locations of the change-points. We consider the following two models, both with $m_0 = 2$. Model M6 is given by

$$X_t = \begin{cases} 0.7e_t, & 1 \leq t \leq \lfloor T\lambda_1^0 \rfloor \\ 0.6X_{t-1} + 0.7e_t, & \lfloor T\lambda_1^0 \rfloor + 1 \leq t \leq \lfloor T\lambda_2^0 \rfloor \\ 0.6X_{t-1} + e_t & \lfloor T\lambda_2^0 \rfloor + 1 \leq t \leq T. \end{cases},$$

while model M7 is the same as model M4. We set $\lambda_1^0 = 0.33$ and $\lambda_2^0 = 0.66$ throughout. We use

$T = 250$, 500, 1000 in model M1-M5 and $T = 1000$ in model M6-M7. The number of simulations is 5,000. Table 1 reports the results for the size of the tests for model M1-M2.

Next, we move to the estimation to the number of change-points m_0. Table 1 reports the summary statistics for $\widehat{m} - m_0$. For each method, we list the estimated change-points in chronological order so that we can focus on $\widehat{T}_1 - T_1^0$ and $\widehat{T}_2 - T_2^0$ where $\widehat{T}_1 < \widehat{T}_2$.
Our results show that raising K from $K = 1$ (default) to $K = m_T/3$ lead to more precise estimates without excessively increasing the computing time.

9 Conclusions

We develop a theoretical framework for inference about the smoothness of the spectral density over time. We provide frequency domain statistical tests for the detection of discontinuities in the spectrum of a segmented locally stationary time series and for changes in the regularity exponent of the spectral density over time. We provide different test statistics depending on whether prior knowledge about the frequency component at which the change-point occurs is available. The null distribution of the test follows an extreme value distribution. We rely on the theory on minimax-optimal testing developed by Ingster (1993). We determine the optimal rate for the minimax distinguishable boundary, i.e., the minimum break magnitude such that we are still able to uniformly control type I and type II errors. We propose a novel procedure for estimation of the change-points based on a wild sequential top-down algorithm and show its consistency under shrinking shifts and possibly growing number of change-points. The advantage of using frequency domain methods to detect change-points is that it does not require to make assumptions about the data-generating process under the null hypotheses beyond the fact that the spectrum is bounded. Furthermore, the method allows for a broader range of alternative hypotheses than time domain methods which usually have power against some specific alternatives but not against others. Therefore, our method can be useful across many fields.

References

Adak, S. (1998): "Time-Dependent Spectral Analysis of Nonstationary Time Series," *Journal of the American Statistical Association*, (444), 1488–1501.

Anderson, T. W. (1971): *The Stastical Analysis of Time Series*. New York: Wiley.

Andrews, D. W. K. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change-Point," *Econometrica*, 61(4), 821–56.

Aue, A., D. Dubart Nourinho, and S. Hormann (2015): "On the Prediction of Stationary Functional Time Series," *Journal of the American Statistical Association*, 110(509), 378–392.

Aue, A., and L. Hórvath (2013): "Structural Breaks in Time Series," *Journal of Time Series Analysis*, 34(1), 1–16.

Aue, A., and A. van Delft (2020): "Testing for Stationarity of Functional Time Series in the Frequency Domain," *Annals of Statistics*, 48(5), 2505–2547.

Bai, J. (1994): "Least Squares Estimation of a Shift in Linear Processes," *Journal of Time Series*, 15(5), 453–472.

Bai, J., and P. Perron (1998): "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66(1), 47–78.

Berkes, I., and W. Philipp (1979): "Approximation Theorems for Independent and Weakly Dependent Random Vectors," *Annals of Probability*, 7(1), 29–54.

Bibinger, M., M. Jirak, and M. Vetter (2017): "Nonparametric Change-Point Analysis of Volatility," *Annals of Statistics*, 45(4), 1542–1578.

Bobkov, S. G., G. P. Chistyakov, and F. Götze (2013): "Rate of Convergence and Edgeworth-Type Expansion in the Entropic Central Limit Theorem," *Annals of Probability*, 41(4), 2479–2512.

Brillinger, D. R. (1975): *Time Series Data Analysis and Theory*. New York: Holt, Rinehart and Winston.

Brillinger, D. R., and M. Rosenblatt (1967): "Asymptotic Theory of k-th Order Spectra," In *Specfral Analysis of Time Series*, Ed. B. Harris, pp. 153–188. New York : Wiley.

Brown, L. D., and M. G. Low (1996): "Asymptotic Equivalence of Nonparametric Regression and White Noise," *Annals of Statistics*, 24(6), 2384–2398.

Casini, A. (2018): "Tests for Forecast Instability and Forecast Failure under a Continuous Record Asymptotic Framework," *arXiv preprint arXiv:1803.10883*.

——— (2021): "Theory of Evolutionary Spectra for Heteroskedasticity and Autocorrelation Robust Inference in Possibly Misspecified and Nonstationary Models," *arXiv preprint arXiv:2103.02981*.

Casini, A., T. Deng, and P. Perron (2021): "Theory of Low Frequency Contamination from Nonstationarity and Misspecification: Consequences for HAR Inference," *arXiv preprint arXiv:2103.01604*.

Casini, A., and P. Perron (2019): "Structural Breaks in Time Series," *Oxford Research Encyclopedia of Economics and Finance*, Oxford University Press.

——— (2020): "Generalized Laplace Inference in Multiple Change-Points Models," *Econometric Theory*, forthcoming.

——— (2021a): "Continuous Record Asymptotics for Change-Point Models," *arXiv preprint arXiv:1803.10881*.

——— (2021b): "Minimax MSE Bounds and Nonlinear VAR Prewhitening for Long-Run Variance Estimation Under Nonstationarity," *arXiv preprint arXiv:2103.02235*.

——— (2021c): "Supplement to "Change-Point Analysis of Time Series with Evolutionary Spectra"," *Unpublished Manuscript, Department of Economics and Finance, University of Rome Tor Vergata.*

Coates, D. S., and P. J. Diggle (1986): "Tests for Comparing Two Estimated Spectral Densities," *Journal of Time Series A*, 7(1), 7–20.

Dahlhaus, R. (1996a): "Asymptotic Statistical Inference for Nonstationary Processes with Evolutionary Spectra," in *Athens Conference on Applied Probability and Time Series Vol II*, ed. by P. M. Robinson, and M. Rosenblatt, Lecture Notes in Statistics 115, pp. 145–159. Springer, New York.

——— (1996b): "On the Kullback-Leibler Information Divergence of Locally Stationary Processes," *Stochastic Processes and their Applications*, 62(1), 139–168.

——— (1997): "Fitting Time Series Models to Nonstationary Processes," *Annals of Statistics*, 25(1), 1–37.

——— (2012): "Locally Stationary Processes," in *Handbook of Statistics*, ed. by T. Subba Rao, S. Subba Rao, and C. Rao, vol. 30, pp. 351–413. Elsevier.

Dahlhaus, R., and L. Giraitis (1998): "On the Optimal Segment Length for Parameter Estimates for Locally Stationary Time Series," *Journal of Time Series Analysis*, 19(6), 629–655.

Fryzlewicz, P. (2014): "Wild Binary Segmentation for Multiple Change-Point Detection," *Annals of Statistics*, 42(6), 2243–2281.

Grenander, U., and M. Rosenblatt (1957): *Statistical Analysis of Stationary Time Series*. New York: Wiley.

Hannan, E. J. (1970): *Multiple Time Series*. New York: Wiley.

Hinkley, D. V. (1971): "Inference About the Change-Point from Cumulative Sum Tests," *Biometrika*, 58(3), 509–523.

Hoffmann, M., and R. Nickl (2011): "On Adaptive Inference and Confidence Bands," *Annals of Statistics*, 39(5), 2383–2409.

Hörmann, S., and P. Kokoszka (2010): "Weakly Dependent Functional Data," *Annals of Statistics*, 38(3), 1845–1884.

Horváth, L. (1993): "The Maximum Likelihood Method for Testing Changes in the Parameters of Normal Observations," *The Annals of Statistics*, 21(2), 671–680.

Huang, H.-Y., H. Ombao, and D. S. Stoffer (2004): "Discrimination and Classification of Nonstationary Time Series Using the SLEX Model," *Journal of the American Statistical Association*, 99(467), 763–774.

Ingster, Y. I. (1993): "Asymptotically Minimax Hypothesis Testing for Nonparametric Alternatives I, II, III," *Mathematical Methods of Statistics*, 2(2), 85–114.

Ingster, Y. I., and I. Suslina (2003): *Nonparametric Goodness-of-Fit Testing under Gaussian Models*, Lecture Notes in Statistics, 169. Springer-Verlag, New York.

Kakizawa, Y., R. H. Shumway, and M. Taniguchi (1998): "Discrimination and Clustering for Multivariate Time Series," *Journal of the American Statistical Association*, 93(441), 328–340.

Last, M., and R. Shumway (2008): "Detecting Abrupt Changes in a Piecewise Locally Stationary Time Series," *Journal of Multivariate Analysis*, 99(2), 191–214.

Loader, C. R. (1996): "Change Point Estimation Using Nonparameteric Regression," *Annals of Statistics*, 24(4), 1667–1678.

Müller, H.-G. (1992): "Change-Points in Nonparametric Regression Analysis," *Annals of Statistics*, 20(2), 737–761.

Müller, H.-G., and U. Stadtmuller (1999): "Discontinuous Versus Smooth Regression," *Annals of Statistics*, 27(1), 299–337.

Panaretos, V. M., and S. Tavakoli (2013): "Fourier Analysis of Stationary Time Series in Function Space," *Annals of Statistics*, 41(2), 568–603.

Parzen, E. (1957): "On Consistent Estimates of the Spectrum of a Stationary Time Series," *Annals of Mathematical Statistics*, 28(2), 329–348.

Richter, S., and R. Dahlhaus (2019): "Cross Validation for Locally Stationary Processes," *Annals of Statistics*, 47(4), 2145–2173.

Rosenblatt, M. (1959): "Statistical Analysis of Stochastic Processes with Stationary Residuals," In *Probability and Statistics*, Ed. U. Grenander, pp 246–275.

Spokoiny, V. G. (1998): "Estimation of a Function with Discontinuities via Local Polynomial Fit with an Adaptive Window Choice," *Annals of Statistics*, 26(4), 1356–1378.

van Delft, A., and M. Eichler (2018): "Locally Stationary Functional Time Series," *Electronic Journal of Statistics*, 12(1), 107–170.

VOGT, M. (2012): "Nonparametric Regression for Locally Stationary Time Series," *Annals of Statistics*, 40(5), 2601–2633.

WU, W. B. (2005): "Non-Linear System Theory: Another Look at Dependence," *Proceedings of the National Academy of Sciences of the United States of America*, 102(40), 14150–14154.

——— (2007a): "Strong Invariance Principles for Dependent Random Variables," *Annals of Probability*, 35(6), 2294–2320.

——— (2007b): "Strong Invariance Principles for Dependent Random Variables," *Annals of Probability*, 35(6), 2294–2320.

WU, W. B., AND Z. ZHAO (2007): "Inference of Trends in Time Series," *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 69(3), 391–410.

WU, W. B., AND Z. ZHOU (2011): "Gaussian Approximation for Non-Stationary Multiple Time Series," *Statistica Sinica*, pp. 1397–1413.

YAO, Y. (1987): "Approximating the Distribution of the ML Estimate of the Change-Point in a Sequence of Independent Random Variables," *Annals of Statistics*, 15, 1321–1328.

ZHANG, T., AND W. B. WU (2012): "Inference of Time-Varying Regression Models," *Annals of Statistics*, 40, 1376–1402.

ZHOU, Z. (2013): "Heteroscedasticity and Autocorrelation Robust Structural Change Detection," *Journal of the American Statistical Association*, 103(502), 726–740.

10 Appendix

10.1 Tables

Table 1: Empirical small-sample size for model M1-M2

	Model M1		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}$	0.039	0.043	0.053
$\mathrm{S}_{\mathrm{Dmax},T}$	0.029	0.049	0.037
$\mathrm{R}_{\max,T}$	0.040	0.058	0.042
$\mathrm{R}_{\mathrm{Dmax},T}$	0.015	0.032	0.034
$\widehat{D}$ statistic	0.581	0.471	0.068
	Model M4		
	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}$	0.061	0.059	0.061
$\mathrm{S}_{\mathrm{Dmax},T}$	0.035	0.055	0.058
$\mathrm{R}_{\max,T}$	0.036	0.035	0.036
$\mathrm{R}_{\mathrm{Dmax},T}$	0.015	0.032	0.035
$\widehat{D}$ statistic	0.731	0.583	0.102

Table 2: Empirical small-sample power for model M3-M5

	Model M3		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}$	0.694	0.850	0.889
$\mathrm{S}_{\mathrm{Dmax},T}$	0.734	0.890	0.921
$\mathrm{R}_{\max,T}$	0.768	0.940	0.973
$\mathrm{R}_{\mathrm{Dmax},T}$	0.456	0.752	0.874
$\widehat{D}$ statistic	0.961	0.967	0.790
	Model M4		
	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}$	0.868	0.977	0.973
$\mathrm{S}_{\mathrm{Dmax},T}$	0.938	0.988	0.996
$\mathrm{R}_{\max,T}$	0.927	0.997	0.999
$\mathrm{R}_{\mathrm{Dmax},T}$	0.775	0.983	0.998
$\widehat{D}$ statistic	1.000	1.000	1.000
	Model M5		
	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}$	0.223	0.475	0.565
$\mathrm{S}_{\mathrm{Dmax},T}$	0.325	0.801	0.918
$\mathrm{R}_{\max,T}$	0.028	0.237	0.369
$\mathrm{R}_{\mathrm{Dmax},T}$	0.001	0.189	0.304
$\widehat{D}$ statistic	0.834	0.695	0.172

Table 3: Summary statistics for the empirical distribution of $\widehat{m} - m_0$

		Summary of $\widehat{m} - m_0$			
	Percent time $\widehat{m} = m_0$		$Q_{0.25}$	Median	$Q_{0.75}$
	Model M6				
Algorithm 1 ($\widehat{T}_j,\ j = 1, \dots, \widehat{m}$)	85.50	$\widehat{T}_1$	299	333	352
		$\widehat{T}_2$	622	663	688
	Model M7				
Algorithm 1 ($\widehat{T}_j,\ j = 1, \dots, \widehat{m}$)	80.12	$\widehat{T}_1$	307	340	350
		$\widehat{T}_2$	603	650	685

Supplemental Material to

Change-Point Analysis of Time Series with Evolutionary Spectra

Alessandro Casini

Department of Economics
Boston University

5th March 2021

Abstract

This supplemental material is structured as follows. Section S.A contains the Mathematical Appendix which includes all proofs of the results in the paper.

S.A Mathematical Appendix

S.A.1 Preliminary Lemmas

Let $L_T : \mathbb{R} \to \mathbb{R}$, $T \in \mathbb{R}_+$ be the 2π-periodic extension of

$$L_T(\omega) \triangleq \begin{cases} T, & |\omega| \leq 1/T, \\ 1/|\omega|, & 1/T \leq |\omega| \leq \pi. \end{cases}$$

For a complex-valued function f define $H_{n_T}(f(\cdot), \omega) = \sum_{s=0}^{n_T-1} f(s) \exp(-i\omega s)$, and, for the taper $h(x)$, $H_{k,n_T}(\omega) = H_{n_T}\left(h^k\left(\frac{\cdot}{n_T}\right), \omega\right)$, and $H_{n_T}(\omega) = H_{1,n_T}(\omega)$.

Lemma S.A.1. *Let $\Pi \triangleq (-\pi, \pi]$. We obtain the following with a constant K independent of T: (i) $L_T(\omega)$ is monotone increasing in T and decreasing in $\omega \in [0, \pi]$; (ii) $\int_\Pi L_T(\alpha) d\alpha \leq K \ln T$ for $T > 1$.*

Proof of Lemma S.A.1. See Lemma A.4 in Dahlhaus (1997). □

Lemma S.A.2. *Let n_T, $T \in \mathbb{N}$. Suppose $h(\cdot)$ satisfies Assumption 3.3 and $\vartheta : [0, 1] \to \mathbb{R}$ is differentiable with bounded derivative. Then we have for $0 \leq t \leq n_T$,*

$$\begin{aligned} H_{n_T}\left(\vartheta\left(\frac{\cdot}{n_T}\right) h\left(\frac{\cdot}{n_T}\right), \omega\right) &= \vartheta\left(\frac{t}{n_T}\right) H_{n_T}(\omega) + O\left(\sup_x |d\vartheta(x)/dx| \frac{n_T}{T} L_{n_T}(\omega)\right) \\ &= O\left(\sup_{x \leq n_T/T} |\vartheta(x)| L_{n_T}(\omega) + \sup_x |d\vartheta(x)/dx| L_{n_T}(\omega)\right). \end{aligned}$$

The same holds, if $\vartheta(\cdot/T)$ is replaced on the left side by numbers $\psi_{s,T}$ with $\sup_s |\vartheta_{s,T} - \vartheta(s/T)| = O(T^{-1})$.

Proof of Lemma S.A.2. Dahlhaus (1997) proved this result under differentiability of $h(\cdot)$. By Abel's transformation [cf. Exercise 1.7.13 in Brillinger (1975)],

$$\begin{aligned} H_{n_T}\left(\vartheta\left(\frac{\cdot}{T}\right) h\left(\frac{\cdot}{n_T}\right), \omega\right) - \vartheta\left(\frac{t}{T}\right) H_{n_T}(\omega) &= \sum_{s=0}^{n_T-1} \left[\vartheta\left(\frac{s}{T}\right) - \vartheta\left(\frac{t}{T}\right)\right] h\left(\frac{s}{n_T}\right) \exp(-i\omega s) \\ &= -\sum_{s=0}^{n_T-1} \left[\vartheta\left(\frac{s}{T}\right) - \vartheta\left(\frac{s-1}{T}\right)\right] H_s\left(h\left(\frac{\cdot}{n_T}\right), \omega\right) \\ &\quad + \left[\vartheta\left(\frac{n_T-1}{T}\right) - \vartheta\left(\frac{t}{T}\right)\right] H_{n_T}\left(h\left(\frac{\cdot}{n_T}\right), \omega\right). \end{aligned} \tag{S.1}$$

By repeated application of Abel's transformation,

$$\begin{aligned} H_s\left(h\left(\frac{\cdot}{n_T}\right), \omega\right) &= \sum_{t=0}^{s-1} h\left(\frac{t}{s}\right) \exp(-i\omega t) \\ &= \sum_{t=0}^{s-1} \left(h\left(\frac{t}{s}\right) - h\left(\frac{t-1}{s}\right)\right) H_t(1, \omega) \\ &\quad + h\left(\frac{n_T-1}{n_T}\right) H_{n_T}(1, \omega) \\ &= \sum_{t=0}^{s-1} \left(h\left(\frac{t}{s}\right) - h\left(\frac{t-1}{s}\right)\right) H_t(1, \omega) + 0, \end{aligned}$$

where we have used $h\left(\left(n_T-1\right)/n_T\right)-h\left(1\right)=O\left(n_T^{-1}\right)$ and $h\left(x\right)=0$ for $x\notin[0,\,1)$. Since $h\left(\cdot\right)$ is of bounded variation, if $|\omega|\leq 1/s$ we have

$$\begin{aligned}\sum_{t=0}^{s-1}\left|\left(h\left(\frac{t}{s}\right)-h\left(\frac{t-1}{s}\right)\right)\right|\left|H_t\left(1,\,\omega\right)\right| &\leq \sum_{t=0}^{s-1}t\left|\left(h\left(\frac{t}{s}\right)-h\left(\frac{t-1}{s}\right)\right)\right|\\ &\leq \left(s-1\right)\sum_{t=0}^{s-1}\left|\left(h\left(\frac{t}{s}\right)-h\left(\frac{t-1}{s}\right)\right)\right|\\ &\leq C\left(s-1\right),\end{aligned}$$

whereas if $1/s\leq|\omega|\leq\pi$ we have,

$$\begin{aligned}\sum_{t=0}^{s-1}\left|\left(h\left(\frac{t}{s}\right)-h\left(\frac{t-1}{s}\right)\right)\right|\left|H_t\left(1,\,\omega\right)\right| &\leq C\frac{1}{|\omega|}\sum_{t=0}^{s-1}\left|\left(h\left(\frac{t}{s}\right)-h\left(\frac{t-1}{s}\right)\right)\right|\\ &\leq C\frac{1}{|\omega|}.\end{aligned}$$

Thus, $H_s\left(h\left(\frac{\cdot}{n_T}\right),\,\omega\right)\leq L_s\left(\omega\right)\leq L_{n_T}\left(\omega\right)$ where the last inequality follows by Lemma S.A.1-(i). It follows from (S.1) that,

$$H_{n_T}\left(\vartheta\left(\frac{\cdot}{T}\right)h\left(\frac{\cdot}{n_T}\right),\,\omega\right)-\vartheta\left(\frac{t}{T}\right)H_{n_T}\left(\omega\right)=O\left(\sup_{x\leq n_T/T}\left|\vartheta\left(x\right)\right|L_{n_T}\left(\omega\right)+\sup_{x}\left|d\vartheta\left(x\right)/dx\right|L_{n_T}\left(\omega\right)\right).\square$$

Lemma S.A.3. *Assume that $h^{(a_j)}\left(x\right)$ satisfies Assumption 3.3 for all $j=1,\ldots,\,p$, then we have for some C with $0<C<\infty$,*

$$\left|\sum_{s=0}^{n_T-1}h_T^{(a_1)}\left(s+k_1\right)\cdots h_T^{(a_{p-1})}\left(s+k_{p-1}\right)h_T^{(a_1)}\left(s\right)\exp\left(-i\omega s\right)-H_T^{(a_1,\cdots,\,a_p)}\left(\omega\right)\right|\leq C\left(|k_1|+\ldots+|k_{p-1}|\right).$$

Proof of Lemma S.A.3. See Lemma P4.1 in Brillinger (1975). $\square$

Lemma S.A.4. *Let $\{Y_T\}$ be a sequence of p vector-valued random variables, with (possibly) complex components, and such that all cumulants of the variate $\left(Y_T^{(a_1)},\,\overline{Y}_T^{(a_2)},\ldots,\,Y_T^{(a_p)},\,\overline{Y}_T^{(a_p)}\right)$ exist and tend to the corresponding cumulants of a variate $\left(Y^{(a_1)},\,\overline{Y}^{(a_1)},\ldots,\,Y^{(a_p)},\,\overline{Y}^{(a_p)}\right)$ that is determined by its moments. Then Y_T tends in distribution to a variate having components $Y^{(a_1)},\ldots,\,Y^{(a_p)}$.*

Proof of Lemma S.A.4. It follows from Lemma P4.5 in Brillinger (1975).

S.A.2 Proofs of the Results of Section 3

S.A.2.1 Proof of Theorem 3.1

For $\lfloor Tu\rfloor-n_T/2+1\leq t_1,\ldots,\,t_p\leq\lfloor Tu\rfloor+n_T/2-1$,

$$\begin{aligned}&\mathrm{cum}\left(X_{t_1,T},\ldots,\,X_{t_p,T}\right)\\ &=\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}\exp\left(it_1\omega_1+\cdots+it_p\omega_p\right)A_{t_1,T}^0\left(\omega_1\right)\cdots A_{t_p,T}^0\left(\omega_p\right)\eta\left(\sum_{j=1}^{p}\omega_j\right)g_p\left(\omega_1,\ldots,\,\omega_{p-1}\right)d\omega_1\cdots d\omega_p.\end{aligned}$$

We can replace $A^0_{t_j,T}(\omega_j)$ by $A(t_j/T, \omega_j)$ using (2.3), and then replace $A(t_j/T, \omega_j)$ by $A(\lfloor Tu \rfloor, \omega_j)$ using the smoothness of $A(u, \cdot)$. Altogether, this gives an error $O(n_T/T)$. Let $t_1 = t_p + k_1, \dots, t_{p-1} = t_p + k_{p-1}$. We have

$$
\begin{aligned}
&\operatorname{cum}(X_{t_1,T}, \dots, X_{t_p,T}) \\
&\quad = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(i\left((\omega_1 + \cdots + \omega_{p-1}) t_p + \omega_1 k_1 + \cdots + \omega_{p-1} k_{p-1} + t_p \omega_p\right)\right) \\
&\qquad \times A(\lfloor Tu \rfloor, \omega_1) \cdots A(\lfloor Tu \rfloor, \omega_p)\, \eta\left(\sum_{j=1}^{p} \omega_j\right) g_p(\omega_1, \dots, \omega_{p-1})\, d\omega_1 \cdots d\omega_p + O(n_T/T) \\
&\quad = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(i\left((\omega_1 + \cdots + \omega_{p-1} + \omega_p) t_p + \omega_1 k_1 + \cdots + \omega_{p-1} k_{p-1}\right)\right) \\
&\qquad \times A(\lfloor Tu \rfloor, \omega_1) \cdots A(\lfloor Tu \rfloor, \omega_p)\, \eta\left(\sum_{j=1}^{p} \omega_j\right) g_p(\omega_1, \dots, \omega_{p-1})\, d\omega_1 \cdots d\omega_p + O(n_T/T) \\
&\quad \triangleq \kappa_{Tu,t_p}(k_1 \dots, k_{p-1}) + O(n_T/T).
\end{aligned}
$$

This shows that $\operatorname{cum}(X_{t_1,T}, \dots, X_{t_p,T})$ depends on t_p only through $\exp(i(\omega_1 + \cdots + \omega_{p-1} + \omega_p) t_p)$. By Lemma P4.1 of Brillinger (1975),

$$
\begin{aligned}
&\left| \sum_{s=0}^{n_T-1} h_{a_1}\left(\frac{s+k_1}{n_T}\right) \cdots h_{a_{p-1}}\left(\frac{s+k_{p-1}}{n_T}\right) h_{a_p}\left(\frac{s}{n_T}\right) \exp\left(i \sum_{j=1}^{p} \omega_j s\right) - H_T^{(a_1,\cdots,a_p)}\left(\sum_{j=1}^{p} \omega_j\right) \right| \\
&\leq C\left(|k_1| + \ldots + |k_{p-1}|\right).
\end{aligned}
$$

The cumulant has then the form,

$$
\begin{aligned}
&\operatorname{cum}\left(d_T^{(a_1)}(u, \omega_1), \dots, d_T^{(a_p)}(u, \omega_p)\right) \\
&\quad = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} H_{n_T}\left(A^{0,(a_1)}_{\lfloor Tu \rfloor - n_T/2 + 1 + \cdot, T}(\gamma_1)\, h\left(\frac{\cdot}{n_T}\right), \omega_1 - \gamma_1\right) \\
&\qquad \times H_{n_T}\left(A^{0,(a_2)}_{\lfloor Tu \rfloor - n_T/2 + 1 + \cdot, T}(\gamma_2)\, h\left(\frac{\cdot}{n_T}\right), \omega_2 - \gamma_2\right) \\
&\qquad \times \cdots \\
&\qquad \times H_{n_T}\left(A^{0,(a_p)}_{\lfloor Tu \rfloor - n_T/2 + 1 + \cdot, T}(\gamma_p)\, h\left(\frac{\cdot}{n_T}\right), \omega_p - \gamma_p\right) \\
&\qquad \times \exp\left\{i\left((\gamma_1 + \cdots + \gamma_p) \lfloor Tu \rfloor\right)\right\} \eta\left(\sum_{j=1}^{p} \gamma_j\right) g_p(\gamma_1, \dots, \gamma_{p-1})\, d\gamma_1 \cdots d\gamma_p.
\end{aligned}
$$

By Lemma S.A.2, the latter is equal to

$$
\begin{aligned}
&\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} A^{(a_1)}(u, \gamma_1) \cdots A^{(a_p)}(u, \gamma_p) \\
&\qquad \times H_{n_T}(\omega_1 - \gamma_1) \cdots H_{n_T}(\omega_p - \gamma_p)
\end{aligned}
$$

$$\times \exp\left(i\left(\left(\gamma_1+\cdots+\gamma_p\right)\lfloor Tu\rfloor\right)\right)\eta\left(\sum_{j=1}^{p}\gamma_j\right)g_p\left(\gamma_1,\ldots,\gamma_{p-1}\right)d\gamma_1\cdots d\gamma_p,$$

plus a remainder term R_u with

$$\begin{aligned}
|R_u| &\leq C\frac{n_T}{T}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}L_{n_T}\left(\omega_1-\gamma_1\right)\cdots L_{n_T}\left(\omega_p-\gamma_p\right)\exp\left(i\left(\left(\gamma_1+\cdots+\gamma_p\right)\lfloor Tu\rfloor\right)\right) \\
&\quad\times\eta\left(\sum_{j=1}^{p}\gamma_j\right)g_p\left(\gamma_1,\ldots,\gamma_{p-1}\right)d\gamma_1\cdots d\gamma_p \\
&\leq C\frac{n_T}{T}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}L_{n_T}\left(\omega_1-\gamma_1\right)\cdots L_{n_T}\left(\omega_p-\gamma_p\right)d\gamma_1\cdots d\gamma_p \\
&\leq C\frac{n_T}{T}\left(\ln n_T\right)^p,
\end{aligned} \tag{S.2}$$

where we have used $g_p\left(\gamma_1,\ldots,\gamma_{p-1}\right)\leq \text{const}_p$, the fact that $\int_{-\pi}^{\pi}\exp\left\{i\left(\gamma\lfloor Tu\rfloor\right)\right\}=2\sin\left(\pi\lfloor Tu\rfloor\right)/\lfloor Tu\rfloor$, and the third equality follows from Lemma S.A.1-(ii).

Next, note that the function $H_{n_T}(\omega)$ will have substantial magnitude only for ω near some multiple of 2π. Thus, by continuity of $A(\cdot,\omega)$, g_p, and of the exponential function we yield

$$\begin{aligned}
&\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}A^{(a_1)}\left(u,\omega_1\right)\cdots A^{(a_p)}\left(u,\omega_p\right) \\
&\qquad\times H_{n_T}\left(\omega_1-\gamma_1\right)\cdots H_{n_T}\left(\omega_p-\gamma_p\right) \\
&\qquad\times\exp\left(i\left(\left(\omega_1+\cdots+\omega_p\right)\lfloor Tu\rfloor\right)\right)\eta\left(\sum_{j=1}^{p}\omega_j\right)g_p\left(\omega_1,\ldots,\omega_{p-1}\right)d\gamma_1\cdots d\gamma_p,
\end{aligned} \tag{S.3}$$

By Lemma P4.1 of Brillinger,

$$\begin{aligned}
&|\sum_{s=0}^{n_T-1}h_{a_1}\left(\frac{s+k_1}{n_T}\right)\cdots h_{a_{p-1}}\left(\frac{s+k_{p-1}}{n_T}\right)h_{a_p}\left(\frac{s}{n_T}\right)\exp\left(i\sum_{j=1}^{p}\omega_j t_p\right)-H_T^{(a_1,\ldots,a_p)}\left(\sum_{j=1}^{p}\omega_j\right)| \\
&\leq C\left(|k_1|+\cdots+|k_p|\right)
\end{aligned}$$

Thus, eq. (S.3) is equal to

$$\sum_{k_1=-n_T}^{n_T}\cdots\sum_{k_{p-1}=-n_T}^{n_T}\exp\left(i\sum_{j=1}^{p-1}\omega_j k_j\right)\left(\kappa_{Tu,t_p}\left(k_1\ldots,k_{p-1}\right)H_T^{(a_1,\ldots,a_p)}\left(\sum_{j=1}^{p}\omega_j\right)+O\left(n_T/T\right)\right) \\ +\varepsilon_T$$

with

$$|\varepsilon_T|\leq C\sum_{k_1=-n_T}^{n_T}\cdots\sum_{k_{p-1}=-n_T}^{n_T}\kappa_{Tu,t_p}\left(k_1\ldots,k_{p-1}\right)\left(|k_1|+\cdots+|k_p|\right)<\infty.$$

Therefore, it remains to show that

$$\sum_{k_1=-n_T}^{n_T} \cdots \sum_{k_{p-1}=-n_T}^{n_T} \exp\left(i \sum_{j=1}^{p-1} \omega_j k_j\right) \kappa_{Tu,t_p}\left(k_1 \ldots, k_{p-1}\right) = (2\pi)^{p-1} f^{(a_1,\ldots,a_p)}\left(u, \omega_1 \ldots, \omega_{p-1}\right).$$

We have

$$\begin{aligned}
&\sum_{k_1=-n_T}^{n_T} \cdots \sum_{k_{p-1}=-n_T}^{n_T} \exp\left(i \sum_{j=1}^{p-1} \omega_j k_j\right) \kappa_{Tu,t_p}\left(k_1 \ldots, k_{p-1}\right) \\
&\quad = \sum_{k_1=-n_T}^{n_T} \cdots \sum_{k_{p-1}=-n_T}^{n_T} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(i\left(\omega_1 + \cdots + \omega_{p-1} + \omega_p\right) t_p + \gamma_1 k_1 + \cdots + \gamma_{p-1} k_{p-1}\right) \\
&\quad\quad \times A\left(\lfloor Tu \rfloor, \omega_1\right) \cdots A\left(\lfloor Tu \rfloor, \omega_p\right) \eta\left(\sum_{j=1}^{p} \gamma_j\right) g_p\left(\omega_1, \ldots, \omega_{p-1}\right) d\gamma_1 \cdots d\gamma_p \\
&\quad\quad \sum_{k_1=-n_T}^{n_T} \cdots \sum_{k_{p-1}=-n_T}^{n_T} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(i\left(\omega_1 + \cdots + \omega_{p-1} + \omega_p\right) \lfloor Tu \rfloor + \gamma_1 k_1 + \cdots + \gamma_{p-1} k_{p-1}\right) \\
&\quad\quad \times A\left(\lfloor Tu \rfloor, \omega_1\right) \cdots A\left(\lfloor Tu \rfloor, \omega_p\right) \eta\left(\sum_{j=1}^{p} \gamma_j\right) g_p\left(\omega_1, \ldots, \omega_{p-1}\right) d\gamma_1 \cdots d\gamma_p.
\end{aligned}$$

Since $\sum_{j=1}^{p} \omega_j \equiv 0 \, (\mathrm{mod}\, 2\pi)$, ω_p is normalized and so the latter is equivalent to

$$\begin{aligned}
&\sum_{k_1=-n_T}^{n_T} \cdots \sum_{k_{p-1}=-n_T}^{n_T} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(i\left(\omega_1 + \cdots + \omega_{p-1} + \omega_p\right) \lfloor Tu \rfloor + \gamma_1 k_1 + \cdots + \gamma_{p-1} k_{p-1}\right) \\
&\quad A\left(\lfloor Tu \rfloor, \omega_1\right) \cdots A\left(\lfloor Tu \rfloor, \omega_p\right) g_p\left(\omega_1, \ldots, \omega_{p-1}\right) d\gamma_1 \cdots d\gamma_{p-1}.
\end{aligned}$$

Then, $A\left(\lfloor Tu \rfloor, \omega_1\right) \cdots A\left(\lfloor Tu \rfloor, \omega_p\right) g_p\left(\omega_1, \ldots, \omega_{p-1}\right) = f^{(a_1,\ldots,a_p)}\left(u, \omega_1, \ldots, \omega_{p-1}\right)$. In view of the following identities [see e.g., Exercise 1.7.5-(c,d) in Brillinger (1975)],

$$\sum_{k=-n_T}^{n_T} \exp\left(-i\omega k\right) = \frac{\sin\left(n_T + 1/2\right)\omega}{\sin \omega/2}, \qquad \int_{-\pi}^{\pi} \frac{\sin\left(n_T + 1/2\right)\omega}{\sin \omega/2} d\omega = 2\pi,$$

we yield

$$\begin{aligned}
&\mathrm{cum}\left(d_T^{(a_1)}\left(u, \omega_1\right), \ldots, d_T^{(a_p)}\left(u, \omega_p\right)\right) \\
&\quad = (2\pi)^{p-1} H_T^{(a_1,\ldots,a_p)}\left(\sum_{j=1}^{p} \omega_j\right) f^{(a_1,\ldots,a_p)}\left(u, \omega_1, \ldots, \omega_{p-1}\right) + O(1) + O\left(n_T/T\right),
\end{aligned}$$

which concludes the proof. □

S.A.2.2 Proof of Theorem 3.2

Proof of Theorem 3.2. We have,

$$\mathbb{E}\left(\mathbf{d}_{h,T}\left(u,\,\omega\right)\right)=\sum_{s=0}^{n_T-1}\exp\left(-i\omega s\right)\mathbb{E}\left(\mathbf{X}_{\lfloor Tu\rfloor-n_T/2+s+1,T}\right)$$
$$=0.$$

By Theorem 3.1,

$$n_T^{-1}\mathrm{Cov}\left(d_{h,T}^{(a_l)}\left(u,\,\pm\omega_j\right),\,d_{h,T}^{(a_r)}\left(u,\,\pm\omega_k\right)\right)$$
$$=n_T^{-1}2\pi H_{n_T}^{(a_l,\,a_r)}\left(\pm\omega_j\mp\omega_k\right)f^{(a_l,\,a_r)}\left(u,\,\pm\omega_j\left(n_T\right)\right)+o\left(1\right)+O\left(n_T^{-1}\right). \quad \text{(S.4)}$$

Note that [see e.g., Lemma P4.6 in Brillinger (1975)],

$$\left|H_{n_T}^{(a_1,\ldots,\,a_p)}\left(\omega\right)\right|\leq\frac{C}{\left|\sin\left(\omega/2\right)\right|}, \quad \text{(S.5)}$$

where C is a constant with $0<C<\infty$. In part (i) we have $\omega_j\pm\omega_k\not\equiv 0\,(\mathrm{mod}\,2\pi)$, thus the first term on the right-hand side of (S.4) tends to zero using (S.5). In part (ii) we have $\pm\omega_j\mp\omega_k\equiv 0\,(\mathrm{mod}\,2\pi)$, thus the right-hand side of (S.4) tends to

$$2\pi H_T^{(a_l,\,a_r)}\left(0\right)f^{(a_l,a_r)}\left(u,\,\pm\omega_j\right)=2\pi\left(\int h^{(a_l)}\left(t\right)h^{(a_r)}\left(t\right)dt\right)f^{(a_l,a_r)}\left(u,\,\pm\omega_j\right).$$

This shows that the second-order cumulants behave as indicated by the theorem. By Theorem 3.1 for $r>2$,

$$n_T^{-r/2}\mathrm{cum}\left(d_{h,T}^{(a_1)}\left(u,\,\pm\omega_{j_1}\right),\ldots,\,d_{h,T}^{(a_r)}\left(u,\,\pm\omega_{j_r}\right)\right)$$
$$=n_T^{-r/2}\left(2\pi\right)^{r-1}H_T^{(a_1,\ldots,\,a_r)}\left(\pm\omega_{j_1}\pm\cdots\pm\omega_{j_r}\right)f_{X^{(a_1)}\cdots X^{(a_r)}}\left(u,\,\pm\omega_{j_1},\ldots,\,\pm\omega_{j_{r-1}}\right)+o\left(n_T^{1-r/2}\right).$$

This last tends to 0 as $n_T\to\infty$ if $r>2$ because $H_T^{(a_1,\ldots,\,a_p)}\left(\omega\right)=O\left(T\right)$. Thus, also the cumulants of order higher than two behave as indicated by the theorem. This implies that the cumulants of the considered variables and the conjugates of those variables tend to the cumulants of Gaussian random variable. Since the distribution of the latter is fully determined by its moments, the theorem follows from Lemma S.A.4. □

S.A.2.3 Proof of Theorem 3.3

The proof of the second equality in (3.3) is similar to Dahlhaus (1996a) who proved the result under stronger assumptions on the data taper. Using the spectral representation (2.1),

$$\mathrm{cum}\left(d_{h,T}\left(u,\,\omega\right),\,d_{h,T}\left(u,\,-\omega\right)\right)$$
$$=\sum_{t=0}^{n_T-1}\sum_{s=0}^{n_T-1}h\left(\frac{t}{T}\right)h\left(\frac{s}{T}\right)\exp\left(-i\left(\omega-\eta\right)\left(s-t\right)\right)A^0_{\lfloor Tu\rfloor-n_T/2+t}\left(\eta\right)A^0_{\lfloor Tu\rfloor-n_T/2+s}\left(-\eta\right)d\eta.$$

We use Abel's transformation to replace $A^0_{\lfloor Tu\rfloor - n_T/2+t}(\eta)$ by $A(u,\omega)$,

$$\begin{aligned}
&\left|\sum_{t=0}^{n_T-1} h\left(\frac{t}{n_T}\right)\left(A^0_{\lfloor Tu\rfloor-n_T/2+t}(\eta)-A(u,\omega)\right)\exp\left(-i(\omega-\eta)t\right)\right| \\
&=\left|\sum_{t=0}^{n_T-1}\left(A^0_{\lfloor Tu\rfloor-n_T/2+t}(\eta)-A^0_{\lfloor Tu\rfloor-n_T/2+t-1}(\eta)\right)H_t\left(h\left(\frac{\cdot}{n_T},\omega-\eta\right)\right)\right| \\
&\quad+\left|\left(A^0_{\lfloor Tu\rfloor-n_T/2+n_T-1}(\eta)-A(u,\omega)\right)H_{n_T}\left(h\left(\frac{\cdot}{n_T}\right),\omega-\eta\right)\right| \\
&\leq O\left(\frac{n_T}{T}\right)L_{n_T}(\omega-\eta)+\left(O\left(\frac{n_T}{T}\right)+O\left(|\omega-\eta|\right)\right)L_{n_T}(\omega-\eta),
\end{aligned}$$

where the inequality follows from using Lemma S.A.2,

$$\left|H_t\left(h\left(\frac{\cdot}{n_T},\omega-\eta\right)\right)\right|\leq L_t(\omega-\eta)\leq L_{n_T}(\omega-\eta). \tag{S.6}$$

Since we are dividing by $\sum_{s=0}^{n_T-1}h(s/n_T)^2\sim n_T$ we get,

$$\begin{aligned}
&n_T^{-1}\left|\sum_{t=0}^{n_T-1}h\left(\frac{t}{n_T}\right)\left(A^0_{\lfloor Tu\rfloor-n_T/2+t}(\eta)-A\left(u+\frac{t-n_T/2}{T},\omega\right)\right)\exp\left(-i(\omega-\eta)t\right)\right| \\
&\leq O\left(\frac{1}{T}\right)L_{n_T}(\omega-\eta)+\left(O\left(\frac{1}{T}\right)+n_T^{-1}O\left(|\omega-\eta|\right)\right)L_{n_T}(\omega-\eta) \\
&\leq O\left(\frac{1}{T}\right)L_{n_T}(\omega-\eta)+\left(O\left(\frac{1}{T}\right)+n_T^{-1}O\left(|\omega-\eta|\right)\right)L_{n_T}(\omega-\eta) \\
&\leq C<\infty
\end{aligned}$$

where we have used the fact that $L_{n_T}(\omega-\eta)\leq n_T$ and that

$$|\omega-\eta|\,L_{n_T}(\omega-\eta)=\begin{cases}|\omega-\eta|\,n_T, & |\omega-\eta|\leq 1/n_T\\ |\omega-\eta|/|\omega-\eta|, & 1/n_T\leq|\omega-\eta|\leq\pi\end{cases}.$$

Using Lemma S.A.2 and (S.6) we have,

$$\begin{aligned}
&n_T^{-1}\left|\sum_{s=0}^{n_T-1}h\left(\frac{s}{T}\right)\exp\left(i(\omega-\eta)s\right)A^0_{\lfloor Tu\rfloor-n_T/2+s}(-\eta)\,d\eta\right| \\
&=n_T^{-1}A\left(\left(\lfloor Tu\rfloor-n_T/2\right)/T,-\eta\right)H_{n_T}(-\omega+\eta)+O\left(T^{-1}\right) \\
&=n_T^{-1}O\left(\sup_{u\leq n_T/T}A(u,-\eta)\right)L_{n_T}(-\omega+\eta)+O\left(T^{-1}\right).
\end{aligned}$$

Thus, after integration over η we yield that error in replacing $A^0_{\lfloor Tu\rfloor-n_T/2+t}(\eta)$ by $A(u,\omega)$ is $O((\log n_T)/n_T)$. Next, we replace $A^0_{\lfloor Tu\rfloor-n_T/2+s}(-\eta)$ by $A(u,\omega)$ and integrate over η using the relation

$$A(u,\omega)\,A(u,-\omega)=|A(u,\omega)|^2=f(u,\omega).$$

In view of

$$\int_{-\pi}^{\pi} |H_{n_T}(\alpha)|^2 d\alpha = 2\pi \sum_{t=0}^{n_T-1} \left(\frac{t}{n_T}\right)^2, \tag{S.7}$$

we then yield

$$\begin{aligned}\mathbb{E}(I_{h,T}(u,\omega)) &= \frac{1}{2\pi H_{2,n_T}(0)} \sum_{t=0}^{n_T-1}\sum_{s=0}^{n_T-1} h\left(\frac{t}{T}\right)\left(\frac{s}{T}\right) \exp(-i(\omega-\eta)(s-t)) f(u,\omega) + O\left(\frac{\log n_T}{n_T}\right) \qquad \text{(S.8)}\\ &= \frac{1}{\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2 d\alpha} \int_{-\pi}^{\pi} |H_{n_T}(\omega-\alpha)|^2 f(u,\alpha)\, d\alpha + O\left(\frac{\log n_T}{n_T}\right)\\ &= \frac{1}{\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2 d\alpha} \int_{-\pi}^{\pi} |H_{n_T}(\alpha)|^2 f(u,\omega-\alpha)\, d\alpha + O\left(\frac{\log n_T}{n_T}\right).\end{aligned}$$

This shows the first equality of (3.3). For the second equality of (3.3) replace $A^0_{\lfloor Tu\rfloor-n_T/2+t}(\eta)$ by $A\left(u+\frac{t-n_T/2}{T},\,\omega\right)$ and $A^0_{\lfloor Tu\rfloor-n_T/2+t}(-\eta)$ by $A\left(u+\frac{t-n_T/2}{T},\,-\omega\right)$ so that (S.8) holds with $f\left(u+\frac{t-n_T/2}{T},\,\omega\right)$ in place of $f(u,\omega)$. Then take a second-order Taylor expansion of f around around u to obtain

$$\begin{aligned}\mathbb{E}(I_{h,T}(u,\omega)) &= \frac{1}{2\pi H_{2,n_T}(0)} \sum_{t=0}^{n_T-1} h\left(\frac{t}{T}\right)^2 f\left(u+\frac{t-n_T/2}{T},\,\omega\right) + O\left(\frac{\log n_T}{n_T}\right)\\ &= f(u,\omega) + \frac{1}{2}\left(\frac{n_T}{T}\right)^2 \int_0^1 x^2 h^2(x+1/2)\, dx \frac{\partial^2}{\partial u^2} f(u,\omega) + o\left(\left(\frac{n_T}{T}\right)^2\right) + O\left(\frac{\log n_T}{n_T}\right). \square\end{aligned}$$

S.A.2.4 Proof of Theorem 3.4

By Theorem 2.3.1-(ix) in Brillinger (1975), $\text{Cov}(Y_j,\,Y_k) = \text{cum}\left(Y_j,\,\overline{Y}_k\right)$ for possibly complex variables Y_j and Y_k. Thus,

$$\begin{aligned}&\text{Cov}(d_{h,T}(u,\omega_j)\, d_{h,T}(u,-\omega_j),\, d_{h,T}(u,\omega_k)\, d_{h,T}(u,-\omega_k))\\ &\quad= \text{cum}\left(d_{h,T}(u,\omega_j)\, d_{h,T}(u,-\omega_j),\, \overline{d_{h,T}(u,\omega_k)\, d_{h,T}(u,-\omega_k)}\right)\\ &\quad= \text{cum}\left(d_{h,T}(u,\omega_j)\, d_{h,T}(u,-\omega_j),\, \overline{d_{h,T}(u,\omega_k) d_{h,T}(u,-\omega_k)}\right)\\ &\quad= \text{cum}(d_{h,T}(u,\omega_j)\, d_{h,T}(u,-\omega_j),\, d_{h,T}(u,-\omega_k)\, d_{h,T}(u,\omega_k))\\ &\quad= \text{cum}(d_{h,T}(u,\omega_j)\, d_{h,T}(u,-\omega_j),\, d_{h,T}(u,\omega_k)\, d_{h,T}(u,-\omega_k)).\end{aligned}$$

By the product theorem for cumulants [cf. Brillinger (1975), Theorem 2.3.2], we have to sum over all indecompasable partitions $\{P_1,\ldots,\,P_m\}$ with $|P_i| = \text{card}(P_i) \geq 2$ of the two-way table,

$$\begin{array}{c} a_{1,1}a_{1,2}\\ \vdots\,\vdots\\ a_{l,1}a_{l,2}\end{array}$$

where $a_{j,1}$ and $a_{j,2}$ stand for the position of $d_{h,T}(u,\omega_j)$ and $d_{h,T}(u,-\omega_j)$, respectively. This results in,

$$\text{cum}(d_{h,T}(u,\omega_j)\, d_{h,T}(u,-\omega_j),\, d_{h,T}(u,\omega_k)\, d_{h,T}(u,-\omega_k))$$

$$
\begin{aligned}
&= \operatorname{cum}\left(d_{h,T}\left(-\omega_j\right),\, d_{h,T}\left(-\omega_j\right),\, d_{h,T}\left(\omega_k\right),\, d_{h,T}\left(-\omega_k\right)\right) \\
&\quad + \operatorname{cum}\left(d_{h,T}\left(\omega_j\right)\right) \operatorname{cum}\left(d_{h,T}\left(-\omega_j\right),\, d_{h,T}\left(\omega_k\right),\, d_{h,T}\left(-\omega_k\right)\right) \\
&\quad + \text{three similar terms} \\
&\quad + \operatorname{cum}\left(d_{h,T}\left(\omega_j\right)\right) \operatorname{cum}\left(d_{h,T}\left(\omega_k\right)\right) \operatorname{cum}\left(d_{h,T}\left(-\omega_j\right),\, d_{h,T}\left(-\omega_k\right)\right) \\
&\quad + \text{three similar terms} \\
&\quad + \operatorname{cum}\left(d_{h,T}\left(\omega_j\right),\, d_{h,T}\left(-\omega_k\right)\right) \operatorname{cum}\left(d_{h,T}\left(-\omega_j\right),\, d_{h,T}\left(-\omega_k\right)\right) \\
&\quad + \operatorname{cum}\left(d_{h,T}\left(\omega_j\right),\, d_{h,T}\left(-\omega_k\right)\right) \operatorname{cum}\left(d_{h,T}\left(-\omega_j\right),\, d_{h,T}\left(\omega_k\right)\right).
\end{aligned}
$$

Then, by Theorem 3.1,

$$
\begin{aligned}
&\operatorname{cum}\left(d_{h,T}\left(u,\, \omega_j\right) d_{h,T}\left(u,\, -\omega_j\right),\, d_{h,T}\left(u,\, \omega_k\right) d_{h,T}\left(u,\, -\omega_k\right)\right) \\
&\quad = (2\pi)^3 H_{4,n_T}(0)\, f\left(u,\, \omega_j,\, -\omega_j,\, \omega_k\right) + O(1) \\
&\qquad + \left[2\pi H_{2,n_T}\left(\omega_j+\omega_k\right) f\left(u,\, \omega_j\right) + O(1)\right]\left[2\pi H_{2,n_T}\left(-\omega_j-\omega_k\right) f\left(u,\, \omega_j\right) + O(1)\right] \\
&\qquad + \left[2\pi H_{2,n_T}\left(\omega_j-\omega_k\right) f\left(u,\, \omega_j\right) + O(1)\right]\left[2\pi H_{2,n_T}\left(-\omega_j+\omega_k\right) f\left(u,\, \omega_j\right) + O(1)\right].
\end{aligned}
\tag{S.9}
$$

Given $H_{2,n_T}(0) = \sum_{t=0}^{n_T-1} h^2(t/T) \sim n_T \int h^2(\alpha)\, d\alpha$ and $H_{2,n_T}\left(\omega_j-\omega_k\right) H_{2,n_T}\left(-\omega_j+\omega_k\right) = \left|H_{2,n_T}\left(\omega_j-\omega_k\right)\right|^2$, the result of the theorem follows because

$$
n_T^{-2} (2\pi)^3 H_{4,n_T}(0)\, f\left(u,\, \omega_j,\, -\omega_j,\, \omega_k\right) = O\left(n_T^{-1}\right),
$$

and because the $O(1)$ terms on the right-hand side of (S.9) becomes negligible when multiplied by $H_{2,n_T}^{-2}(0)$.

Next, we prove that second result of the theorem. Recall that $\mathbf{z} \sim \mathscr{N}_p^{\mathrm{C}}\left(\mu_z,\, \Sigma_z\right)$ means that the $2p$ vector

$$
\begin{bmatrix} \operatorname{Re} \mathbf{z} \\ \operatorname{Im} \mathbf{z} \end{bmatrix}
$$

is distributed as

$$
\mathscr{N}_{2p}\left(\begin{bmatrix} \operatorname{Re} \mu_z \\ \operatorname{Im} \mu_z \end{bmatrix};\, \frac{1}{2}\begin{bmatrix} \operatorname{Re} \Sigma_z & -\operatorname{Im} \Sigma_z \\ -\operatorname{Im} \Sigma_z & \operatorname{Re} \Sigma_z \end{bmatrix}\right),
$$

where Σ_z is a $p \times p$ hermitian positive semidefinite matrix. By Theorem 3.2 we know that $\operatorname{Re} \mathbf{d}_{h,T}\left(\omega_j\right)$, $\operatorname{Im} \mathbf{d}_{h,T}\left(\omega_j\right)$ are asymptotically independent $\mathscr{N}\left(0,\, \pi n_T f\left(u,\, \omega_j\right)\right)$ variates. Hence, by the Mann-Wald Theorem,

$$
I_{h,T}\left(u,\, \omega_j\left(n_T\right)\right) = \left(2\pi n_T\right)^{-1}\left\{\left(\operatorname{Re} d_{h,T}\left(u,\, \omega_j\left(n_T\right)\right)\right)^2 + \left(\operatorname{Im} d_{h,T}\left(\omega_j\left(n_T\right)\right)\right)^2\right\},
$$

is asymptotically distributed as $f\left(u,\, \omega_j\right) \chi_2^2/2$ if $2\omega_j \not\equiv 0 \,(\operatorname{mod} 2\pi)$. In addition, if $\omega = \pm\pi,\, \pm 3\pi, \ldots$ then $I_{h,T}(u,\, \omega)$ is asymptotically distributed as $f\left(u,\, \omega_j\right) \chi_1^2$ independently from the previous variates. □

S.A.2.5 Proof of Theorem 3.5

Using Theorem 3.3,

$$\begin{aligned}\mathbb{E}\left(f_{h,T}\left(u,\,\omega\right)\right) &= \frac{2\pi}{n_T}\sum_{s=0}^{n_T-1} W_T\left(\omega-\frac{2\pi s}{n_T}\right)\mathbb{E}\left(I_{h,T}\left(u,\,\frac{2\pi s}{n_T}\right)\right)\\ &= \frac{2\pi}{n_T}\sum_{s=0}^{n_T-1} W_T\left(\omega-\frac{2\pi s}{n_T}\right)f\left(\frac{2\pi s}{T}\right)+O\left(\log\left(n_T\right)n_T^{-1}\right).\end{aligned}$$

Next, we approximate the sum appearing by an integral, then we see that the first term on the right-hand side is

$$\begin{aligned}&= \frac{2\pi}{n_T}\sum_{s=0}^{n_T-1} W_T\left(\omega-\frac{2\pi s}{n_T}\right)f\left(u,\,\frac{2\pi s}{n_T}\right)\\ &= \int_0^{2\pi} W_T\left(\omega-\alpha\right)f\left(u,\,\alpha\right)d\alpha+O\left(\left(n_Tb_T\right)^{-1}\right)\\ &= \int_0^{2\pi}\sum_{j=-\infty}^{\infty} b_T^{-1}W\left(b_T^{-1}\left(\omega-\alpha+2\pi j\right)\right)f\left(u,\,\alpha\right)d\alpha+O\left(\left(n_Tb_T\right)^{-1}\right)\\ &= \int_{-\infty}^{\infty} W\left(\beta\right)f\left(u,\,\omega-\beta b_T\right)d\beta+O\left(\left(n_Tb_T\right)^{-1}\right),\end{aligned}$$

where the last equality follows from the change in variable $\beta=b_T^{-1}\left(\omega-\alpha\right)$. This yields the first equality of (3.5). The second equality follows from the first and Theorem 3.3. □

S.A.2.6 Proof of Theorem 3.6

Let $c_T\left(u\right) = H_{2,T}\left(0\right)^{-1}\sum_{s=0}^{n_T-1} h\left(\frac{s+k}{T}\right)h\left(\frac{s}{T}\right)X_{\lfloor Tu\rfloor-n_T/2+s+k+1,T}X_{\lfloor Tu\rfloor-n_T/2+s+1,T}$. $I_{h,T}\left(u,\,\omega\right)$ can be written as $\left(2\pi\right)^{-1}\sum_{k=-\infty}^{\infty}\exp\left(-i\omega k\right)c_T\left(k\right)$. Note that

$$f_{h,T}\left(u,\,\omega\right) = \int_0^{2\pi} W_{2,T}\left(\omega-\alpha\right)I_{h,T}\left(u,\,\alpha\right)d\alpha,$$

where $W_{2,T}\left(\omega\right) = \sum_{k=-\infty}^{\infty} w\left(b_Tk\right)\exp\left(-i\omega k\right)$ and $w\left(k\right) = \int_{-\infty}^{\infty} W\left(\alpha\right)\exp\left(i\alpha k\right)d\alpha$ for $k\in\mathbb{R}$. From Theorem 3.3 we have $\mathbb{E}\left(I_{h,T}\left(u,\,\omega\right)\right) = f_{h,T}\left(u,\,\omega\right)+O\left(\left(n_T/T\right)^2\right)+O\left(\log\left(n_T\right)n_T^{-1}\right)$, and so

$$\begin{aligned}\mathbb{E}\left(f_{h,T}\left(u,\,\omega\right)\right) &= \int_0^{2\pi} W^{\left(n_T\right)}\left(\omega-\alpha\right)f\left(u,\,\alpha\right)d\alpha+O\left(\log\left(n_T\right)n_T^{-1}\right)\\ &= \int_0^{2\pi}\sum_{k=-\infty}^{\infty} w\left(b_Tk\right)\exp\left(-i\left(\omega-\alpha\right)k\right)f\left(u,\,\alpha\right)d\alpha\\ &\quad+O\left(\left(n_T/T\right)^2\right)+O\left(\log\left(n_T\right)n_T^{-1}\right)\\ &= \int_0^{2\pi}\sum_{k=-\infty}^{\infty} w\left(b_Tk\right)\exp\left(-i\left(\omega-\alpha\right)k\right)\sum_{s=-\infty}^{\infty}\exp\left(-i\alpha s\right)c\left(u,\,s\right)d\alpha\\ &\quad+O\left(\left(n_T/T\right)^2\right)+O\left(\log\left(n_T\right)n_T^{-1}\right)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=-\infty}^{\infty} w\left(b_T k\right) \exp\left(-i\omega k\right) c\left(u,\, k\right) d\alpha + O\left(\left(n_T/T\right)^2\right) + O\left(\log\left(n_T\right) n_T^{-1}\right) \\
&= \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} W\left(\alpha\right) \exp\left(i\alpha b_T k\right) \exp\left(-i\omega k\right) c\left(u,\, k\right) d\alpha \\
&\quad + O\left(\left(n_T/T\right)^2\right) + O\left(\log\left(n_T\right) n_T^{-1}\right) \\
&= \sum_{k=-\infty}^{\infty} b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega-\alpha\right)\right) \exp\left(-i b_T^{-1}\left(\omega-\alpha\right) b_T k\right) \exp\left(-i\omega k\right) c\left(u,\, k\right) d\alpha \\
&\quad + O\left(\left(n_T/T\right)^2\right) + O\left(\log\left(n_T\right) n_T^{-1}\right) \\
&= \sum_{k=-\infty}^{\infty} b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega-\alpha\right)\right) \exp\left(-i\left(\omega-\alpha\right) k\right) \exp\left(-i\omega k\right) c\left(u,\, k\right) d\alpha \\
&\quad + O\left(\left(n_T/T\right)^2\right) + O\left(\log\left(n_T\right) n_T^{-1}\right) \\
&= \sum_{k=-\infty}^{\infty} b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega-\alpha\right)\right) \exp\left(-i\omega k\right) c\left(u,\, k\right) d\alpha \\
&\quad + O\left(\left(n_T/T\right)^2\right) + O\left(\log\left(n_T\right) n_T^{-1}\right) \\
&= b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega-\alpha\right)\right) f\left(u,\, \omega\right) d\alpha + O\left(\left(n_T/T\right)^2\right) + O\left(\log\left(n_T\right) n_T^{-1}\right) \\
&= \int_{-\infty}^{\infty} W\left(\alpha\right) f\left(u,\, \omega - b_T\alpha\right) d\alpha + O\left(\left(n_T/T\right)^2\right) + O\left(\log\left(n_T\right) n_T^{-1}\right).
\end{aligned}
$$

Thus, $\mathbb{E}\left(f_{h,T}\left(u,\, \omega\right)\right) \to f\left(u,\, \omega\right)$. Next, from Theorem 3.4,

$$
\begin{aligned}
&\operatorname{Cov}\left(f_{h,T}\left(u,\, \omega_j\right),\, f_{h,T}\left(u,\, \omega_k\right)\right) \\
&\quad = \int_0^{2\pi}\int_0^{2\pi} W_T\left(\omega_j-\alpha\right) W_T\left(\omega_k-\beta\right) \operatorname{Cov}\left(I_{h,T}\left(u,\, \alpha\right),\, I_{h,T}\left(u,\, \beta\right)\right) d\alpha d\beta \\
&\quad = H_{2,n_T}\left(0\right)^{-1} H_{2,n_T}\left(0\right)^{-1} \int_0^{2\pi}\int_0^{2\pi} W_T\left(\omega_j-\alpha\right) W_T\left(\omega_k-\beta\right) \\
&\qquad \times \left\{\left|H_{2,n_T}\left(\alpha-\beta\right)\right|^2 + \left|H_{2,n_T}\left(\alpha+\beta\right)\right|^2\right\} \left|f\left(u,\, \alpha\right)\right|^2 d\alpha d\beta + O\left(n_T^{-1}\right).
\end{aligned}
$$

We now show that

$$
\begin{aligned}
&\int_0^{2\pi} W_T\left(\omega_k-\beta\right) \left|H_{2,n_T}\left(\alpha-\beta\right)\right|^2 d\beta \\
&\qquad\qquad = 2\pi W_T\left(\omega_k-\alpha\right) \sum_{s=0}^{n_T-1} h^4\left(s\right) + O\left(b_T^{-2}\right),
\end{aligned}
\tag{S.10}
$$

uniformly in α. As

$$
H_{2,n_T}\left(\omega\right) = \sum_{s=0}^{n_T-1} h\left(s/n_T\right)^2 \exp\left(-i\omega s\right),
$$

we may write (S.10) as

$$\begin{aligned}
&\sum_{t=0}^{n_T-1}\sum_{s=0}^{n_T-1} h^2\left(t/n_T\right) h^2\left(s/n_T\right) \int_0^{2\pi} W_T\left(\omega_k-\beta\right) \\
&\qquad \times \exp\left\{-i\left(\alpha-\beta\right)t+i\left(\alpha-\beta\right)s\right\} d\beta \\
&= \sum_{t=0}^{n_T-1}\sum_{s=0}^{n_T-1} h^2\left(t\right) h^2\left(s\right) \int_0^{2\pi} \sum_{k=-\infty}^{\infty} w\left(b_T k\right) \exp\left(-i\left(\omega_k-\beta\right)k\right) \\
&\qquad \times \exp\left\{-i\left(\alpha-\beta\right)t+i\left(\alpha-\beta\right)s\right\} d\beta \\
&= \sum_{t=0}^{n_T-1}\sum_{s=0}^{n_T-1} h^2\left(t/n_T\right) h^2\left(s/n_T\right) w\left(b_T\left(t-s\right)\right) \exp\left(i\left(\omega_k-\alpha\right)\left(t-s\right)\right) \\
&\qquad = \sum_{k=-\infty}^{\infty} w\left(b_T k\right) \exp\left(i\left(\omega_k-\alpha\right)k\right) \sum_{s=0}^{n_T-1} h^2\left(\left(s+k\right)n_T\right) h^2\left(s/n_T\right) \\
&\qquad = 2\pi W_T\left(\omega_k-\alpha\right) \sum_{s=0}^{n_T-1} h^4\left(s/n_T\right) + R_T.
\end{aligned}$$

where we have applied Lemma S.A.3 to $\exp\left(i\left(\omega_k-\alpha\right)k\right)\sum_{s=0}^{n_T-1} h^2\left(s+k\right)h^2\left(s\right)$ to yield,

$$\left|\exp\left(i\left(\omega_k-\alpha\right)k\right)\sum_{s=0}^{n_T-1} h^2\left(s+k\right)h^2\left(s\right) - \exp\left(i\left(\omega_k-\alpha\right)k\right)\sum_{s=0}^{n_T-1} h^4\left(s/n_T\right)\right| \leq C\left|k\right|,$$

so that for $0<C<\infty$,

$$\left|R_T\right| \leq C\sum_{k=-\infty}^{\infty}\left|w\left(b_T k\right)\right|\left|k\right| \sim H b_T^{-2}\int\left|x\right|\left|w\left(x\right)\right|dx.$$

The latter result follows because

$$\begin{aligned}
C\sum_{k=-\infty}^{\infty}\left|w\left(b_T k\right)\right|\left|k\right| &= C b_T^{-1}\sum_{k=-\infty}^{\infty}\left|w\left(b_T k\right)\right|\left|b_T k\right| \\
&= C b_T^{-2} b_T\sum_{k=-\infty}^{\infty}\left|w\left(b_T k\right)\right|\left|b_T k\right| = C b_T^{-2}\int\left|x\right|\left|w\left(x\right)\right|dx
\end{aligned}$$

for a finite $0<C<\infty$. A similar result holds for the second term involving $\left|H_{2,n_T}\left(\alpha+\beta\right)\right|^2$. The covariance being evaluated thus has the form,

$$\begin{aligned}
&\mathrm{Cov}\left(f_{h,T}\left(u,\,\omega_j\right),\, f_{h,T}\left(u,\,\omega_k\right)\right) \\
&\quad = 2\pi H_{2,T}\left(0\right)^{-1} H_{2,T}\left(0\right)^{-1}\sum_{s=0}^{n_T-1} h\left(s/n_T\right)^4 \\
&\qquad \int_0^{2\pi}\left\{W_T\left(\omega_j-\alpha\right)W_T\left(\omega_k-\alpha\right)\left|f\left(u,\,\alpha\right)\right|^2\right. \\
&\quad \left.+ W_T\left(\omega_j-\alpha\right)W_T\left(\omega_k-\alpha\right)\left|f\left(u,\,\alpha\right)\right|^2\right\}d\alpha + O\left(n_T^{-1}\right),
\end{aligned}$$

Equation (3.6) follows from

$$
\begin{aligned}
&\lim_{T\to\infty} n_T b_T \mathrm{Cov}\left(f_{h,T}\left(u,\,\omega_j\right),\, f_{h,T}\left(u,\,\omega_k\right)\right) \\
&\qquad = b_T n_T 2\pi n_T^{-1} n_T H_{2,T}\left(0\right)^{-1} n_T^{-1} n_T H_{2,T}\left(0\right)^{-1} n_T n_T^{-1} \sum_{s=0}^{n_T-1} h\left(s/n_T\right)^4 \\
&\qquad\quad \int_0^{2\pi} \{W_T\left(\omega_j-\alpha\right) W_T\left(\omega_k-\alpha\right)\left|f\left(u,\,\alpha\right)\right|^2 \\
&\qquad\quad + W_T\left(\omega_j-\alpha\right) W_T\left(\omega_k+\alpha\right)\left|f\left(u,\,\alpha\right)\right|^2\} d\alpha + O\left(\left(n_T b_T\right)^{-1}\right) + O\left(b_T\right), \\
&\qquad = b_T 2\pi n_T H_{2,T}\left(0\right)^{-1} n_T H_{2,T}\left(0\right)^{-1} n_T^{-1} \sum_{t=0}^{n_T-1} h\left(t/n_T\right)^4 \\
&\qquad\quad \int_0^{2\pi} \{ \sum_{l=-\infty}^{\infty} b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega_j-\alpha+2\pi l\right)\right) \\
&\qquad\quad \times \sum_{l=-\infty}^{\infty} b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega_k-\alpha+2\pi l\right)\right)\left|f\left(u,\,\alpha\right)\right|^2 \\
&\qquad\quad + \sum_{l=-\infty}^{\infty} b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega_j-\alpha+2\pi l\right)\right) \\
&\qquad\quad \times \sum_{l=-\infty}^{\infty} b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega_k+\alpha+2\pi l\right)\right)\left|f\left(u,\,\alpha\right)\right|^2\} d\alpha + O\left(\left(n_T b_T\right)^{-1}\right) + O\left(b_T\right) \\
&\qquad = 2\pi n_T H_{2,T}\left(0\right)^{-1} n_T H_{2,T}\left(0\right)^{-1} \int h^4\left(t\right) dt \\
&\qquad\quad \int_0^{2\pi} \{b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega_j-\alpha\right)\right) \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega_k-\alpha\right)\right)\left|f\left(u,\,\alpha\right)\right|^2 \\
&\qquad\quad + b_T^{-1} \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega_j-\alpha\right)\right) \int_{-\infty}^{\infty} W\left(b_T^{-1}\left(\omega_k+\alpha\right)\right)\left|f\left(u,\,\alpha\right)\right|^2\} d\alpha \\
&\qquad\quad + O\left(\left(n_T b_T\right)^{-1}\right) + O\left(b_T\right) \\
&\qquad = 2\pi \left(\int h^2\left(t\right) dt\right)^{-2} \int h^4\left(t\right) dt \\
&\qquad\quad \int_0^{2\pi} \left[\eta\left\{\omega_j-\omega_k\right\}\left|f\left(u,\,\omega_j\right)\right|^2 + \eta\left\{\omega_j+\omega_k\right\}\left|f\left(u,\,\omega_j\right)\right|^2\right] \int_{-\infty}^{\infty} W^2\left(\alpha\right) d\alpha \\
&\qquad\quad + O\left(\left(n_T b_T\right)^{-1}\right) + O\left(b_T\right).
\end{aligned}
$$

Finally, we consider the magnitude of the joint cumulants of order L. We have

$$
\begin{aligned}
&\mathrm{cum}\left(f_{h,T}\left(u,\,\omega_1\right),\ldots,\, f_{h,T}\left(u,\,\omega_L\right)\right) \\
&\quad = 2\pi \left\{H_{2,n_T}\left(0\right)\right\}^{-L} \\
&\quad \times \sum_{t_1=0}^{n_T-1} \cdots \sum_{2L=0}^{n_T-1} w\left(b_T\left(t_1-t_2\right)\right) \cdots w\left(b_T\left(t_{2L-1}-t_{2L}\right)\right) \\
&\quad \times \exp\left(-i\omega_1\left(t_1-t_2\right) - \ldots - i\omega_L\left(t_{2L-1}-t_{2L}\right)\right) h_{n_T}\left(t_1\right) \cdots h_{n_T}\left(t_{2L}\right)
\end{aligned}
\tag{S.11}
$$

$$\times \operatorname{cum}\left(X_{\lfloor Tu\rfloor-n_T/2+t_1+1,T}X_{\lfloor Tu\rfloor-n_T/2+t_2+1,T},\ldots,\,X_{\lfloor Tu\rfloor-n_T/2+t_{2L-1}+1,T}X_{\lfloor Tu\rfloor-n_T/2+t_{2L}+1,T}\right).$$

Note that

$$\begin{aligned}&\operatorname{cum}\left(X_{\lfloor Tu\rfloor-n_T/2+t_1+1,T}X_{\lfloor Tu\rfloor-n_T/2+t_2+1,T},\ldots,\,X_{\lfloor Tu\rfloor-n_T/2+t_{2L-1}+1,T}X_{\lfloor Tu\rfloor-n_T/2+t_{2L}+1,T}\right)\\&\quad=\sum_{\mathbf{v}}\left|c_{X\cdots X}\left(u;\,t_j;\,j\in v_1\right)\cdots c_{X\cdots X}\left(u;\,t_j;\,j\in v_p\right)\right|,\end{aligned}$$

where the summation is over all indecomposable partitions $\mathbf{v}=(v_1,\ldots,\,v_{\mathrm{p}})$ of the table

$$\begin{matrix}1 & 2\\ 3 & 4\\ \vdots & \vdots\\ 2L-1 & 2L.\end{matrix}$$

As the partition is indecomposable, in each set v_p of the partition we may find an element t_p^*, so that none of $t_j-t_p^*$, $j\in v_{\mathrm{p}}$, $\mathrm{p}=1,\ldots,\,\mathrm{P}$ is $t_{2l-1}-t_{2l}$, $l=1,\,2,\ldots,\,L$. Define $2L-\mathrm{P}$ new variables $k_1,\ldots,\,k_{2L-\mathrm{P}}$ as the nonzero $t_j-t_{\mathrm{p}}^*$. Since for each v_{p} we get one t_{p}^* we have $2L-\mathrm{P}$ variables. In the case $L=1$ we have $t_1,\,t_2$ and get $u=t_1-t_2$. The cumulant (*) is now bounded by

$$\begin{aligned}M^Ln_T^{-L}\sum_{\mathbf{v}}\sum_{t_1^*}\cdots\sum_{t_{\mathrm{P}}^*}\sum_{k_1}\cdots\sum_{k_{2L-\mathrm{P}}}&\left|w\left(b_T\left(k_{\alpha_1}+t_{\beta_1}^*-k_{\alpha_1}-t_{\beta_2}^*\right)\right)\right.\\&\cdots w\left(b_T\left(k_{\alpha_{2L-1}}+t_{\beta_{2L-1}}^*-k_{\alpha_{2L}}-t_{\beta_{2L}}^*\right)\right)\\&\left|h\left(t_1^*/n_T\right)\right|^{2L}\left|c_{X\cdots X}\left(u;\,k_1,\ldots\right)\cdots c_{X\cdots X}\left(u;\,\ldots,\,k_{2L-\mathrm{P}}\right)\right|,\end{aligned}$$

for some finite M where $\alpha_1,\ldots,\,\alpha_{2L}$ are selected from $1,\ldots,\,2L$ and $\beta_1,\ldots,\,\beta_{2L}$ from $1,\ldots,\,\mathrm{P}$. Note that n_T^{-L} comes from the H_T's in the denominator. Defining $\phi\left(t_j\right)=t_p^*$, $j\in v_1$, we apply Lemma 2.3.1 in Brillinger (1975) to see that there are $\mathrm{P}-1$ linearly independent differences among the $t_{\beta_1^*}-t_{\beta_2^*},\ldots,\,t_{\beta_{2L-1}^*}-t_{\beta_{2L}^*}$. For convenience suppose these are $t_{\beta_1^*}-t_{\beta_2^*},\ldots,\,t_{\beta_{2\mathrm{P}-2}^*}-t_{\beta_{2\mathrm{P}-1}^*}$. Making a final change of variables

$$\begin{aligned}v_1&=u_{\alpha_1}+t_{\beta_1}^*-u_{\alpha_1}-t_{\beta_2}^*\\&\vdots\\v_{\mathrm{P}-1}&=u_{\alpha_{2\mathrm{P}-3}}+t_{\beta_{2\mathrm{P}-3}}^*-u_{\alpha_{2\mathrm{P}-2}}-t_{\beta_{2\mathrm{P}-2}}^*,\end{aligned}$$

we see that the cumulant (*) is bounded by

$$\begin{aligned}M^Ln_T^{-L}\sum_{\mathbf{v}}\sum_{t_1^*}\sum_{v_1}\cdots\sum_{v_{\mathrm{P}-1}}\sum_{u_1}\cdots\sum_{u_{2L-\mathrm{P}}}&\left|w\left(b_Tv_1\right)\cdots w\left(b_Tv_{P-1}\right)\right|\\&\left|h^{(T)}\left(t_1^*\right)\right|^{2L}\left|c_{X\cdots X}\left(u;\,k_1,\ldots\right)\cdots c_{X\cdots X}\left(u;\,\ldots,\,k_{2L-\mathrm{P}}\right)\right|,\\&\leq M^Ln_T^{-L+1}b_T^{-(\mathrm{P}-1)}\sum_{\mathbf{v}}C_{n_{2,1}}\left(u\right)\cdots C_{n_{2,\mathrm{P}}}\left(u\right)\\&=O\left(n_T^{-L+1}b_T^{-(\mathrm{P}-1)}\right),\qquad\text{as}\,\mathrm{P}\leq L,\end{aligned}$$

where $C_n = \sum_{t_1,\ldots,t_{n-1}} |c_{X\cdots X}(u;\, t_1, \ldots, t_{n-1})|$ and $n_{2,j}$ denotes the number of elements in the jth set of the partition $\mathbf{v}$. We see that the standardized joint cumulant,

$$\text{cum}\left\{(n_T b_T)^{1/2} f_{h,T}(u,\, \omega_1), \ldots, (n_T b_T)^{1/2} f_{h,T}(u,\, \omega_L)\right\},$$

for $L > 2$, tends to 0 as $T \to \infty$. This means that the variates $f_{h,T}(u,\, \omega_1), \ldots, f_{h,T}(u,\, \omega_L)$ are asymptotically normal with the moment structure of the theorem. □

S.A.3 Proof of the Results of Section 4

S.A.3.1 Proof of Theorem 4.1

Without loss of generality, we assume that we observe $\{X_t\}$ for $t \in \mathbb{Z}$. For $\omega \in [-\pi,\, \pi]$ let $S_{r+1}(\omega) = \sum_{j \in \{\{\mathbf{S}_r,\, r=1,\ldots,r+1\}\}} I^*_{h,T}(j/T,\, \omega)$ and

$$R_{r,T}(\omega) = \frac{1}{M_{S,T}} \left(S_{r+1}(\omega) - \sum_{j \in \{\{\mathbf{S}_r,\, r=1,\ldots,r+1\}\}} \mathscr{W}_j(\omega) - \left(S_{rm_T+m_T/2} - \sum_{j=1}^{rm_T+m_T/2} \mathscr{W}_j(\omega) \right) \right),$$

where $\mathscr{W}_j(\omega) = \sigma_j(\omega) Z_j$ with $Z_j \sim i.i.d.\, \mathscr{N}(0,\, 1)$. Write

$$\begin{aligned}
\widetilde{I}_{r,T}(\omega) &= \frac{1}{m_T} \sum_{j=(r+1)m_T-m_T/2+1}^{(r+1)m_T+m_T/2} I_{h,T}(j/T,\, \omega) \qquad (S.12)\\
&= \frac{1}{m_T} \sum_{j=(r+1)m_T-m_T/2+1}^{(r+1)m_T+m_T/2} \left(I^*_{h,T}(j/T,\, \omega) + \mathbb{E}(I_{h,T}(j/T,\, \omega)) \right)\\
&= \frac{1}{m_T} \left(\sum_{j=(r+1)m_T-m_T/2+1}^{(r+1)m_T+m_T/2} \mathscr{W}_j(\omega) \right) + R_{r,T} + \frac{1}{m_T} \sum_{j=(r+1)m_T-m_T/2+1}^{(r+1)m_T+m_T/2} \mathbb{E}(I_{h,T}(j/T,\, \omega)).
\end{aligned}$$

Under Assumption 4.1-(i), Theorem 1 in Wu and Zhou (2011) yields $\max_{0 \le r \le M_{S,T}-1} |R_{r,T}| = O_{\mathbb{P}}(\tau_T / M_{S,T})$. The same bound holds under Assumption 4.1-(ii) by Corollary 1 in Wu and Zhou (2011). By Theorem 3.3, $\mathbb{E}(I_{h,T}(j/T,\, \omega)) = f(j/T,\, \omega) + O((n_T/T)^2) + O(\log(n_T)/n_T)$. By Assumption 2.1, we have,

$$f(((r+1)m_T + j)/T,\, \omega) - f((rm_T + j)/T,\, \omega) = O\left((m_T/T)^\theta\right), \qquad \text{uniformly in } r \text{ and } j. \qquad (S.13)$$

Altogether, we yield

$$\begin{aligned}
\sqrt{m_T}\left(\widetilde{I}_{r+1,T}(\omega) - \widetilde{I}_{r,T}(\omega)\right) &= \frac{1}{\sqrt{m_T}} \left(\left(\sum_{j=(r+1)m_T-m_T/2+1}^{(r+1)m_T+m_T/2} \mathscr{W}_j(\omega) \right) - \left(\sum_{j=rm_T-m_T/2+1}^{rm_T+m_T/2} \mathscr{W}_j(\omega) \right) \right) \qquad (S.14)\\
&\quad + O\left(M_{S,T}^{1/2} m_T^\theta / T^\theta\right) + O_{\mathbb{P}}\left(\tau_T / M_{S,T}^{1/2}\right) + O_{\mathbb{P}}\left(M_{S,T}^{1/2} (n_T/T)^2 + M_{S,T}^{1/2} \log(n_T)/n_T\right)\\
&= \frac{1}{\sqrt{m_T}} \left(\left(\sum_{j=(r+1)m_T-m_T/2+1}^{(r+1)m_T+m_T/2} \mathscr{W}_j(\omega) \right) - \left(\sum_{j=rm_T-m_T/2+1}^{rm_T+m_T/2} \mathscr{W}_j(\omega) \right) \right)\\
&\quad + o_{\mathbb{P}}\left((\log M_T)^{-1/2}\right).
\end{aligned}$$

Theorem 4.1 then follows from Lemma 1 in Wu and Zhao (2007). □

S.A.3.2 Proof of Theorem 4.2

Lemma S.A.5. *Let $\mathscr{V}(\omega)$ denote a random variable defined by $\mathbb{P}(\mathscr{V}(\omega)\leq v)=\exp(-\pi^{-1/2}\exp(-v))$ for $\omega\in\Pi$. Assume that for $\omega,\,\omega'\in\Pi$ the variables $\mathscr{V}(\omega)$ and $\mathscr{V}(\omega')$ are independent. Let $\mathscr{V}^{*}\triangleq\max_{\omega\in\Pi}\mathscr{V}(\omega)-\log(n_{\omega})$. Then, $\mathbb{P}(\mathscr{V}^{*}\leq v)=\exp\left(-\pi^{-1/2}\exp(-v)\right)$.*

Proof. Since $\mathscr{V}(\omega)$ is independent from any $\mathscr{V}(\omega')$ with $\omega\neq\omega'$, we have

$$\begin{aligned}\log\mathbb{P}(\mathscr{V}^{*}\leq v)&=\sum_{j=1}^{n_{\omega}}\log\mathbb{P}(\mathscr{V}(\omega_{j})\leq(\log(n_{\omega})+v))\\&=\sum_{j=1}^{n_{\omega}}\left(-\pi^{-1/2}\exp\left(\log\left(n_{\omega}^{-1}\right)\right)\exp(-v)\right)\\&=-\pi^{-1/2}\exp(-v).\end{aligned}$$

Thus, $\mathbb{P}(\mathscr{V}^{*}\leq v)=\exp\left(-\pi^{-1/2}\exp(-v)\right)$. □

Proof of Theorem 4.2. From Theorem 3.4 it follows that $I_{h,T}(u,\,\omega_{j})$ and $I_{h,T}(u,\,\omega_{k})$ are asymptotically independent if $2\omega_{j},\,\omega_{k}\pm\omega_{k}\not\equiv 0\,(\mathrm{mod}\,2\pi)$, $1\leq j<k\leq n_{\omega}$. The result then follows from Lemma S.A.5. □

S.A.3.3 Proof of Theorem 4.3

Proof of Theorem 4.3. For $\omega\in[-\pi,\,\pi]$ let $S_{(r+1)m_{T}}(\omega)=\sum_{j=1}^{(r+1)m_{T}}f_{h,T}(j/T,\,\omega)$ and $R_{r,T}(\omega)$ be defined as in the proof of Theorem 4.1. Write

$$\begin{aligned}\widetilde{f}_{r,T}(\omega)&=\frac{1}{m_{T}}\sum_{j=(r+1)m_{T}-m_{T}/2+1}^{(r+1)m_{T}+m_{T}/2}f_{h,T}(j/T,\,\omega)\\&=\frac{1}{m_{T}}\sum_{j=(r+1)m_{T}-m_{T}/2+1}^{(r+1)m_{T}+m_{T}/2}\left(f_{h,T}^{*}(j/T,\,\omega)+\mathbb{E}(f_{h,T}(j/T,\,\omega))\right)\\&=\frac{1}{m_{T}}\left(\sum_{j=(r+1)m_{T}-m_{T}/2+1}^{(r+1)m_{T}+m_{T}/2}\mathscr{W}_{j}(\omega)\right)+R_{r,T}+\frac{1}{m_{T}}\sum_{j=(r+1)m_{T}-m_{T}/2+1}^{(r+1)m_{T}+m_{T}/2}\mathbb{E}(f_{h,T}(j/T,\,\omega)).\end{aligned}\tag{S.15}$$

As in the proof of Theorem 4.1 $\max_{0\leq r\leq M_{T}-1}|R_{r,T}|=O_{\mathbb{P}}(\tau_{T}/m_{T})$. By Theorem 3.5, $\mathbb{E}(f_{h,T}(j/T,\,\omega))=f(j/T,\,\omega)+O\left((n_{T}/T)^{2}\right)+O\left(b_{W,T}^{2}\right)+O(\log(n_{T})/n_{T})$. Note that eq. (S.13) continues to hold. Thus, we yield

$$\begin{aligned}&\sqrt{m_{T}}\left(\widetilde{f}_{r+1,T}(\omega)-\widetilde{f}_{r,T}(\omega)\right)\\&=\frac{1}{\sqrt{m_{T}}}\left(\left(\sum_{j=(r+1)m_{T}-m_{T}/2+1}^{(r+1)m_{T}+m_{T}/2}\mathscr{W}_{j}(\omega)\right)-\left(\sum_{j=rm_{T}-m_{T}/2+1}^{rm_{T}+m_{T}/2}\mathscr{W}_{j}(\omega)\right)\right)\\&\quad+O\left(m_{T}^{\theta+1/2}/T^{\theta}\right)+O_{\mathbb{P}}\left(\tau_{T}/m_{T}^{1/2}\right)+O_{\mathbb{P}}\left(m_{T}^{1/2}(n_{T}/T)^{2}+m_{T}^{1/2}b_{W,T}^{2}+m_{T}^{1/2}\log(n_{T})/n_{T}\right)\end{aligned}\tag{S.16}$$

$$= \frac{1}{\sqrt{m_T}} \left(\left(\sum_{j=(r+1)m_T - m_T/2+1}^{(r+1)m_T + m_T/2} \mathscr{W}_j(\omega) \right) - \left(\sum_{j=rm_T - m_T/2+1}^{rm_T + m_T/2} \mathscr{W}_j(\omega) \right) \right)$$
$$+ o_{\mathbb{P}} \left((\log M_T)^{-1/2} \right).$$

The result then follows from combining Lemma 1 in Wu and Zhao (2007), Lemma S.A.5 and the same argument as in Theorem 4.2. □

S.A.4 Proofs of the Results in Section 5

For a sequence of random variables $\{\xi_j\}$, let $\mathbb{P}_{\{\xi_j\}}$ denote the law of the observations $\{\xi_j\}$. Let $||\mathbb{P}_{\{\xi_j\}} - \mathbb{P}_{\{\xi_j^*\}}||_{\mathrm{TV}}$ define the total variation distance between the probability measures $\mathbb{P}_{\{\xi_j\}}$ and $\mathbb{P}_{\{\xi_j^*\}}$. For two random variables Y and X with distributions $\mathbb{P}_Y$ and $\mathbb{P}_X$, respectively, denote the Kullback-Leibler divergence by $D_{\mathrm{KL}}(Y||\,X) = D_{\mathrm{KL}}(\mathbb{P}_Y||\,\mathbb{P}_X) = \int \log(d\mathbb{P}_Y/d\mathbb{P}_X)\, d\mathbb{P}_Y$.

S.A.4.1 Proof of Theorem 5.1

The proof is based on several steps of information-theoretic reductions that allow us to show the asymptotic equivalence in the strong Le Cam sense of our statistical problem to a special high-dimensional signal detection problem. The minimax lower bound is then obtained by using classical arguments as in Ingster and Suslina (2003). Information-theoretic reductions were also used by Bibinger, Jirak, and Vetter (2017) to establish a minimax lower bound for change-point testing in volatility in the context of high-frequency data. Our derivations differ from theirs in several ways because we deal with serially correlated observations while they had independent observations. Furthermore, our testing problem is more complex because our observations have an unknown distribution while their observations are squared of standard normal variables.

Define

$$d_{L,h,T}(u,\,\omega) \triangleq \sum_{s=0}^{n_T-1} h\left(\frac{s}{n_T}\right) X_{\lfloor Tu\rfloor - n_T+s+1,T} \exp(-i\omega s),$$
$$I_{L,h,T}(u,\,\omega) \triangleq \frac{1}{2\pi H_{2,n_T}(0)} |d_{L,h,T}(u,\,\omega)|^2. \tag{S.17}$$

Note that the difference between $d_{L,h,T}(u,\,\omega)$ and $d_{h,T}(u,\,\omega)$ as defined in Section 3 is that the former uses only observations to the left of $\lfloor Tu\rfloor$. We first consider alternatives as in $\mathcal{H}_1$. Throughout the proof we set

$$m_T = C_T \left(\sqrt{\log(M_T)} T^{\theta}/D \right)^{\frac{2}{2\theta+1}}, \tag{S.18}$$

with a constant $C_T > 0$. We begin by granting the experimenter additional knowledge thereby focusing on a simpler sub-model. This additional knowledge can only decrease the lower bound on minimax distinguishability and therefore such lower bound carries over to the original model. We restrict attention to a sub-class of $\boldsymbol{F}_{1,\lambda_b^0,\omega_0}(\theta,\, b_T,\, D)$ which is characterized by a break at time $\lambda_b^0 \in (0,\,1)$ with $|f(\lambda_b^0,\,\omega_0) - f(\lambda_b^0+,\,\omega_0)| \geq b_T$, where $f(\lambda_b^0+,\,\omega) = \lim_{s\downarrow\lambda_b^0} f(s,\,\omega)$. We further assume that the break point is an integer multiple of m_T, i.e., $T\lambda_b^0 m_T^{-1} \in \{1,\,2,\ldots,\,\lfloor T/m_T\rfloor - 1\}$.

In order to simplify the proof, we consider a simplified version of the problem following Bibinger, Jirak, and Vetter (2017). We set $f_{-}(\omega_0) = 1$ and let

$$f(j/T, \omega_0) = \begin{cases} 1 + (m_T - j \bmod m_T)^{\theta} T^{-\theta}, & T\lambda_b^0 < j \leq T\lambda_b^0 + m_T \\ 1, & \text{else} \end{cases}. \tag{S.19}$$

We discuss the general case $f_{-}(\omega_0) \neq 1$ at the end of this proof. Eq. (S.19) specifies that the spectrum at frequency ω_0 exhibits a break of order b_T at λ_b^0 and then decays on the interval $(\lambda_b^0, \lambda_b^0 + T^{-1}m_T]$ smoothly with regularity θ and is constant elsewhere. Name this sub-class $\boldsymbol{F}^{+}_{\lambda_b^0, \omega_0}$. Note that here the location of λ_b^0 is still unknown. To establish the lower bound, it suffices to focus on the sub-class of the above form.

Next, we introduce a stepwise approximation to $f(j/T, \omega_0)$. Define, for a given sequence a_T with $a_T \to \infty$ and $a_T m_T^{-1} = o(1/\log(M_T))$,

$$\widetilde{f}(j/T, \omega_0) = \begin{cases} 1 + (m_T - la_T)^{\theta} T^{-\theta}, & T\lambda_b^0 + (l-1)a_T < j \leq T\lambda_b^0 + la_T, \quad 1 \leq l \leq m_T/a_T \\ 1 & \text{else} \end{cases}.$$

We are given the observations $I_{L,h,T}(j/T, \omega)$ for $j = n_T + 1, \ldots, T$ and $\omega \in [-\pi, \pi]$. Assume without loss of generality that $\omega_0 \neq \pm\pi, \pm 3\pi, \ldots$. By Theorem 3.4-(ii), $I_{L,h,T}(j/T, \omega_0)$ is approximately $f(j/T, \omega_0)\chi_2^2/2$ for $j/T \neq \lambda_b^0$. For $j/T = \lambda_b^0$, $I_{L,h,T}(j/T, \omega_0)$ is approximately $f(j/T, \omega_0)\chi_2^2/2$ which also follows from Theorem 3.4-(ii) since by Assumption 2.1 is continuous from the left at λ_b^0. However, note that $I_{L,h,T}(j/T, \omega_0)$ is not asymptotically independent of $I_{L,h,T}(l/T, \omega_0)$ for $l = j - n_T + 1, \ldots, j$. Let $S_J = \{n_T + 1, \ldots, T\}$. Let $\zeta_j = f(j/T, \omega_0)\chi_2^2/2$ and $\zeta_j^* = f(j/T, \omega_0)\chi_2^2/2$ where ζ_j^* are independent across $j \in S_J$. Define $\widetilde{\zeta}_j^* = \widetilde{f}(j/T, \omega_0)\chi_2^2/2$ where $\widetilde{\zeta}_j$ are independent across $j \in S_J$.

We distinguish between two cases: (i) $\theta > 1/2$ and (ii) $\theta \leq 1/2$.

(i) Case $\theta > 1/4$. Let us consider the following distinct experiments:

$\mathcal{E}_1$: Observe $\{\zeta_j\}_{j \in S_J}$ and information $T\lambda_b^0 m_T^{-1} \in \{1, 2, \ldots, \lfloor T/m_T \rfloor - 1\}$ is provided.

$\mathcal{E}_2$: Observe $\{\zeta_j^*\}_{j \in S_J}$and information $T\lambda_b^0 m_T^{-1} \in \{1, 2, \ldots, \lfloor T/m_T \rfloor - 1\}$ is provided.

$\mathcal{E}_3$: Observe $\{\widetilde{\zeta}_j^*\}_{j \in S_J}$and information $T\lambda_b^0 m_T^{-1} \in \{1, 2, \ldots, \lfloor T/m_T \rfloor - 1\}$ is provided.

$\mathcal{E}_4$: Observe $\chi = ((\widetilde{f}(jm_T/T, \omega_0)\chi^2_{2m_T,j})_{j \in \mathcal{I}_1}, (\widetilde{f}(\lambda_b^0 + ((j-1)a_T + 1)/T, \omega_0)\widetilde{\chi}^2_{2m_T,j})_{j \in \mathcal{I}_2})$, where $\mathcal{I}_1 = \{1, \ldots, \lambda_b^0 T m_T^{-1}, \lambda_b^0 T m_T^{-1}+2, \ldots, \lfloor T/m_T \rfloor\}$, $\mathcal{I}_2 = \{1, 2, \ldots, m_T a_T^{-1}\}$, and $\{\chi^2_{2m_T,j}\}_{j \in \mathcal{I}_1}$ and $\{\widetilde{\chi}^2_{2a_T,j}\}_{j \in \mathcal{I}_2}$ are i.i.d. sequences of chi-square random variables with $2m_T$ and $2a_T$ degrees of freedom, respectively. Further, information $T\lambda_b^0 m_T^{-1} \in \{1, 2, \ldots, \lfloor T/m_T \rfloor - 1\}$ is provided.

$\mathcal{E}_5$: Observe $\xi = ((m_T^{1/2}\xi_j\widetilde{f}(jm_T/T, \omega_0) + \widetilde{f}(jm_T/T, \omega_0))_{j \in \mathcal{I}_1}, (a_T^{1/2}\widetilde{\xi}_j\widetilde{f}(\lambda_b^0 + ((j-1)a_T + 1)/T, \omega_0) + \widetilde{f}(\lambda_b^0 + ((j-1)a_T + 1)/T, \omega_0))_{j \in \mathcal{I}_2})$, where $\{\xi_j\}_{j \in \mathcal{I}_1}$ and $\{\widetilde{\xi}_j\}_{j \in \mathcal{I}_2}$ are i.i.d standard normal random variables. Further, information $T\lambda_b^0 m_T^{-1} \in \{1, 2, \ldots, \lfloor T/m_T \rfloor - 1\}$ is provided.

We assume that $\{\zeta_j\}_{j \in S_J}$ and $\{\zeta_j^*\}_{j \in S_J}$ are realized on the same probability space which is rich enough to allow for both sequences to be realized there. This is richer then the probability space in which $\{\zeta_j\}_{j \in S_J}$ is realized. Thus, the latter probability space is extended in the usual way using product spaces. The symbol $\approx$ denotes asymptotic equivalence while $\sim$ denotes strong Le Cam equivalence. Our proof consists of showing the following strong Le Cam equivalence of statistical experiments:

$$\mathcal{E}_1 \approx \mathcal{E}_2 \approx \mathcal{E}_3 \sim \mathcal{E}_4 \approx \mathcal{E}_5. \tag{S.20}$$

Therefore, given the relation (S.20), the lower bound for $\mathcal{E}_5$ carries over to the less informative experiment $\mathcal{E}_1$. We prove (S.20) in steps.

Step 1: $\mathcal{E}_1 \approx \mathcal{E}_2$. Given $\zeta_j = f(j/T, \omega_0)\chi_2^2/2$ and the boundness of $f(\cdot, \cdot)$, Theorem 1 in Berkes and

Philipp (1979) implies that there exists a sequence $\{\zeta_j^*\}_{j\in S_J}$ of independent random variables such that ζ_j^* has the same distribution as ζ_j and $\mathbb{P}(|\zeta_j-\zeta_j^*|\geq \nu_j)\leq \nu_j$ with $\nu_j>0$. In view of Assumption 4.1 we have $\sum_{j=1}^{\infty}\nu_j<\infty$ which in turn yields,

$$\sum_{j=1}^{\infty}\left|\zeta_j-\zeta_j^*\right|<\infty \qquad \mathbb{P}-\text{almost surely.} \tag{S.21}$$

Note that

$$|S_J|^{-1}\sum_{j\in S_J}\left|\zeta_j-\zeta_j^*\right|=|S_J|^{-1}\sum_{j=1}^{J_1}\left|\zeta_j-\zeta_j^*\right|+|S_J|^{-1}\sum_{j=J_1+1}^{T}\left|\zeta_j-\zeta_j^*\right|.$$

Choose J_1 large enough such that $\sum_{j=J_1+1}^{T}\left|\zeta_j-\zeta_j^*\right|=o_{\mathrm{a.s}}(|S_J|)$. Thus, $|S_J|^{-1}\sum_{j=1}^{S_J}\left|\zeta_j-\zeta_j^*\right|\to 0$ $\mathbb{P}$-almost surely. This implies that $||\mathbb{P}_{\{|S_J|^{-1}\zeta_j\}}-\mathbb{P}_{\{|S_J|^{-1}\zeta_j^*\}}||_{\mathrm{TV}}\to 0$. The latter shows that $\mathcal{E}_1\approx\mathcal{E}_2$.

Step 2: $\mathcal{E}_2\approx\mathcal{E}_3$. Note that $c\chi_2^2$ with $c>0$ is approximately distributed as $\Gamma(1,\,2c)$ where $\Gamma(a,\,b)$ is the Gamma distribution with parameters $(a,\,b)$. The Kullback-Leibler divergence of $\Gamma(1,\,2c)$ from $\Gamma(1,\,2\widetilde{c})$ is given by

$$D_{\mathrm{KL}}\left(\mathbb{P}_c||\,\mathbb{P}_{\widetilde{c}}\right)=(\log c-\log\widetilde{c})+\frac{\widetilde{c}-c}{c}.$$

For $c=\widetilde{c}+\delta$ with $\delta\to 0$, we obtain

$$\begin{aligned}D_{\mathrm{KL}}\left(\mathbb{P}_c||\,\mathbb{P}_{\widetilde{c}}\right)&=\log\left(\frac{\widetilde{c}+\delta}{\widetilde{c}}\right)+\frac{\widetilde{c}-(\widetilde{c}+\delta)}{\widetilde{c}+\delta}\\&=-\frac{\delta^2}{2\widetilde{c}^2}+O\left(\delta^2\right)+O\left(\delta^3\right).\end{aligned} \tag{S.22}$$

By Pinsker's inequality,

$$\left\|\mathbb{P}_{\{\zeta_j^*\}}-\mathbb{P}_{\{\widetilde{\zeta}_j^*\}}\right\|_{\mathrm{TV}}^2\leq\frac{1}{2}D_{\mathrm{KL}}\left(\mathbb{P}_{\zeta_j^*}||\,\mathbb{P}_{\widetilde{\zeta}_j^*}\right).$$

Thus, using (S.22) and the additivity of Kullback-Leibler divergence for independent distributions, we have

$$D_{\mathrm{KL}}\left(\mathbb{P}_{\zeta_j^*}||\,\mathbb{P}_{\widetilde{\zeta}_j^*}\right)=C\sum_{s=1}^{m_Ta_T^{-1}}\sum_{j=1}^{a_T}\left(jT^{-1}\right)^{2\theta}=CO\left(a_TT^{-1}\right)^{2\theta}m_T.$$

This tends to zero in view of (S.18) and $m_T^{-1}a_T=o(1/\log(M_T))$.

Step 3: $\mathcal{E}_3\sim\mathcal{E}_4$. The vector of averages

$$\left(\left((2m_T)^{-1}\sum_{s=1}^{m_T}\widetilde{\zeta}_{jm_T+s-1}^*\right)_{j\in\mathcal{I}_1},\,\left((2a_T)^{-1}\sum_{s=1}^{a_T}\widetilde{\zeta}_{T\lambda_b^0+(j-1)a_T+s}^*\right)_{j\in\mathcal{I}_2}\right),$$

forms a sufficient statistic for $\left\{\widetilde{f}\,(j/T,\,\omega_0)\right\}_{(j/T)\in[0,\,1]}$. Hence, by Lemma 3.2 of Brown and Low (1996) this yields the strong Le Cam equivalence.

Step 4: $\mathcal{E}_4 \approx \mathcal{E}_5$. Let

$$\begin{aligned}\chi^* = &(m_T^{-1/2}(\widetilde{f}(jm_T/T,\,\omega_0)(\chi^2_{2m_T,j}-2m_T))_{j\in\mathcal{I}_1},\\ &a_T^{-1/2}(\widetilde{f}(\lambda_b^0+((j-1)a_T+1)/T,\,\omega_0)(\widetilde{\chi}^2_{2m_T,j}-2a_T))_{j\in\mathcal{I}_2})\\ \xi^* = &((\xi_j\widetilde{f}(jm_T/T,\,\omega_0))_{j\in\mathcal{I}_1},\,(\widetilde{\xi}_j\widetilde{f}(\lambda_b^0+((j-1)a_T+1)/T,\,\omega_0))_{j\in\mathcal{I}_2}).\end{aligned}$$

Note that $\|\mathbb{P}_\chi-\mathbb{P}_\xi\|^2_{\mathrm{TV}}=\|\mathbb{P}_{\chi^*}-\mathbb{P}_{\xi^*}\|^2_{\mathrm{TV}}$. By Pinsker's inequality and independence,

$$\begin{aligned}&||\mathbb{P}_{\chi^*}-\mathbb{P}_{\xi^*}||^2_{\mathrm{TV}}\\ &\quad\leq 2^{-1}D_{\mathrm{KL}}\left(\mathbb{P}_{\chi^*}||\,\mathbb{P}_{\xi^*}\right)\\ &\quad\leq 2^{-1}\sum_{j\in\mathcal{I}_1}D_{\mathrm{KL}}\left(m_T^{-1/2}\left(\widetilde{f}\left(jm_T/T,\,\omega_0\right)\left(\chi^2_{2m_T,j}-2m_T\right)\right)||\,\xi_j\widetilde{f}\left(jm_T/T,\,\omega_0\right)\right)\\ &\qquad+2^{-1}\sum_{j\in\mathcal{I}_2}D_{\mathrm{KL}}(a_T^{-1/2}(\widetilde{f}(\lambda_b^0+((j-1)a_T+1)/T,\,\omega_0)(\chi^2_{2a_T,j}-2a_T))\\ &\qquad||\,\xi_j\widetilde{f}(\lambda_b^0+((j-1)a_T+1)/T,\,\omega_0)).\end{aligned}$$

We now apply Theorem 1.1 in Bobkov, Chistyakov, and Götze (2013) with $c_1=12^{-1}\kappa_3^2$ in (1.3) there, where κ_3 is the third-order cumulant of the variable in question. This gives the following bounds,

$$D_{\mathrm{KL}}\left((4m_T)^{-1/2}\left(\widetilde{f}\left(jm_T/T,\,\omega_0\right)\left(\chi^2_{2m_T,j}-2m_T\right)\right)||\,\xi_j\widetilde{f}\left(jm_T/T,\,\omega_0\right)\right)=\frac{1}{12}\left(\frac{8}{2m_T}\right)+o\left(\frac{1}{m_T\log m_T}\right),$$

and

$$\begin{aligned}&D_{\mathrm{KL}}(a_T^{-1/2}(\widetilde{f}(\lambda_b^0+((j-1)a_T+1)/T,\,\omega_0)(\chi^2_{2a_T,j}-2a_T))\\ &\qquad||\,\xi_j\widetilde{f}(\lambda_b^0+((j-1)a_T+1)/T,\,\omega_0))=\frac{1}{12}\left(\frac{8}{2a_T}\right)+o\left(\frac{1}{a_T\log a_T}\right).\end{aligned}$$

Hence, $||\mathbb{P}_{\chi^*}-\mathbb{P}_{\xi^*}||^2_{\mathrm{TV}}=O(Tm_T^{-2})+O(m_Ta_T^{-2})$. Since $\theta>1/2$ we have $Tm_T^{-2}\to 0$. Finally, since $m_T^{-1}a_T\to 0$ we can choose a_T sufficiently fast to yield $m_Ta_T^{-2}\to 0$. This shows that $||\mathbb{P}_\chi-\mathbb{P}_\xi||_{\mathrm{TV}}\to 0$.

By step 1-4, it is sufficient to establish the minimax lower bound for experiment $\mathcal{E}_5$. After adding an additional drift ξ, which gives an equivalent problem, we cast the problem as a high dimensional location signal detection problem [cf. Ingster and Suslina (2003)] from which the bound can be derived using classical arguments. Consider observations

$$\begin{aligned}\xi^* = &((m_T^{-1/2}\xi_j\widetilde{f}(jm_T/T,\,\omega_0)+\widetilde{f}(jm_T/T,\,\omega_0)-1)_{j\in\mathcal{I}_1},\\ &(a_T^{-1/2}\widetilde{\xi}_j\widetilde{f}(\lambda_b^0+((j-1)a_T+1)/T,\,\omega_0)+\widetilde{f}(\lambda_b^0+((j-1)a_T+1)/T,\,\omega_0)-1)_{j\in\mathcal{I}_2}),\end{aligned}$$

and the hypothesis

$$\mathcal{H}_0:\,\sup_j\left(\widetilde{f}\left(j/T,\,\omega_0\right)-1\right)=0\qquad\text{versus}\qquad\mathcal{H}_1:\,\sup_j\left(\widetilde{f}\left(j/T,\,\omega_0\right)-1\right)\geq b_T.\tag{S.23}$$

The goal is to find the maximal value $b_T\to 0$ such that the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are non-distinguishable in the minimax sense or $\lim_{T\to\infty}\inf_\psi\gamma_\psi\left(\theta,\,b_T\right)=1$. Here the detection rate is $b_T\propto\left(T^{-1}m_T\right)^\theta\propto T^{-\frac{\theta}{2\theta+1}}$.

Consider the product measures $\mathbb{P}_{\mathcal{H}_0} = \mathbb{P}_{\xi^*} \times \mathbb{P}_0$ and $\mathbb{P}_{\mathcal{H}_1} = \mathbb{P}_{\xi^*} \times \mathbb{P}_{\lambda_b^0,1}$ where $\mathbb{P}_{\xi^*}$ is the probability law of ξ^* and $\mathbb{P}_0$ is the measure for the no break case. Thus, $\mathbb{P}_{\mathcal{H}_0}$ is the probability measure under $\mathcal{H}_0$ while $\mathbb{P}_{\mathcal{H}_1}$ is the probability measure under $\mathcal{H}_1$ which draws a break at time λ_b^0 with $T\lambda_b^0 m_T^{-1} \in \{1, 2, \ldots, \lfloor T/m_T \rfloor - 1\}$ uniformly from this set. From similar derivations that yield eq. (2.20)-(2.22) in Ingster and Suslina (2003), it follows that

$$\inf_{\psi} \gamma_\psi (\theta,\, b_T) \geq 1 - \frac{1}{2} \left\| \mathbb{P}_{\mathcal{H}_1} - \mathbb{P}_{\mathcal{H}_0} \right\|_{\mathrm{TV}} \geq 1 - \frac{1}{2} \left| \mathbb{E}_{\mathbb{P}_{\mathcal{H}_0}} \left(\mathscr{L}_{0,1}^2 - 1 \right) \right|^{1/2},$$

where $\mathscr{L}_{0,1} = d\mathbb{P}_{\mathcal{H}_1}/d\mathbb{P}_{\mathcal{H}_0}$ is the likelihood ratio between $\mathbb{P}_{\mathcal{H}_1}$ and $\mathbb{P}_{\mathcal{H}_0}$. By the above inequality, it is sufficient to show $\mathbb{E}_{\mathbb{P}_{\mathcal{H}_0}}(\mathscr{L}_{0,1}^2) \to 1$. The proof of the latter result follows similar arguments as in Bibinger, Jirak, and Vetter (2017).

It remains to consider the case $\theta \leq 1/2$. In a different setting, Bibinger, Jirak, and Vetter (2017) considered separately the case where their regularity exponent $\mathfrak{a} \leq 1/2$ to obtain the minimax lower bound. The same arguments can be applied in our context which lead to the same result as for the case $\theta > 1/2$.

The general case with $f_-(\omega_0) > 0$ rather than with $f_-(\omega_0) = 1$ as discussed above follows from the same arguments after we rescale the equations in (S.23). The only difference is the form of the detection rate which is now $b_T \leq f_-(\omega_0) D(T^{-1} m_T)^\theta$.

The proof for the lower bound for the alternative $\mathcal{H}_1'$ is similar to the proof discussed above. The minor differences in the proof outlined by Bibinger, Jirak, and Vetter (2017) also apply here. □

S.A.4.2 Proof of Theorem 5.2

We present the proof for the statistic $\max_{\omega \in \Pi} \mathrm{I}_{\max,T}(\omega)$ constructed with $I_{L,h,T}(j/T,\, \omega)$ in place of $I_{h,T}(j/T,\, \omega)$. The proof for the other test statistics discussed in Section 4 is similar and omitted. Without loss of generality we assume that $\omega_0 \neq \pm\pi$. Let $\widetilde{I}_{L,r,T}(\omega)$ be defined as $\widetilde{I}_{r,T}(\omega)$ but with $I_{L,h,T}(j/T,\, \omega)$ in place of $I_{h,T}(j/T,\, \omega)$. Let $\widetilde{I}^*_{L,r,T}(\omega)$ be defined as $\widetilde{I}^*_{r,T}(\omega)$ but with $I^*_{L,T}(r/T,\, \omega) = I_{L,h,T}(j/T,\, \omega) - \mathbb{E}(I_{L,h,T}(j/T,\, \omega))$ in place of $I^*_T(r/T,\, \omega)$. As in (S.12)-(S.13), if $T\lambda_b^0 \notin [rm_T^* - m_T^*/2 + 1,\, (r+1)\, m_T^* + m_T^*/2]$ or if $\omega \neq \omega_0$ then

$$\begin{aligned}
&\left| \frac{\widetilde{I}_{L,r+1,T}(\omega) - \widetilde{I}_{L,r,T}(\omega)}{\sigma_r(\omega)} \right| \\
&= \left| \frac{(m_T^*)^{-1} \sum_{j=(r+1)m_T^* - m_T^*/2+1}^{(r+1)m_T^* + m_T^*/2} \left(I^*_{L,T}(j/T,\, \omega) + \mathbb{E}(I_{L,h,T}(j/T,\, \omega)) \right)}{\sigma_r(\omega)} \right. \\
&\quad \left. - \frac{(m_T^*)^{-1} \sum_{j=rm_T^* - m_T^*/2+1}^{rm_T^* + m_T^*/2} \left(I^*_{L,T}(j/T,\, \omega) + \mathbb{E}(I_{L,h,T}(j/T,\, \omega)) \right)}{\sigma_r(\omega)} \right| \\
&= \left| \frac{(m_T^*)^{-1} \sum_{j=(r+1)m_T^* - m_T^*/2+1}^{(r+1)m_T^* + m_T^*/2} I^*_{L,T}(j/T,\, \omega) - (m_T^*)^{-1} \sum_{j=rm_T^* - m_T^*/2+1}^{rm_T^* + m_T^*/2} I^*_{L,T}(j/T,\, \omega)}{\sigma_r(\omega)} \right| \\
&\quad + O\left((m_T^*/T)^\theta \right) + O_{\mathbb{P}}\left((n_T/T)^2 + \log(n_T)/n_T \right) \\
&\triangleq \mathring{I}_{r,T}(\omega) + O\left((m_T^*/T)^\theta \right) + O_{\mathbb{P}}\left((n_T/T)^2 + \log(n_T)/n_T \right)
\end{aligned}$$

$$= \mathring{I}_{r,T}(\omega) + O_{\mathbb{P}}\left(\left(\sqrt{\log(M_T^*)}\,(m_T^*)^{1/2}\right)^{-1}\right).$$

If $T\lambda_b^0 \in [rm_T^* - m_T^*/2 + 1,\, (r+1)\,m_T^* + m_T^*/2]$ and $\omega = \omega_0$,

$$\frac{1}{\sqrt{m_T^*}} \sum_{j=vm_T^*-m_T^*/2+1}^{vm_T^*+m_T^*/2} I_{L,T}^*(j/T,\,\omega_0) = O_{\mathbb{P}}(1), \qquad v = r,\, r+1.$$

Hence, we yield that $\sqrt{m_T^*}\mathring{I}_{r,T}(\omega) = O_{\mathbb{P}}(1)$ for $1 \le r \le M_T^* - 2$. This can be used to obtain the following inequality, for $rm_T^* - m_T^*/2 + 1 \le T\lambda_b^0 \le rm_T^* + m_T^*/2$

$$\begin{aligned}
\mathrm{I}_{\max,T}(\omega_0) \ge & -\mathring{I}_{r,T}(\omega_0) \\
& + \frac{1}{m_T^*}\left|\int_{(rm_T^*-m_T^*/2+1)/T}^{\lambda_b^0} f(u,\,\omega_0)\,du - \int_{\lambda_b^0}^{(r+1)m_T^*+m_T^*/2} f(u,\,\omega_0)\,du\right| \frac{(1-o_{\mathbb{P}}(1))}{\sup_r \sigma_r(\omega_0)}, \\
\ge & -O_{\mathbb{P}}\left((m_T^*)^{-1/2}\right) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(S.24)} \\
& + \frac{1}{m_T^*}\left|\int_{(rm_T^*-m_T^*/2+1)/T}^{\lambda_b^0} f(u,\,\omega_0)\,du - \int_{\lambda_b^0}^{(r+1)m_T^*+m_T^*/2} f(u,\,\omega_0)\,du\right| \frac{(1-o_{\mathbb{P}}(1))}{\sup_r \sigma_r(\omega_0)}.
\end{aligned}$$

In order to prove $\gamma_{\psi^*}(\theta,\, b_T^*) \to 0$, it suffices to show that

$$\mathbb{P}\left(\mathrm{I}_{\max,T}(\omega) < 2D^*\sqrt{\log(M_T^*)/m_T^*}\right) \to 1,\ \forall\omega \in [-\pi,\,\pi] \qquad \text{under } \mathcal{H}_0 \tag{S.25}$$

$$\mathbb{P}\left(\mathrm{I}_{\max,T}(\omega) \ge 2D^*\sqrt{\log(M_T^*)/m_T^*}\right) \to 1, \text{for some}\,\omega \in [-\pi,\,\pi], \quad \text{under } \mathcal{H}_1 \text{ or } \mathcal{H}_1', \tag{S.26}$$

We first show (S.25). Note that

$$2D^*\sqrt{\log(M_T^*)/m_T^*} \ge 2\sqrt{\log(M_T^*)/m_T^*} + D\,(m_T^*/T)^{\theta}.$$

Under $\mathcal{H}_0$, since $\theta' < \theta$ we have

$$\mathrm{I}_{\max,T}(\omega_0) \le \max_{1\le r\le M_T^*-2} \mathring{I}_{r,T}(\omega_0) + D\,(m_T^*/T)^{\theta'} + O_{\mathbb{P}}\left((n_T/T)^2 + \log(n_T)/n_T\right).$$

Since $(n_T/T)^2 + \log(n_T)/n_T \le 2D^*\sqrt{\log(M_T^*)/m_T^*}$, to conclude the proof we have to show

$$\mathbb{P}\left(\max_{1\le r\le M_T^*-2} \mathring{I}_{r,T}(\omega_0) \le \sqrt{\log(M_T^*)/m_T^*}\right) \to 1.$$

The latter result follows from Theorem 4.1.

We now prove (S.25) under $\mathcal{H}_1$. We have to show that the second term on the right hand side of (S.24) is greater than or equal to $2D^*\sqrt{\log(M_T^*)/m_T^*}$. The term in question is larger than $b_T^* - 2D\,(m_T^*/T)^{\theta}$. In view of (5.2) with $\theta' = 0$ the result follows.

We now prove S.25 under $\mathcal{H}_1'$. For $h \le 2m_T^*/T$ we have $f(\lambda_b^0 + h,\,\omega_0) \ge f(\lambda_b^0,\,\omega_0) + b_T^* h^{\theta'}$ or

$f\left(\lambda_b^0+h,\,\omega_0\right)\leq f\left(\lambda_b^0,\,\omega_0\right)-b_T^*h^{\theta'}$. Thus,

$$\frac{1}{m_T^*}\left|\int_{\lambda_b^0+n_T^*/T+m_T^*/T}^{\lambda_b^0+2\left(n_T^*/T+m_T^*/T\right)}f\left(u,\,\omega_0\right)du\right|\geq b_T\left(m_T^*/T\right)^{\theta'}$$
$$\geq 2D^*\sqrt{\log\left(M_T^*\right)/m_T^*},$$

where the second equality follows from (5.2). □

S.A.5 Proofs of the Results of Section 6

S.A.5.1 Proof of Proposition 6.1

The following lemma is simple to verify. It was also used by Bibinger, Jirak, and Vetter (2017) in a different context.

Lemma S.A.6. *Let $C(u)$ and $d(u)$ be functions on $\left[0,\,\lambda_b^0\right]$ such that $d(u)$ is increasing. As long as $d\left(\lambda_b^0\right)-d\left(\lambda_b^0-\kappa\right)\geq\sup_{0\leq u\leq\lambda_b^0}|C(u)|$ for some $\kappa\in\left[0,\,\lambda_b^0\right]$ we have that,*

$$\underset{0\leq u\leq\lambda_b^0}{\operatorname{argmax}}\left(d(u)+C(u)\right)\geq\lambda_b^0-\kappa. \tag{S.27}$$

An analogous results holds if $C(u)$ and $d(u)$ are functions on $\left[\lambda_b^0,\,1\right]$ and $d(u)$ is decreasing.

Proof of Proposition 6.1. For $\lambda_b^0\in(0,\,1)$ define $\overline{r}_b=\left\lceil T\lambda_b^0+1\right\rceil$, i.e., the smallest integer such that $\overline{r}_b/T$ is larger or equal than λ_b^0+1/T. Denote by $\left\{\widetilde{f}\left(u,\,\omega_0\right)\right\}_{u\in[0,\,1]}$ the path of the spectrum $f\left(\cdot,\,\omega_0\right)$ without the break: $f\left(r/T,\,\omega\right)=\widetilde{f}\left(r/T,\,\omega\right)+\delta_T\mathbf{1}\left\{r\geq\overline{r}_b\right\}$. Without loss of generality, we assume $\delta_T>0$. Define $d\left(r/T,\,\omega\right)=0$ for $\omega\neq\omega_0$ and

$$d\left(r/T,\,\omega_0\right)=\begin{cases}0 & \text{if } r+m_T<\overline{r}_b,\\ \left(r+m_T-\overline{r}_b\right)m_T^{-1/2}\delta_T & \text{if } r=\overline{r}_b-m_T,\ldots,\overline{r}_b,\\ m_T^{1/2}\delta_T & \text{if } r>\overline{r}_b,\end{cases}$$

and $\left\{d\left(u,\,\omega_0\right)\right\}_{u\in[0,\,1]}$ is the associated piecewise constant increasing step function. For $r=m_T,\ldots,T-m_T$, write

$$\begin{aligned}\sum_{j=r-m_T+1}^{r}I_T\left(j/T,\,\omega_0\right)-\sum_{j=r+1}^{r+m_T}I_T\left(j/T,\,\omega_0\right)\\ =\sum_{j=r-m_T+1}^{r}\left(I_T\left(j/T,\,\omega_0\right)-\mathbb{E}\left(I_T\left(j/T,\,\omega_0\right)\right)\right)-\sum_{j=r+1}^{r+m_T}\left(I_T\left(j/T,\,\omega_0\right)-\mathbb{E}\left(I_T\left(j/T,\,\omega_0\right)\right)\right)\\ +\sum_{j=r-m_T+1}^{r}\left(\mathbb{E}\left(I_T\left(j/T,\,\omega_0\right)\right)-\widetilde{f}\left(j/T,\,\omega_0\right)\right)-\sum_{j=r+1}^{r+m_T}\left(\mathbb{E}\left(I_T\left(j/T,\,\omega_0\right)-f\left(j/T,\,\omega_0\right)\right)\right)\\ +\sum_{j=r-m_T+1}^{r}\widetilde{f}\left(j/T,\,\omega_0\right)-\sum_{j=r+1}^{r+m_T}\widetilde{f}\left(j/T,\,\omega_0\right)-\sum_{j=r+1}^{r+m_T}\left(f\left(j/T,\,\omega_0\right)-\widetilde{f}\left(j/T,\,\omega_0\right)\right).\end{aligned}$$

For $r = m_T, \ldots, \overline{r}_b$, let $C(r/T, \omega) = \mathrm{D}_{r,T}(\omega)$ for $\omega \neq \omega_0$ and

$$C(r/T, \omega_0) = m_T^{-1/2} \Bigg(\sum_{j=r-m_T+1}^{r} I_T(j/T, \omega_0) - \sum_{j=r+1}^{r+m_T} I_T(j/T, \omega_0) + \sum_{j=\overline{r}_b+1}^{r+m_T} \left(f(j/T, \omega_0) - \widetilde{f}(j/T, \omega_0) \right) \Bigg).$$

Note that $C(s/T, \omega)$ does not involve any break for any ω. Thus, we can proceed similarly as in the proofs of Section 4. That is, we exploit the smoothness of $f(\cdot, \cdot)$ as under $\mathcal{H}_0$ to yield $\sup_{u \in [0, \lambda_b^0]} \sup_{\omega \in [-\pi, \pi]} |C(u, \omega)| = O_{\mathbb{P}}\left(\sqrt{\log(T)}\right)$. This combined with the definition of $d(r/T, \omega_0)$ implies

$$|d(r/T, \omega_0)| > \max_{\omega \in [-\pi, \pi]} \left(|d(r/T, \omega)|\right) > 0,$$

with probability approaching one and

$$\mathrm{D}_{r,T}(\omega) = |d(r/T, \omega) + C(r/T, \omega)| = d(r/T, \omega) + \mathrm{sign}(C(r/T, \omega)) |C(r/T, \omega)|,$$

for each $r = \overline{r}_b - \lfloor m_T/B \rfloor, \ldots, \overline{r}_b$ where B is any finite integer with $B > 1$. By definition of $d(\cdot, \omega_0)$, for $\kappa_T \in [0, m_T/(BT)]$,

$$d(\overline{r}_b/T, \omega_0) - d(\overline{r}_b/T - \kappa_T, \omega_0) = \lfloor \kappa_T T \rfloor \delta_T m_T^{-1/2}.$$

In order to apply Lemma S.A.6, we need to choose κ_T such that $\lfloor \kappa_T T \rfloor \delta_T m_T^{-1/2} / \sqrt{\log(T)} \geq 1$ or $\sqrt{m_T \log(T)}/\delta_T T = o(\kappa_T)$. Lemma S.A.6 then yields

$$\frac{\overline{r}_b}{T} \geq \underset{r=m_T,\ldots,\overline{r}_b}{\mathrm{argmax}} \max_{\omega \in [-\pi, \pi]} T^{-1}\mathrm{D}_{r,T}(\omega) = \underset{r=m_T,\ldots,\overline{r}_b}{\mathrm{argmax}} T^{-1}\mathrm{D}_{r,T}(\omega_0) \geq \frac{\overline{r}_b}{T} - \kappa_T.$$

The case $r > \overline{r}_b$ can be treated similarly by symmetry. It results in

$$\frac{\overline{r}_b}{T} \leq \underset{r=\overline{r}_b,\ldots,T-m_T}{\mathrm{argmax}} \max_{\omega \in [-\pi, \pi]} T^{-1}\mathrm{D}_{r,T}(\omega) = \underset{r=\overline{r}_b,\ldots,T-m_T}{\mathrm{argmax}} T^{-1}\mathrm{D}_{r,T}(\omega_0) \leq \frac{\overline{r}_b}{T} + \kappa_T.$$

Therefore, we conclude $\left|\widehat{\lambda}_b - \overline{r}_b/T\right| = O_{\mathbb{P}}(\kappa_T) \to 0$. □

S.A.5.2 Proof of Proposition 6.2

Set $\widehat{\mathcal{I}} = \{m_T, \ldots, T - m_T\} \setminus \{m_T\}$ and $\widehat{\mathcal{T}} = \emptyset$. Under $\mathcal{H}_{1,\mathrm{M}}$, the arguments in the proof of Theorem 4.1 yields,

$$\underset{r=m_T,\ldots,T-m_T}{\mathrm{argmax}} \max_{\omega \neq \omega_0} \mathrm{D}_{r,T}(\omega) = O_{\mathbb{P}}\left(\sqrt{\log(T)}\right).$$

Let $r_l, r_{l+1} \in \widehat{\mathcal{I}}$ such that $r_{l+1} = r_l + m_T$ and $r_l \leq T_l^0 < r_{l+1}$, $l = 1, \ldots, m_0$. For $\omega = \omega_0$ we have

$$\underset{r \in \widehat{\mathcal{I}} \setminus \{r_1, \ldots, r_{m+1}\}}{\mathrm{argmax}} \mathrm{D}_{r,T}(\omega_0) = O_{\mathbb{P}}\left(\sqrt{\log(T)}\right).$$

For each $r \in \widehat{\mathcal{I}}$, we draw K points $r^{\diamond}_{k,r}$ with $k = 1, \ldots, K$ uniformly (without replacement) from $\mathbf{I}(r)$. Consider the following events,

$$\mathbf{D}_1 = \left\{\forall r \in \widehat{\mathcal{I}}\, \forall\, k = 1, \ldots, K, \, (\exists!\, 1 \leq l \leq m_0) \vee (\nexists 1 \leq l \leq m_0) \text{ s.t.}\, T^0_l \in \left[r^{\diamond}_{k,r} - m_T, \, r^{\diamond}_{k,r} + m_T\right]\right\}$$
$$\mathbf{D}_2 = \left\{\forall l = 1, \ldots, m_0\, \exists\, r \in \widehat{\mathcal{I}} \,\text{s.t.}\, \exists\, k = 1, \ldots, K, \text{ s.t. } \left|T^0_l - r^{\diamond}_{k,r}\right| = Cm_T \,\text{for some}\, C \in [0, 1)\right\}.$$

Let $\mathbf{A}^c$ denote the complement of a set $\mathbf{A}$. Note that $\mathbb{P}\left((\mathbf{D}_1 \cap \mathbf{D}_2)^c\right) = \mathbb{P}\left((\mathbf{D}_2)^c\right)$ by Assumption 6.2 and that $\mathbb{P}\left((\mathbf{D}_2)^c\right) = 0$ if there are still undetected breaks. The remaining arguments will be valid on the set $\mathbf{D}_1 \cap \mathbf{D}_2$ as long as there are undetected breaks.

Let $r_l, \, r_{l+1} \in \widehat{\mathcal{I}}$ be such that $T^0_l \in [r_l, \, r_{l+1})$. As in the proof of Proposition 6.1,

$$\mathrm{D}_{r_l,T}\left(\omega_0\right) = |O_{\mathbb{P}}\left(m^{-1/2}_T \delta_{l,T}\left(m_T - \left(r_l - T^0_l\right)\mathbf{1}\left\{r_{l-1} < T^0_l \leq r_l\right\} + \left(r_{l+1} - T^0_l\right)\mathbf{1}\left\{r_l < T^0_l < r_{l+1}\right\}\right)\right)|.$$

Note that if $\mathrm{D}_{r_l,T}\left(\omega_0\right) / \left(\delta_{l,T}\sqrt{m_T}\right) \xrightarrow{\mathbb{P}} 0$ then we must have $\mathrm{D}_{r_{l+1},T}\left(\omega_0\right) = O_{\mathbb{P}}\left(\delta_{l,T}\sqrt{m_T}\right)$. Thus, in step (2) $\psi^*\left(\{X_r\}_{r\in\widehat{\mathcal{I}}}\right) = 1$ because for large enough T,

$$\begin{aligned}\max_{r\in\widehat{\mathcal{I}}}\max_{k\in K}\mathrm{D}_{r^{\diamond}_{k,r},T}\left(\omega_0\right) &\geq \max_{r\in\widehat{\mathcal{I}}}\mathrm{D}_{r,T}\left(\omega_0\right)\\ &= |\delta_{l,T}O_{\mathbb{P}}\left(\sqrt{m_T}\right)|\\ &\geq \inf_{1\leq l\leq m_0}|\delta_{l,T}O_{\mathbb{P}}\left(\sqrt{m_T}\right)|\\ &= 2D^*\left(\log\left(T\right)\right)^{2/3}\\ &> 2D^*\sqrt{\log\left(M^*_T\right)},\end{aligned}$$

where the last inequality follows from Assumption 6.2. We now move to step (3). By the arguments in the proof of Proposition 6.1, there exists $1 \leq l \leq m_0$ such that $\left|\lambda^0_l - \widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right)\right| \leq m_T/T$. Since $m_T/v_T \to 0$ there can exist exactly one l that satisfies $\left|\lambda^0_l - \widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right)\right| \leq m_T/T$. For such a λ^0_l define $\overline{r}_{l,b} = \lceil T\lambda^0_l + 1\rceil$, the smallest integer such that $\overline{r}_{l,b}/T$ is larger or equal than $\lambda^0_l + 1/T$. Denote by $\left\{\widetilde{f}\left(u, \, \omega\right)\right\}_{u\in[0,\,1]}$ the path of the spectrum $f\left(\cdot, \, \omega\right)$ without the break $\delta_{l,T}$:

$$f\left(r/T, \, \omega\right) = \widetilde{f}\left(r/T, \, \omega\right) + \delta_{l,T}\mathbf{1}\left\{r \geq \overline{r}_{l,b}\right\}.$$

Without loss of generality, we assume $\delta_{l,T} > 0$. Define $d\left(r/T, \, \omega\right) = 0$ for $\omega \neq \omega_0$ and

$$d_l\left(r/T, \, \omega_0\right) = \begin{cases}0 & \text{if } r + m_T < \overline{r}_{l,b}\\ \left(r + m_T - \overline{r}_{l,b}\right)m^{-1/2}_T\delta_{l,T} & \text{if } r = \overline{r}_{l,b} - m_T, \ldots, \overline{s}_{l,b}.\\ m^{1/2}_T\delta_{l,T} & \text{if } r > \overline{r}_{l,b}\end{cases}$$

Let $\left\{d\left(u\right)\right\}_{u\in[0,\,1]}$ is the associated piecewise constant increasing step function. For any $r \in \widehat{\mathcal{I}}$, write

$$\sum^{r}_{j=r-m_T+1} I_T\left(j/T, \, \omega_0\right) - \sum^{r+m_T}_{j=r+1} I_T\left(j/T, \, \omega_0\right)$$

$$= \sum_{j=r-m_T+1}^{r} \left(I_T(j/T,\,\omega_0) - \mathbb{E}\left(I_T(j/T,\,\omega_0)\right)\right) - \sum_{j=r+1}^{r+m_T} \left(I_T(j/T,\,\omega_0) - \mathbb{E}\left(I_T(j/T,\,\omega_0)\right)\right)$$
$$+ \sum_{j=r-m_T+1}^{r} \left(\mathbb{E}\left(I_T(j/T,\,\omega_0)\right) - \widetilde{f}(j/T,\,\omega_0)\right) - \sum_{j=r+1}^{r+m_T} \left(\mathbb{E}\left(I_T(j/T,\,\omega_0) - f(j/T,\,\omega_0)\right)\right)$$
$$+ \sum_{j=r-m_T+1}^{r} \widetilde{f}(j/T,\,\omega_0) - \sum_{j=r+1}^{r+m_T} \widetilde{f}(j/T,\,\omega_0) - \sum_{j=r+1}^{r+m_T} \left(f(j/T,\,\omega_0) - \widetilde{f}(j/T,\,\omega_0)\right).$$

For $r = m_T, \ldots, \overline{r}_b$, let $C_l(r/T,\,\omega) = \mathrm{D}_{r,T}(\omega)$ for $\omega \neq \omega_0$ and

$$C_l(r/T,\,\omega_0) = m_T^{-1/2} \left(\sum_{j=r-m_T+1}^{r} I_T(j/T,\,\omega_0) - \sum_{j=r+1}^{r+m_T} I_T(j/T,\,\omega_0) \right.$$
$$\left. + \sum_{j=\overline{r}_{l,b}+1}^{r+m_T} \left(f(j/T,\,\omega_0) - \widetilde{f}(j/T,\,\omega_0)\right) \right),$$

We proceed as in the proof of Proposition 6.1. We have $d(r/T,\,\omega_0) \geq \max_{\omega\in[-\pi,\,\pi]} |d(r/T,\,\omega)| > 0$ with probability approaching one . Exploiting the smoothness on $(\lambda_{l-1}^0,\,\lambda_l^0]$, we have

$$\sup_{u\in(\lambda_{l-1}^0,\,\lambda_b^0]} \sup_{\omega\in[-\pi,\,\pi]} |C_l(u,\,\omega)| = O_{\mathbb{P}}\left(\sqrt{\log(T)}\right).$$

This implies

$$\mathrm{D}_{r,T}(\omega) = |d_l(r/T,\,\omega) + C_l(r/T,\,\omega)| = d_l(r/T,\,\omega) + \mathrm{sign}\left(C_l(r/T,\,\omega)\right)|C_l(r/T,\,\omega)|,$$

for each $r = \overline{r}_b - \lfloor m_T/B \rfloor, \ldots, \overline{r}_b$ where B is any integer with $1 < B < \infty$. By definition of $d_l(\cdot,\,\omega_0)$, for $\kappa_T \in [0,\,m_T/(BT)]$ we have

$$d_l(\overline{r}_{l,b}/T,\,\omega_0) - d_l(\overline{r}_{l,b}/T - \kappa_T,\,\omega_0) = \lfloor \kappa_T T \rfloor \delta_{l,T} m_T^{-1/2}.$$

In order to apply Lemma S.A.6, we need to choose κ_T such that $\lfloor \kappa_T T \rfloor \delta_{l,T} m_T^{-1/2}/\sqrt{\log(T)} \geq 1$ or $\sqrt{m_T \log(T)}/\delta_{l,T} T = o(\kappa_T)$. Lemma S.A.6 then yields

$$\frac{\overline{r}_{l,b}}{T} \geq \underset{r\in\left(\widehat{\mathcal{I}}\setminus\left\{r:\,r>\overline{r}_{l,b}\right\}\right)}{\mathrm{argmax}} \max_{\omega\in[-\pi,\,\pi]} T^{-1}\mathrm{D}_{r,T}(\omega) = \underset{r\in\left(\widehat{\mathcal{I}}\setminus\left\{r:\,r>\overline{r}_{l,b}\right\}\right)}{\mathrm{argmax}} T^{-1}\mathrm{D}_{r,T}(\omega_0) \geq \frac{\overline{r}_{l,b}}{T} - \kappa_T.$$

The case $r > \overline{r}_b$ can be treated similarly by symmetry. It results in

$$\frac{\overline{r}_{l,b}}{T} \leq \underset{r\in\left(\widehat{\mathcal{I}}\setminus\left\{r:\,r<\overline{r}_{l,b}\right\}\right)}{\mathrm{argmax}} \max_{\omega\in[-\pi,\,\pi]} T^{-1}\mathrm{D}_{r,T}(\omega) = \underset{r\in\left(\widehat{\mathcal{I}}\setminus\left\{r:\,r<\overline{r}_{l,b}\right\}\right)}{\mathrm{argmax}} T^{-1}\mathrm{D}_{r,T}(\omega_0) \leq \frac{\overline{r}_{l,b}}{T} + \kappa_T.$$

Therefore, we conclude $\left|\widehat{\lambda}_T - \overline{r}_{l,b}/T\right| = O_{\mathbb{P}}(\kappa_T) \to 0$. Now set $\widehat{\mathcal{I}} = \widehat{\mathcal{I}}\setminus\left\{T\widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right) - v_T, \ldots, T\widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right) + v_T\right\}$ and $\widehat{\mathcal{T}} = \widehat{\mathcal{T}} \cup \left\{T\widehat{\lambda}_T\left(\widehat{\mathcal{I}}\right)\right\}$. Since $\mathbb{P}\left((\mathbf{D}_2)^c\right) = 0$ if there are still undetected breaks, we can repeat the above steps (1)-(4). The final results are $\mathbb{P}\left(\left|\widehat{\mathcal{T}}\right| = m_0\right) \to 1$ and, after ordering the elements of $\widehat{\mathcal{T}}$ in chronological order, $\sup_{1\leq l\leq m_0}\left|\widehat{\lambda}_{l,T}-\lambda_l^0\right| = O_{\mathbb{P}}\left(\sqrt{m_T\log(T)}/\left(T\inf_{1\leq l\leq m_0}\delta_{l,T}\right)\right)$.

Assume without loss of generality that $\delta_{1,T}\geq\delta_{2,T}\geq\cdots\geq\delta_{m_0,T}$. Let $\widehat{\lambda}_T^{(q)}$ $(q=1,\ldots,m_0)$ denote the qth break detected by the procedure. It remains to prove that if $K\to\infty$ then $\widehat{\lambda}_T^{(q)}$ is consistent for λ_q^0 $(q=1,\ldots,m_0)$. Consider the first break λ_1^0. In order for the algorithm to return $\widehat{\lambda}_T^{(1)}$ such that $\left|\widehat{\lambda}_T^{(q)}-\lambda_1^0\right|\overset{\mathbb{P}}{\to}0$ we need the following event to occur with sufficiently high probability, $\mathbf{W}=\left\{\text{For } l=1\,\exists r\in\widehat{\mathcal{I}} \text{ and } k=1,\ldots,K \text{ s.t. } r_{r,k}^{\diamond}=T_l^0\right\}$. Note that

$$\mathbf{W}^c=\left\{T_1^0 \text{ not sample in } K \text{ draws from } T_1^0-m_T+1,\ldots,T_1^0 \text{ without replacement}\right\}.$$

Thus,

$$\begin{aligned}1-\mathbb{P}\left(\mathbf{W}^c\right)&=1-\frac{m_T-1}{m_T}\times\frac{m_T-2}{m_T-1}\times\cdots\times\frac{m_T-K}{m_T-K+1}\\&=1-\frac{m_T-K}{m_T}\\&\to 0,\end{aligned}$$

only if $K=O\left(a_Tm_T\right)$ with $a_T\in(0,1]$, such that $a_T\to1$. Note that $K\leq m_T$ by construction. The same argument applies for $l=2,\ldots,m_0$. $\square$