Change-Point Analysis of Time Series with Evolutionary Spectra*

Alessandro Casini[†] Pierre Perron[‡]
University of Rome Tor Vergata Boston University

7th October 2023

Abstract

This paper develops change-point methods for the spectrum of a locally stationary time series. We focus on series with a bounded spectral density that change smoothly under the null hypothesis but exhibits change-points or becomes less smooth under the alternative. We address two local problems. The first is the detection of discontinuities (or breaks) in the spectrum at unknown dates and frequencies. The second involves abrupt yet continuous changes in the spectrum over a short time period at an unknown frequency without signifying a break. Both problems can be cast into changes in the degree of smoothness of the spectral density over time. We consider estimation and minimax-optimal testing. We determine the optimal rate for the minimax distinguishable boundary, i.e., the minimum break magnitude such that we are able to uniformly control type I and type II errors. We propose a novel procedure for the 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.

JEL Classification: C12, C13, C22.
Keywords: Change-point, Locally stationary, Frequency domain, Minimax-optimal test.

*We thank an Associate Editor and two referees for helpful comments. We also 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. Early contributions include, among others, Hinkley (1971), Yao (1987), Andrews (1993), Horváth (1993) and Bai and Perron (1998), who assume the presence of a single or multiple change-points in the parameters 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 about 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), Neumann and von Sachs (1997), Hörmann and Kokoszka (2010), Dette, Preuß and Vetter (2011), Zhang and Wu (2012), Panaretos and Tavakoli (2013), Aue, Nourinho and Hormann (2015) and van Delft and Eichler (2018)]. Several authors have extended the stationarity tests originally introduced by Priestley and Subba Rao (1969), and further developed by, e.g., Dwivedi and Subba Rao (2010), Jentsch and Subba Rao (2015) and Bandyopadhyay, Carsten and Subba Rao (2017), to these settings. In the context of local stationarity, Paparoditis (2009) proposed a test based on comparing a local estimate of the spectral density to a global estimate and Preuß, Vetter and Dette (2013) proposed a test for stationarity using empirical process theory. In the context of functional time series, tests for stationarity were considered by Horváth, Kokoszka and Rice (2014) and Aue, Rice and Sönmez (2018) using time-domain methods, and by Aue and van Delft (2020) and van Delft, Characiejus and Dette (2018) using frequency-domain methods.

There is wide agreement in empirical work that besides breaks in the mean the detection of breaks in the variance or the correlation structure of a time series is of importance. For example, the discrimination between regimes of high and low asset volatility is of central interest in finance and the detection of changes in the parameters of an autoregressive process is important to build superior forecasting procedures. In addition, discerning the type of the changes (continuous or abrupt) is useful in applied work. On the one hand, if one assumes local stationarity but the true data-generating process involves structural breaks, the parameter estimates may be severely biased and inference may be misleading. On the other hand, if one assumes a structural break model but the parameters actually change gradually then similar issues may arise. A general approach that allows for both continuous changes as well as breaks is needed to avoid these issues. Therefore, the detection of breaks in an otherwise locally stationary time series is important.

We develop inference methods about the changes in the degree of smoothness of the spectrum of a locally stationary time series, and hence, about change-points in the spectrum as a special case. The key parameter is the regularity exponent that governs how smooth the path of the local spectral density is over time. We address two local problems. The first is the detection of discontinuities (or breaks) in the spectrum at some unknown date 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. For example, the spectrum becomes rougher over a short time period, meaning that the paths are less smooth as quantified by the regularity exponent. This can occur for a stationary process whose parameters start to evolve smoothly according to Lipschitz continuity, or for a locally stationary process with Lipschitz parameters that change to continuous but non-differentiable functional parameters. For example, the volatility of high-frequency stock prices or of other macroeconomic variables is known to become rougher (i.e., less smooth) without signifying a structural break after central banks' official announcements, especially in periods of high market uncertainty. In seismology, earthquakes are made up of several seismic waves that arrive at different times and so changes in the smoothness properties of each wave is important for locating the epicenter and identifying what materials the waves have passed through. We consider minimax-optimal testing and estimation for both problems, following 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. These results are different from the developments on minimax optimality obtained recently in the statistics literature for the classical change-point problem where the mean of the series is piecewise constant [see, e.g., Liu, Gao and Samworth (2021) and Verzelen, Fromont, Lerasle and Reynaud-Bouret (2020)].

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 exceptions [Müller (1992), Spokoiny (1998), Müller and Stadtmuller (1999), Wu and Zhao (2007) and Bibinger, Jirak and Vetter (2017)]. These focused on nonparametric regression and high-frequency volatility, and considered time-domain methods while we consider frequency-domain methods. This adds a difficulty in that, e.g., the search for a break or smooth change has to run over two dimensions, the time and frequency indices. Our test statistics are the maximum of local two-sample t-tests based on the local smoothed periodogram. We construct statistics that allow the researcher to test for a change-point in the spectrum at a prespecified frequency and others that allow to detect a break in the spectrum without prior knowledge about the frequency. These test statistics can detect both discontinuous and smooth changes, and therefore are useful for both inference problems discussed above. The asymptotic null distribution follows an extreme value distribution. In order to derive this result, we first establish several asymptotic results, including bounds for higher-order cumulants and spectra of locally stationary processes. These results are complementary to some in Dahlhaus (1997), Paparoditis (2009), Panaretos and Tavakoli (2013), Aue and van Delft (2020) and Casini (2023), and extend some classical frequency-domain results for stationary processes [e.g., Brillinger (1975)] to locally stationary processes.

Change-point problems have also been studied in the frequency-domain in several fields, though with less generality. Adak (1998) investigated the detection of change-points in piecewise stationary time series by looking at the difference in the power spectral density for two adjacent regimes. She compared several distance metrics such as the Kolmogorov-Smirnov, Crámer-Von Mises and 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. Dette, Wu and Zhou (2019) considered testing for change-points in the autocorrelation coefficient at some pre-specified lag k while allowing for local stationarity under the null hypothesis. They proposed a time-domain method based on a CUSUM-type test. To the extent that the autocorrelation at a given lag is only one of the features contained in the spectrum, our setting is more general. Zhang (2016) considered testing for and estimating change-points in the mean of a piecewise locally stationary time series. This is a different problem from ours since the spectrum involves the second-order properties which are complex to study. Yu, Chatterjee and Xu (2022) studied change-point estimation in piecewise polynomials of general degrees. Preuß, Puchstein and Dette (2015) considered the detection and estimation of change-points in the autocovariance function but they required stationarity under the null hypothesis. Several authors considered methods based on segmenting the wavelet spectrum for piecewise stationary time series [see, e.g., Barigozzi, Cho and Fryzlewicz (2018) and Cho and Fryzlewicz (2012; 2017)]. Outside of the wavelet context other contributions are Kirch, Muhsal and Ombao (2015) and Schröder and Ombao (2019). Vogt and Dette (2015) investigated the detection of gradual changes in a locally stationary time series. Our contribution is different since we provide a general change-point analysis about the time-varying spectrum of a time series and establish the relevant asymptotic theory of the proposed test statistics under both the null and alternative hypotheses.

We also address the problem of estimating the change-points, allowing their number to increase with the sample size and the distance between change-points to shrink to zero. We propose a procedure based on a wild sequential top-down algorithm that exploits the idea of bisection combining it with a wild resampling technique similar to the one proposed by 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 (2021b) and Casini and Perron (2020)]. We verify the performance of our methods via simulations which show their benefits. The advantage of using frequency-domain methods to detect change-points is that they do not require to make assumptions about the data-generating process under the null hypothesis 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 good power only against some specific alternatives. For example, tests for changes in the volatility do not have power for changes in the dependence and vice versa. Our methods are readily available for use in many fields such as speech processing, biomedical signal processing, seismology, failure detection, economics and finance. It can also be used as a pre-test before constructing the recently introduced double kernel long-run variance estimator that accounts more flexibly for nonstationarity [cf. Casini (2023), Casini and Perron (2023a) and Casini, Deng and Perron (2023)].

The rest of the paper is organized as follows. Section 2 introduces the statistical setting and the hypothesis testing problems. Section 3 presents the test statistics and states their null limit distributions. Section 4 addresses the consistency of the tests and their minimax optimality. Section 5 discusses the estimation of the change-points while Section 6 provides details for the implementation of the methods. The results of some Monte Carlo simulations are presented in Section 7. An empirical application is presented in Section 8. Section 9 reports brief concluding comments. An online supplement [cf. Casini and Perron (2023b)] contains additional theoretical and empirical results, and all mathematical proofs.

2 Statistical Environment and the Testing Problems

Section 2.1 introduces the statistical setting and Section 2.2 presents the hypotheses testing problems. We work in the frequency-domain under the locally stationary framework introduced by Dahlhaus (1997) who formalized the ideas of Priestley (1965). Casini (2023) extended his framework to allow for discontinuities in the spectrum which then results in a segmented locally stationary process. This corresponds to the relevant process under the alternative hypothesis of breaks in the spectrum. Since local stationarity is a special case of segmented local stationarity we begin with the latter. We use an infill asymptotic setting whereby we rescale the original discrete time horizon $[1,\, T]$ by dividing each t by T.

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. Let $i \triangleq \sqrt{-1}$. We use the notation $\overline{A}$ for the complex conjugate of $A \in \mathbb{C}$.

Definition 2.1. A sequence of stochastic processes $\{X_{t,T}\}_{t=1}^{T}$ is called segmented locally stationary

(SLS) with m_0+1 regimes, transfer function A^0 and trend μ if there exists a representation

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

for $j=1,\ldots,\,m_0+1$, where by convention $T_0^0=0$ and $T_{m_0+1}^0=T$ ($\mathcal{T}\triangleq\{T_1^0,\,\ldots,\,T_{m_0}^0\}$), and the following holds:

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

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

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

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

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

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

(iii) $\mu_{\cdot}\left(\cdot\right)$ is piecewise continuous where the change-points are λ_j^0 ($j=1,\ldots,\,m_0$).

The smoothness properties of A in u guarantees that $X_{t,T}$ has a piecewise locally stationary behavior. This means that $X_{t,T}$ is locally stationary in each segment where the notion of local stationarity is as introduced by Dahlhaus (1997). We refer to Casini (2023) 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. In particular, Casini (2023) also defined (and worked with) the covariance between observations belonging to different segments whereas previous works considered only the covariance between observations belonging to the same segment, thereby using smoothness which simplifies the analysis.

2.2 The Testing Problems

We focus on time-varying spectra that are bounded, thereby excluding unit root, and long memory processes. Unit roots or trending processes can be handled by, for example, taking first differences or using other de-trending techniques. For some $D<\infty$, we consider the following class of time-varying spectra under the null hypothesis,

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

At all continuity points u, $f(u,\,\omega)$ is related to $A(u,\,\omega)$ by the relationship $f(u,\,\omega)=|A(u,\,\omega)|^2$. The key parameter of the testing problem under the null hypotheses is $\theta>0$. This 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 $\theta>0$ and f being differentiable [see Dahlhaus (1996)]. The latter is the setting that we consider under the null hypothesis. To avoid redundancy, we do not require differentiability directly for the functions in $\boldsymbol{F}(\theta,\,D)$ since below we assume that the transfer function $A(u,\,\omega)$ is differentiable in u which in turn implies that f is differentiable in u. Since most of the applied work concerning local stationarity relies on Lipschitz continuity (i.e., $\theta=1$), our specification under the null hypothesis is more general.

We now discuss features that are relevant under the alternative hypothesis. Our focus is on (i) discontinuities of f in u which corresponds to $\theta=0$ and (ii) decreases in the smoothness of the trajectory $u\mapsto f(u,\,\omega)$ for each ω which corresponds to a decrease in θ. We shall refer to a series affected by the changes of case (ii) as "becoming more rough or less smooth". Both cases refer to the properties of the spectral density and thus to the second-order properties of $X_{t,T}$.

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 $\theta>0$ to $\theta'\in(0,\,\theta)$ after some λ_b^0 for some period of time and some ω; i.e., the spectrum becomes rougher after some $\lambda_b^0\in(0,\,1)$ for some time period before returning to θ-smoothness. The case of an increase in θ is technically more complex to handle (see Section 4). As an example, consider a locally stationary AR(1),

$$X_{t,T}=a(t/T)\,X_{t-1,T}+\sigma(t/T)\,e_t,\qquad t=1,\ldots,\,T,$$

where $a:\,[0,\,1]\to(-1,\,1)$ and $\sigma:\,[0,\,1]\to\mathbb{R}_+$ are functional parameters satisfying a Lipschitz condition and $\{e_t\}$ is an i.i.d. mean-zero sequence. Additionally, if $\sup_{u\in[0\,1]}|\sigma(u)|<\infty$ and the initial condition satisfies some regularity condition, then $f(u,\,\omega)$ is uniformly bounded and $\theta=1$. Problem (ii) refers to either $a(\cdot)$ or $\sigma(\cdot)$, or both, becoming less smooth, i.e., we have a change from $\theta=1$ (since the parameters satisfy a Lipschitz condition) to some $\theta'\in(0,\,1)$. In this case, $X_{t,T}$ becomes an AR(1) process with functional parameters that are still continuous but less smooth and may not be differentiable.

Case (i) 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 and can be of considerable interest in several fields even though it requires larger sample sizes than problem (i). For example, the changes in the volatility dynamics of financial or macroeconomic variables over time. Bibinger et al. (2017) provided evidence that the volatility of stock prices can change substantially its path properties after a press release following a meeting of the Federal Open Market Committee; in particular, its path may become more rough. Since $f(u, \omega)$ is a smooth function of the parameters of the data-generating process, if $\sigma(u)$ becomes more rough, then so does $f(u, \omega)$ becomes more rough as u varies. Case (ii) can also be relevant in seismology since the study of the path properties of the seismic waves is important for locating the epicenter of an earthquake. We show below that our tests are consistent and have minimax optimality properties for both cases (i) and (ii). Note that case (ii) is a local problem. In this paper, we do not consider more global problems where for example the spectrum is such that a fall in θ to $\theta' \in (0, \theta)$ occurs on $(\lambda_b^0, 1]$. This represents a continuous change in the smoothness of the spectrum that persists until the end of the time interval. Different test statistics are needed for this case, as will be discussed later.

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 spectrum can change over a short time without indicating a break. 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 as indicated below. To address the minimax-optimal testing [cf. Ingster (1993)], we now introduce the testing problems (i) and (ii) and defer a more general treatment to Section 4.

Testing Problem for Case (i)

Given the discussion above, 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 class of alternative hypotheses:

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

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

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

Observe that $\mathcal{H}_1^{\mathrm{B}}$ requires at least one break but allows for multiple breaks even across different ω. For the testing problem (2.6), we establish the minimax-optimal rate of convergence of the tests suggested [see Ch. 2 in Ingster and Suslina (2003) for an introduction]. A conventional definition is the following. For a nonrandomized test ψ that maps a sample $\{X_{t,T}\}$ 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

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

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 it is to distinguish between $\mathcal{H}_0$ and $\mathcal{H}_1^{\mathrm{B}}$ but we may incur at the same time a larger type II error $\beta_\psi(\theta, b_T)$. The optimal value b_T^{opt}, labelled 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 $b_T \geq b_T^{\mathrm{opt}}$ is called minimax-optimal.

Minimax-optimality has been considered in other change-point problems. Loader (1996) and Spokoiny (1998) considered the nonparametric estimation of a regression function with break size fixed. Bibinger et al. (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-domain. Another difference from previous work is that we do not deal with i.i.d. observations; hence, we cannot use the same approach to derive the minimax bound as in Bibinger et al. (2017) because their information-theoretic reductions exploit independence. We need to rely on approximation theorems [cf. Berkes and Philipp (1979)] to establish that our statistical experiment is asymptotically equivalent in a strong Le Cam sense to a high dimensional signal detection problem. This allows us to derive the minimax bound using classical arguments based on the results in Ingster and Suslina (2003), Ch. 8. The relevant results are stated in Section 4.

Testing Problem for Case (ii)

We consider alternative hypotheses where f is less smooth than under $\mathcal{H}_0$, including the case of breaks as a special case. Suppose that under $\mathcal{H}_0$ the spectrum $f(u, \omega)$ is differentiable in both arguments and behaves until time $T\lambda_b^0$ as specified in $\boldsymbol{F}(\theta, D)$ for some $\theta > 0$ 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 $\mathcal{H}_1^{\mathrm{S}}$ (to be defined below) are well-separated. To this end, define for some function g_u with $u \in [0, 1]$, $\Delta_h^{\theta'} g_u = (g_{u+h} - g_u) / |h|^{\theta'}$ for $h \in [-u, 1-u]$. The set of possible alternatives is then defined as

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

where $m_T \to \infty$ as $T \to \infty$ such that $m_T^{-1}T^{\epsilon} \to 0$ with $\epsilon > 0$. The requirement $|h| \leq 2m_T/T$ means that the exceedance period is at least two blocks of length m_T. This ensures that a test statistic looking at the difference in the spectral density for two adjacent blocks is able to detect roughness, even when this occurs close to the starting sample point. This leads to the following testing problem,

$$\begin{aligned} &\mathcal{H}_0 : \{f(u, \omega)\}_{u\in[0,1]} \in \boldsymbol{F}(\theta, D) \\ &\mathcal{H}_1^{\mathrm{S}} : \exists \lambda_b^0 \in (0, 1) \text{ and } \omega_0 \in [-\pi, \pi] \text{ with } \left\{ \{f(u, \omega)\}_{u\in[0,1],\, \omega\in[-\pi,\pi]} \right\} \in \boldsymbol{F}'_{1,\lambda_b^0,\omega_0}(\theta, \theta', b_T, D). \end{aligned} \tag{2.7}$$

Note that $\mathcal{H}_1^{\mathrm{S}}$ allows for multiple changes. $\mathcal{H}_1^{\mathrm{B}}$ is a special case of $\mathcal{H}_1^{\mathrm{S}}$ since it can be seen as the limiting case of $\mathcal{H}_1^{\mathrm{S}}$ as $\theta' \to 0$. 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 at least to remove the set of functions in $\boldsymbol{F}(\theta, D)$ from those in $\boldsymbol{F}(\theta', D)$. Still, as discussed by Ingster and Suslina (2003), this would not be enough since 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 $\mathcal{H}_1^{\mathrm{S}}$. These are rough enough so as not to be close to functions in $\boldsymbol{F}(\theta, D)$. This is captured by the requirement that the difference 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 et al. (2017) for similar discussions in different contexts.

3 Tests for Changes in the Spectrum and Their Limiting Distributions

Section 3.1 introduces the test statistics while Section 3.2 presents the results concerning their asymptotic distributions under the null hypothesis. These results apply also to the case of smooth alternatives which we discuss formally in Section 4.

3.1 The Test Statistics

We first define the quantities needed to define the tests. Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a data taper with $h(x) = 0$ for $x \notin [0, 1)$,

$$H_{k,T}(\omega) = \sum_{s=0}^{T-1} h(s/T)^k \exp(-i\omega s),$$

and (for n_T even),

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

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

where $I_{L,h,T}(u, \omega)$ (resp., $I_{R,h,T}(u, \omega)$) is the local periodogram over a segment of length $n_T \rightarrow \infty$ that uses observations to the left (resp., right) of $\lfloor Tu \rfloor$. $I_{L,h,T}(u, \omega)$ (resp., $I_{R,h,T}(u, \omega)$) is a near unbiased estimator of $f(u - n_T/T, \omega)$ (resp., $f(u + n_T/T, \omega)$). We allow for a data taper since one may want to put more weight on observations that are closer to u. The smoothed local periodogram is defined as

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

with $f_{R,h,T}(u, \omega)$ defined similarly to $f_{L,h,T}(u, \omega)$ but with $I_{R,h,T}(u, \omega)$ in place of $I_{L,h,T}(u, \omega)$, where $W_T(\omega)$ $(-\infty < \omega < \infty)$ is a family of weight functions of period 2π,

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

with $b_{W,T}$ a bandwidth and $W(\beta)$ $(-\infty < \beta < \infty)$ a fixed function. We define

$$\widetilde{f}_{L,r,T}(\omega) = M_{S,T}^{-1} \sum_{j \in \mathbf{S}_r} f_{L,h,T}(j/T, \omega) \quad \text{and} \quad \widetilde{f}_{R,r,T}(\omega) = M_{S,T}^{-1} \sum_{j \in \mathbf{S}_r} f_{R,h,T}(j/T, \omega),$$

where

$$\begin{aligned} \mathbf{S}_r = \{ & rm_T - m_T/2 + \lfloor n_T/2 \rfloor + 1,\, rm_T - m_T/2 + \lfloor n_T/2 \rfloor + 1 + m_{S,T}, \\ & \ldots,\, rm_T + \lfloor n_T/2 \rfloor + 1 + m_{S,T} M_{S,T}/2 \}, \end{aligned}$$

with $m_{S,T} = \lfloor m_T^{1/2} \rfloor$ and $M_{S,T} = \lfloor m_T/m_{S,T} \rfloor$. $\widetilde{f}_{a,r,T}(\omega)$ $(a = L, R)$ denotes the average local spectral density around time rm_T computed using $f_{a,h,T}(j/T, \omega)$ where $r = 1, \ldots, M_T = \lfloor T/m_T \rfloor - 1$. We do not use all the m_T local spectral densities $f_{a,h,T}(j/T, \omega)$ $(a = L, R)$ in the block r but only those separated by $m_{S,T}$ points. Thus, $\mathbf{S}_r$ is a subset of the indices in the block r. We need to consider a sub-sample of $f_{a,h,T}(j/T, \omega)$'s $(a = L, R)$ because there is strong dependence among the adjacent terms, e.g., $f_{a,h,T}(j/T, \omega)$ and $f_{a,h,T}((j+1)/T, \omega)$ $(a = L, R)$. A large deviation between $\widetilde{f}_{L,r,T}(\omega)$ and $\widetilde{f}_{R,r+1,T}(\omega)$ suggests the presence of a break in the spectrum close to time $(r+1)\, m_T$ at frequency ω. Note that the latest observation used in the construction of $\widetilde{f}_{L,r,T}$ is $X_{rm_T + m_T/2 + \lfloor n_T/2 \rfloor + 1}$ while the earliest observation used in the construction of $\widetilde{f}_{R,r+1,T}$ is $X_{rm_T + m_T/2 + \lfloor n_T/2 \rfloor + 2}$. This shows that there is no overlapping in the time points used in the construction of $\widetilde{f}_{L,r,T}$ and $\widetilde{f}_{R,r+1,T}$. The reason behind this is that in order to maximize power $\widetilde{f}_{L,r,T}$ and $\widetilde{f}_{R,r+1,T}$ should not use common observations, otherwise the effect of the common observations to the difference in the averages would offset the effect of the change-point.

Let

$$\begin{aligned} \mathbf{S}_{r,+}(j) = \{ \{ & rm_T - m_T/2 + \lfloor n_T/2 \rfloor + 1,\, rm_T - m_T/2 + \lfloor n_T/2 \rfloor + 1 + \widetilde{m}_{S,T}, \\ & \ldots,\, rm_T + \lfloor n_T/2 \rfloor + 1 + \widetilde{m}_{S,T} \widetilde{M}_{S,T}/2 \} / \{ \ldots,\, rm_T - m_T/2 + 1 + \lfloor n_T/2 \rfloor + \widetilde{m}_{S,T} j \} \}, \\ \mathbf{S}_{r,-}(j) = \{ \{ & rm_T - m_T/2 + \lfloor n_T/2 \rfloor + 1,\, rm_T - m_T/2 + \lfloor n_T/2 \rfloor + 1 + \widetilde{m}_{S,T}, \\ & \ldots,\, rm_T + \lfloor n_T/2 \rfloor + 1 + \widetilde{m}_{S,T} \widetilde{M}_{S,T}/2 \} / \{ \ldots,\, rm_T - m_T/2 + 1 + \lfloor n_T/2 \rfloor - \widetilde{m}_{S,T} j \} \}, \end{aligned}$$

where $\widetilde{m}_{S,T} = \lfloor m_T^{1/3} \rfloor$ and $\widetilde{M}_{S,T} = \lfloor m_T/\widetilde{m}_{S,T} \rfloor$. Define $\widehat{\sigma}_{L,r}^2(\omega) = \sum_{j=-\widetilde{M}_{S,T}+1}^{\widetilde{M}_{S,T}-1} K_1(b_{1,T} j)\, \widehat{\Gamma}_r(j)$ where

$$\widehat{\Gamma}_r(j) = \begin{cases} \widetilde{M}_{S,T}^{-1} \sum_{t \in \mathbf{S}_{r,+}(j)} \widehat{f}_{L,h,T}(t/T, \omega)\, \widehat{f}_{L,h,T}((t - j\widetilde{m}_{S,T})/T, \omega), & j \geq 0 \\ \widetilde{M}_{S,T}^{-1} \sum_{t \in \mathbf{S}_{r,-}(j)} \widehat{f}_{L,h,T}(t/T, \omega)\, \widehat{f}_{L,h,T}((t + j\widetilde{m}_{S,T})/T, \omega), & j < 0 \end{cases},$$

and $\widehat{f}_{L,h,T}(j/T,\,\omega) = f_{L,h,T}(j/T,\,\omega) - \widetilde{f}_{L,r,T}(\omega)$ for $j \in \mathbf{S}_r$. The quantity $\widehat{\sigma}^2_{L,r}(\omega)$ is a local long-run variance estimator where K_1 is a kernel and $b_{1,T}$ is the associated bandwidth.

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 detects 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 at which frequency the spectrum is discontinuous. We begin with the following test statistic,

$$\mathrm{S}_{\max,T}(\omega) \triangleq \max_{r=1,\dots,\,M_T-2} \left| \frac{\widetilde{f}_{L,r,T}(\omega) - \widetilde{f}_{R,r+1,T}(\omega)}{\widehat{\sigma}_{L,r}(\omega)} \right|, \qquad \omega \in [-\pi,\,\pi]\,. \tag{3.1}$$

Test statistics of the form of (3.1) were also used in the time-domain in the context of non-parametric change-point analysis under a less general framework [cf. Eichinger and Kirch (2018), Bibinger et al. (2017) and Wu and Zhao (2007)], high-dimensional setting [cf. Chen, Wang and Wu (2022)] and forecasting [cf. Casini (2018)]. The derivation of the null distribution uses a (strong) invariance principle for nonstationary processes [see, e.g., Wu (2007) and Wu and Zhou (2011)].

The test statistic $\mathrm{S}_{\max,T}(\omega)$ aims at detecting a break in the spectrum at some given frequency ω. An alternative would be to consider a double-sup statistic which takes the maximum over $\omega \in [-\pi,\,\pi]$. Theorem S.A.4 in the supplement shows that $I_{h,L,T}(u,\,\omega_j)$ and $I_{h,L,T}(u,\,\omega_k)$ are asymptotically independent if $2\omega_j,\,\omega_j \pm \omega_k \not\equiv 0\,(\mathrm{mod}\,2\pi)$.[1] However, the smoothing over frequencies introduces short-range dependence over ω. Due to this short-range dependence, we cannot consider the maximum over all frequencies in Π because the statistics would not be independent. Thus, we specify a framework based on an infill procedure over the frequency-domain $[-\pi,\,\pi]$ by assuming that there are n_ω frequencies $\omega_1, \dots, \omega_{n_\omega}$, with $\omega_1 = -\pi$ and $\omega_{n_\omega} = \pi - \epsilon$, $\epsilon > 0$, and $|\omega_j - \omega_{j+1}| = O(n_\omega^{-1})$ for $j = 1, \dots, n_\omega - 2$. Assume that $n_\omega \to \infty$ as $T \to \infty$. Let $\Pi \triangleq \{\omega_1, \dots, \omega_{n_\omega}\}$. The maximum is taken over the following set of frequencies

$$\Pi' \triangleq \{\omega_1,\,\omega_{2+\lfloor n_T b_{W,T}\rfloor}, \dots, \omega_{n_\omega - \lfloor n_T b_{W,T}\rfloor - 1},\,\omega_{n_\omega}\}.$$

Let $n'_\omega = \lfloor n_\omega / (\lfloor n_T b_{W,T}\rfloor + 1)\rfloor$ with $n'_\omega \to \infty$. Note that $\Pi' \subset \Pi$. The double-sup statistic is

$$\mathrm{S}_{\mathrm{D}\max,T} \triangleq \max_{\omega_k \in \Pi'} \sqrt{\log(M_T)}(M_{S,T}^{1/2}\mathrm{S}_{\max,T}(\omega_k) - \gamma_{M_T}) - \log(n'_\omega)\,, \tag{3.2}$$

where $\gamma_{M_T} = [4\log(M_T) - 2\log(\log(M_T))]^{1/2}$. This double-sup form is a new feature for change-point testing under the frequency-domain.

[1]The notation $2\omega_j,\,\omega_j \pm \omega_k \not\equiv 0\,(\mathrm{mod}\,2\pi)$ means $2\omega_j \not\equiv 0\,(\mathrm{mod}\,2\pi)$ and $\omega_j \pm \omega_k \not\equiv 0\,(\mathrm{mod}\,2\pi)$.

Next, we consider alternative test statistics that are self-normalized such that one does not need to estimate $\sigma_{L,r}^2(\omega)$. We consider the following test statistic,

$$\mathrm{R}_{\max,T}(\omega) \triangleq \max_{r=1,\ldots,M_T-2} \left| \frac{\widetilde{f}_{L,r,T}(\omega)}{\widetilde{f}_{R,r+1,T}(\omega)} - 1 \right|, \tag{3.3}$$

where $\omega \in [-\pi, \pi]$. We can consider a test statistic corresponding to $\mathrm{S}_{\mathrm{Dmax},T}$ defined by

$$\mathrm{R}_{\mathrm{Dmax},T} \triangleq \max_{\omega_k \in \Pi'} \sqrt{\log(M_T)}(M_{S,T}^{1/2}\mathrm{R}_{\max,T}(\omega_k) - \gamma_{M_T}) - \log(n'_\omega).$$

3.2 The Limiting Distribution Under the Null Hypothesis

Denote by $\kappa_{X,t}(k_1, \ldots, k_{r-1})$ the time-t cumulant of order r of $(X_{t+k_1,T}, \ldots, X_{t+k_{r-1},T}, X_{t,T})$ with $r \leq p$.

Assumption 3.1. *(i) $\{X_{t,T}\}$ is a mean-zero locally stationary process (i.e., $m_0 = 0$); (ii) $A(u, \omega)$ is 2π-periodic in ω and the periodic extensions are differentiable in u and ω with uniformly bounded derivative $(\partial/\partial u)(\partial/\partial\omega)A(u, \omega)$; (iii) g_4 is continuous.*

Assumption 3.1 requires $\{X_{t,T}\}$ to be locally stationary. Without loss of generality, we assume that $\{X_{t,T}\}$ has zero mean. All results go through when the mean is non-zero when using demeaned series. The differentiability of $A(u, \omega)$ implies that $f(u, \omega)$ is also differentiable. This means that under the null hypothesis we require $f(u, \omega)$ to be differentiable in u and to have some regularity exponent $\theta > 0$. The differentiability of $A(u, \omega)$ in u can be relaxed at the expense of more complex proofs in the supplement to establish the results on high-order cumulants. Without differentiability, for any $\theta > 0$ the test statistics above follow the same asymptotic distribution as when differentiability holds, though we do not discuss this case formally.

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

$$\begin{aligned}\phi_{w,q} &= \sup_t \left\| X_{t,T} - X_{t,T,\{w\}} \right\|_q \\ &= \sup_t \left\| H_T(t/T, \mathscr{F}_t) - H_T\left(t/T, \mathscr{F}_{t,\{w\}}\right) \right\|_q,\end{aligned} \tag{3.4}$$

where $\mathscr{F}_{t,\{w\}}$ is a coupled version of $\mathscr{F}_t$ with e_w replaced by an i.i.d. copy e'_w. For $\{X_{t,T}\}$ locally stationary, $H_T(\cdot, \mathscr{F}_t)$ is stochastic θ-Hölder continuous, i.e., there exists $C_{T,H} < \infty$ such that

$$\sup_{0 \leq u < u' \leq 1} \frac{\| H_T(u, \mathscr{F}_t) - H_T(u', \mathscr{F}_t) \|}{|u - u'|^{\theta}} \leq C_{T,H}. \tag{3.5}$$

Assume $\Upsilon_{n,q} = \sum_{j=n}^{\infty} \phi_{j,q} < \infty$ for some $n \in \mathbb{Z}$. 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 3.2. *For $q \geq 2$, $\sum_{n=0}^{\infty} n^{l+q-1} (\sum_{j=n}^{\infty} \phi_{j,q}^2)^{1/2} < \infty$ where $l \geq 0$.*

Assumption 3.2 was also used by Shao and Wu (2007) who showed that it is satisfied for many nonlinear time series processes. Using Assumption 3.2, we can give a sufficient condition for the summability of the joint cumulant up to a certain order. The latter is a common assumption in spectral analysis and we use it to establish results on the high-order cumulants and spectra in Section S.A. These are used to obtain the null limiting distribution of the test statistics.

Lemma 3.1. *Let Assumption 3.2 hold and $X_{t,T} \in \mathscr{L}^r$, $r \geq 2$. We have*

$$\sum_{k_1, \ldots, k_{r-1} = -\infty}^{\infty} \left(1 + |k_j|^l\right) \sup_{1 \leq t \leq T} \left| \kappa_{X,t}^{(a_1, \ldots, a_r)}(k_1, \ldots, k_{r-1}) \right| < \infty, \tag{3.6}$$

for $l \geq 0$, $j = 1, \ldots, r-1$, and any r-tuple $a_1, \ldots, a_r$.

Shao and Wu (2007) proved (3.6) for $l = 0$ and $\{X_{t,T}\}$ a stationary process.

Assumption 3.3. *(i) The data taper $h : \mathbb{R} \to \mathbb{R}$ with $h(x) = 0$ for $x \notin [0, 1)$ is bounded and of bounded variation; (ii) The sequence $\{n_T\}$ satisfies $n_T \to \infty$ as $T \to \infty$ with $n_T/T \to 0$; (iii) $W(\beta)$ $(-\infty < \beta < \infty)$ is real-valued, even, of bounded variation, and satisfies $\int_{-\infty}^{\infty} W(\beta) d\beta = 1$; (iv) $b_{1,T} \to 0$ such that $Tb_{1,T} \to \infty$ and $(\widetilde{M}_{S,T} b_{1,T})^{-1/2} M_{S,T}^{1/2} \log T \to 0$, and $K_1(\cdot) : \mathbb{R} \to [-1, 1]$, $K_1(0) = 1$, $K_1(x) = K_1(-x)$, $\forall x \in \mathbb{R}$, $\int_{-\infty}^{\infty} K_1^2(x) dx < \infty$, $K_1(\cdot)$ is continuous at 0 and at all but a finite number of other points, and $\widetilde{K}_1(\omega) = (2\pi)^{-1} \int_{-\infty}^{\infty} K_1(x) e^{-ix\omega} dx$ satisfies $\widetilde{K}_1(\omega) \geq 0$ for all $\omega \in \mathbb{R}$*

Assumption 3.3-(i,ii) are standard in the nonparametric estimation literature while Assumption 3.3-(iii) is also used for spectral density estimation under stationarity [e.g., Brillinger (1975)]. The conditions on $b_{1,T}$ in Assumption 3.3-(iv) are necessary for the consistency of the long-run variance estimator. The class of kernels allowed by Assumption 3.3-(iv) includes popular kernels such as the Truncated, Bartlett, Parzen, Quadratic Spectral (QS) and Tukey-Hanning kernels. For technical reasons inherent to the proofs, we need to assume that the spectral density is strictly positive. Theorem S.A.6 in the supplement shows that the variance of $f_{L,h,T}(u,\,\omega)$ depends on $f(u,\,\omega)$. Thus, the denominator of the test statistic depends on $f(u,\,\omega)$. Assumption 3.4 requires the latter to be bounded away from zero. In practice, if one suspects that at some frequencies $f(u,\,\omega)$ can be close to zero, then one can add a small number $\epsilon_f > 0$ to the denominator of the test statistic to guarantee numerical stability.

Assumption 3.4. $f_- = \min_{u\in[0,\,1],\,\omega\in[-\pi,\,\pi]} f(u,\,\omega) > 0.$

The next assumption ensures that the local spectral density estimates are asymptotically independent when evaluated at some given frequencies (see Theorem S.A.6). It is used to derive the asymptotic null distribution of the double-sup test statistics $\mathrm{S}_{\mathrm{Dmax},T}$ and $\mathrm{R}_{\mathrm{Dmax},T}$.

Assumption 3.5. *Assume that* $2\omega_j,\,\omega_j \pm \omega_k \not\equiv 0 \,(\mathrm{mod}\, 2\pi)$ *for* $\omega_j,\,\omega_k \in \Pi$.

Assumption 3.6. *(i) The sequence* $\{m_T\}$ *satisfies* $m_T \to \infty$ *as* $T \to \infty$, *and*

$$M_{S,T}^{1/2} m_T^{\theta} T^{-\theta} \left(\log\left(M_T\right)\right)^{1/2} + \tau_T^2 \log\left(M_T\right) M_{S,T}^{-1} + M_{S,T} n_T^4 \log\left(M_T\right) T^{-4} + M_{S,T} \left(\log\left(n_T\right)\right)^2 \log\left(M_T\right) n_T^{-2} \to 0; \tag{3.7}$$

(ii) $b_{W,T} \to 0$ *such that* $Tb_{W,T} \to \infty$, $\log\left(M_T\right) M_{S,T} b_{W,T}^4 \to 0$ *and* $\log\left(M_T\right) M_{S,T} (n_T b_{W,T})^{-1} \to 0$.

Part (i) imposes lower and upper bounds on the growth condition of the sequence $\{m_T\}$. The upper bound relates to the smoothness of $A(u,\,\omega)$, the value of θ under the null hypothesis, n_T and the number of summands $M_{S,T}$ in $\widetilde{f}^{*}_{L,r,T}(\omega)$.

Let $\mathscr{V}$ denote a random variable with an extreme value distribution defined by $\mathbb{P}(\mathscr{V} \leq v) = \exp(\,-\pi^{-1/2} \exp(-v))$.

Theorem 3.1. *Let Assumption 3.1-3.4 and 3.6 hold. Under* $\mathcal{H}_0$, $\sqrt{\log\left(M_T\right)}(M_{S,T}^{1/2} \mathrm{S}_{\max,T}(\omega) - \gamma_{M_T}) \Rightarrow \mathscr{V}$ *for any* $\omega \in [-\pi,\,\pi]$.

Theorem 3.1 shows that the asymptotic null distribution follows an extreme value distribution. The derivation of the null distribution uses a (strong) invariance principle for nonstationary processes [see, e.g., and Wu and Zhou (2011)]. The following theorems show that the asymptotic null distribution of the tests $\mathrm{S}_{\mathrm{Dmax},T}$, $\mathrm{R}_{\max,T}(\omega)$ and $\mathrm{R}_{\mathrm{Dmax},T}$ also follows an extreme value distribution, though the additional Assumption 3.5 and the extra factor $\log(n'_\omega)$ are needed for $\mathrm{S}_{\mathrm{Dmax},T}$ and $\mathrm{R}_{\mathrm{Dmax},T}$.

Theorem 3.2. *Let Assumption 3.1-3.6 hold. Under* $\mathcal{H}_0$, $\mathrm{S}_{\mathrm{Dmax},T} \Rightarrow \mathscr{V}$.

Theorem 3.3. *Let Assumption 3.1-3.4 and 3.6 hold. Under* $\mathcal{H}_0$, $\sqrt{\log\left(M_T\right)}(M_{S,T}^{1/2} \mathrm{R}_{\max,T}(\omega_k) - \gamma_{M_T}) \Rightarrow \mathscr{V}$ *and, in addition if Assumption 3.5 holds, then* $\mathrm{R}_{\mathrm{Dmax},T} \Rightarrow \mathscr{V}$.

The limiting distributions in Theorem 3.1-3.3 are pivotal, so that critical values can be obtained without the need to rely on simulations. The tests have power against both breaks (alternative hypothesis (i)) and changes in the smoothness (alternative hypothesis (ii)). Unfortunately, discerning between the two types of alternative hypotheses is a hard technical problem. Although knowing whether the rejection of the null hypothesis is due to a break or a change in the smoothness would be useful, in many cases knowing that there has been some change in the data-generating process is sufficient to modify the estimation and/or inference strategy to account for the change. If the change involves a break, a reasonable approach would be to apply some sample-splitting method for estimation. For example, in the context of long-run variance estimation the existence of a break suggests to modify the double kernel HAC (DK-HAC) estimator to avoid mixing two different regimes [cf. Casini (2023)]. On the other hand, for a change in the smoothness a modification of a standard nonparametric kernel smoothing could be enough. However, using a sample-splitting technique even when there is a change in the smoothness would be robust to the change and would result in better estimation and inference. In general, treating a change as a break and applying some sample-splitting method would be valid even when the rejection was due to a change in the smoothness, though it may not be efficient.

The property of being robust to different alternative hypotheses is not specific to our method. It is shared by most of the existing structural break tests. For example, Andrews (1993) showed that structural break tests also have some power against some forms of smoothly-varying parameters. This property was seen as a positive feature in the structural break literature. Following the same reasoning, the property of being robust to breaks as well as changes in the smoothness can actually be seen as a virtue of our method.

4 Consistency and Minimax Optimal Rate of Convergence

In this section, we discuss the consistency and minimax-optimal lower bound for the testing problem (2.7) (i.e., case (ii)). The discussion also covers the testing problem (i) since $\mathcal{H}_1^{\mathrm{B}}$ can be seen as the limiting case of $\mathcal{H}_1^{\mathrm{S}}$ as $\theta' \to 0$. We assume that $X_{t,T}$ is segmented locally stationary with transfer function $A(u, \omega)$ satisfying the following smoothness properties.

Assumption 4.1. *(i)* $\{X_{t,T}\}$ *is a mean-zero segmented locally stationary process; (ii)* $A(u, \omega)$ *is twice continuously differentiable in* u *at all* $u \neq \lambda_j^0$ $(j = 1, \ldots, m_0 + 1)$ *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 *at* $u = \lambda_j^0$ $(j = 1, \ldots, m_0 + 1)$ *with uniformly bounded derivatives* $(\partial/\partial_- u) A(u, \cdot)$ *and* $(\partial^2/\partial_- u^2) A(u, \cdot)$.

Assumption 4.2. *(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.*

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

Theorem 4.1. *Let Assumption 3.2-3.3, 4.1-4.2 and $f_{-}(\omega_0) > 0$ hold. Consider either set of hypotheses $\{\mathcal{H}_0, \mathcal{H}_1^{\mathrm{B}}\}$ with $\theta' = 0$ or $\{\mathcal{H}_0, \mathcal{H}_1^{\mathrm{S}}\}$ with $0 < \theta' < \theta$. Then, for*

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

we have $\lim_{T\to\infty} \inf_{\psi} \gamma_{\psi}(\theta, b_T) = 1$.

The theorem implies the need for

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

otherwise there cannot exist a minimax-optimal test yielding $\lim_{T\to\infty} \inf_{\psi} \gamma_{\psi}(\theta, b_T) = 0$. Note that the lower bound does not depend on ω. In Theorem 4.2 we establish a corresponding upper bound. From the lower and upper bounds we deduce the optimal rate for the minimax distinguishable boundary. We can also derive tests based on b_T^{opt}. For example, using the test statistic (3.1) for $\{\mathcal{H}_0, \mathcal{H}_1^{\mathrm{B}}\}$ we obtain the following test $\psi^* : \psi^*(\{X_{t,T}\}) = 1$ if $\mathrm{S}_{\max,T}(\omega) \geq 2D^* \sqrt{\log(M_T^*)/m_T^*}$ for $\omega \in [-\pi, \pi]$ where $D^* > 2$, $m_T^* = (\sqrt{\log(M_T^*)} T^{\theta}/D)^{\frac{2}{2\theta+1}}$ and $M_T^* = \lfloor T/m_T^* \rfloor$. Hence, in order to construct such a test, we need knowledge of θ under $\mathcal{H}_0$. We discuss this in Section 6.

Next, we establish the optimal rate for minimax distinguishability. Note that either alternatives $\mathcal{H}_1^{\mathrm{B}}$ or $\mathcal{H}_1^{\mathrm{S}}$ allows for multiple breaks. The following results require further restrictions on the relation between n_T and m_T.

Theorem 4.2. *Let Assumption 3.2-3.3, 4.1-4.2 hold. Consider either alternative hypotheses $\mathcal{H}_1^{\mathrm{B}}$ with $\theta' = 0$ and $\lambda_j^0 < \lambda_{j+1}^0$ for $j = 1, \ldots, m_0$, or $\mathcal{H}_1^{\mathrm{S}}$ with $0 < \theta' < \theta$. If*

$$\left(\sqrt{\log(M_T^*)/m_T^*}\right)^{-1} \left((n_T/T)^2 + \log(n_T)/n_T + b_{W,T}^2\right) \to 0, \tag{4.1}$$

and

$$b_T^* > \left(4D^* \sup_{u\in[0,1]} f(u, \omega_0) + 2\right)^{-\frac{\theta-\theta'}{2\theta+1}} (T/\log(M_T))^{-\frac{\theta-\theta'}{2\theta+1}} D^{-\frac{2\theta'+1}{2\theta+1}}, \tag{4.2}$$

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 distinguished from a break of magnitude smaller than b_T^{opt} because the change from 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 et al. (2017) for the volatility of a Itô semimartingale. The theorem suggests that knowledge of the frequency ω_0 at 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. This may be of interest for economic and financial time series since they often exhibit a break simultaneously at high and low frequencies. We leave this to future research.

5 Estimation of the Change-Points

We now discuss the estimation of the break locations for the case of discontinuities in the spectrum (i.e., $\mathcal{H}_1^{\mathrm{B},m_0}$ where m_0 is the number of breaks, recall Definition 2.1). The same estimator is valid for the locations of the smooth changes as under $\mathcal{H}_1^{\mathrm{S}}$. For the latter case we later provide intuitive remarks about the consistency result and the conditions needed for it. We first consider the case of a single break (i.e., $\mathcal{H}_1^{\mathrm{B},1}$) and then present the results for the case of multiple breaks (i.e., $\mathcal{H}_1^{\mathrm{B},m_0}$).

5.1 Single Break Alternatives $\mathcal{H}_1^{\mathrm{B},1}$

Let

$$\mathrm{D}_{r,T}(\omega) \triangleq M_{S,T}^{-1/2} \left| \sum_{j \in \mathbf{S}_{L,r}} f_{L,h,T}(j/T,\, \omega) - \sum_{j \in \mathbf{S}_{R,r}} f_{R,h,T}(j/T,\, \omega) \right|, \qquad \omega \in [-\pi,\, \pi].$$

where

$$\mathbf{S}_{L,r} = \{r - m_T + 1,\, r - m_T + 1 + m_{S,T}, \ldots,\, r - m_T + 1 + m_{S,T} M_{S,T}\},$$
$$\mathbf{S}_{R,r} = \{r + 1,\, r + 1 + m_{S,T}, \ldots,\, r + 1 + m_{S,T} M_{S,T}\},$$

and $r = 2m_T,\, 2m_T + \dot{m}_T,\, 2m_T + 2\dot{m}_T \ldots$ with $\dot{m}_T = m_T/\sqrt{M_{S,T}}$ and $r < (M_T - 1)\, m_T - n_T$. Note that the maximum of the statistics $\mathrm{D}_{r,T}(\omega)$ is a version of $\mathrm{S}_{\max,T}$ that does not involve the normalization. The change-point estimator is defined as

$$T\widehat{\lambda}_{b,T} = \underset{r=2m_T,\, 2m_T+\dot{m}_T,\ldots}{\operatorname{argmax}} \max_{\omega\in[-\pi,\,\pi]} \mathrm{D}_{r,T}(\omega).$$

Recall that we consider the following alternative hypothesis:

$$\mathcal{H}_1^{\mathrm{B},1}: \left\{ f\left(T_b^0/T,\, \omega_0\right) - \lim_{s\downarrow T_b^0} f\left(s/T,\, \omega_0\right) = \delta_T \neq 0, \quad \omega_0 \in [-\pi,\, \pi] \right\}.$$

Note that a break does not need to occur simultaneously at all frequencies for the procedure to work. The break magnitude can be either fixed or converge to zero as specified by the following assumption.

Assumption 5.1. *$\delta_T = \delta \neq 0$ is fixed or $\delta_T \to 0$ and $\delta_T M_{S,T}^{1/2}/\sqrt{\log(T)} \to (0,\, \infty]$.*

Proposition 5.1. *Let Assumption 3.2-3.3-(i-iii), 3.6 and 4.1 with $m_0 = 1$ hold. Under $\mathcal{H}_1^{\mathrm{B},1}$, if δ_T satisfies Assumption 5.1, we have $\widehat{\lambda}_{b,T} - \lambda_b^0 = O_{\mathbb{P}}(m_{S,T}\sqrt{M_{S,T}\log(T)}/(T\delta_T))$.*

We compare the rate of convergence in Proposition 5.1 with that of classical change-point estimators in the piecewise constant mean model. For fixed shifts, the latter rate of convergence is $O_{\mathbb{P}}(T^{-1})$ while for shrinking shifts it 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)].[2] Unlike the classical change-point problem where the mean is piecewise constant, our problem involves a spectrum that can vary smoothly. The latter represents a local problem that cannot be addressed by standard sample-splitting methods. Our method is local in nature and so it is sub-optimal for the classical change-point problem but it is valid for the more general case of a piecewise smooth spectrum. Hence, for fixed shifts, the rate of convergence in our problem is slower. The smallest break magnitude allowed by Proposition 5.1 is $\delta_T = O(\sqrt{\log(T)}/M_{S,T}^{1/2})$. Under this condition the convergence rate for the classical estimator is $O_{\mathbb{P}}(M_{S,T}(T\log(T))^{-1})$ which is faster by a factor $O(m_{S,T}\sqrt{\log(T)})$ than the one suggested by Proposition 5.1. In addition, 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.

Under the smooth alternative $\mathcal{H}_1^{\mathrm{S}}$ the estimator is consistent when θ-regularity is violated only once in the sample and also when the violation occurs in a small interval around λ_b^0 which does not exceed $O(m_{S,T}\sqrt{M_{S,T}\log(T)}/T\delta_T)$. If that interval is longer then this becomes a global problem which cannot be addressed by the estimation method considered in this section. This also relates

[2] See also Verzelen et al. (2020) for recent developments on minimax optimality for change-point estimation in the piecewise constant mean model. They considered as a significance of a change-point a measure that depends on both the break magnitude and the location of the break. They named it the *energy* of the change-point. They established the uniform detection threshold for the energy.

to the discussion in Section 4 that one cannot perfectly separate functions with θ-smoothness from functions with θ'-smoothness such that $\theta' < \theta$.

5.2 Multiple Breaks Alternatives $\mathcal{H}_1^{\mathrm{B},m_0}$

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

$$\mathcal{H}_1^{\mathrm{B},m_0} : \left\{ f\left(T_l^0/T,\,\omega_l\right) - \lim_{s \downarrow T_l^0} f\left(s/T,\,\omega_l\right) = \delta_{l,T} \neq 0, \quad \omega_l \in [-\pi,\,\pi] \text{ for } 1 \leq l \leq m_0 \right\}.$$

We provide a consistency result for both m_0 and the actual locations of the breaks λ_l^0 $(1 \leq l \leq m_0)$. Let j^* be the largest integer such that $2m_T + \left(\dot{M}_T - j^*\right)\dot{m}_T \leq (M_T - 1)\,m_T - n_T$ where $\dot{M}_T = \lfloor T/\dot{m}_T \rfloor$, and $\mathcal{I} \subseteq \left\{2m_T,\, 2m_T + \dot{m}_T, \ldots,\, 2m_T + \left(\dot{M}_T - j^*\right)\dot{m}_T\right\}$ denote a generic index set. One can test for a break at some time index in $\mathcal{I}$ by using the test $\psi^*(\{X_r\}_{r\in\mathcal{I}})$ based on $\max_{\omega_k \in \Pi'} \mathrm{S}_{\max,T}(\omega_k)$ and if the test rejects one can estimate the break location using

$$T\widehat{\lambda}_T(\mathcal{I}) = \underset{r\in\mathcal{I}}{\mathrm{argmax}} \max_{\omega\in[-\pi,\,\pi]} \mathrm{D}_{r,T}(\omega). \tag{5.1}$$

We can 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 may not be efficient. For example, consider the first step of the algorithm in which we test for the first break; this is associated with the largest break magnitude ($\delta_{1,T} > \delta_{l,T}$ for all $l = 2, \ldots, m_0$). If the true break date T_1^0 falls in between two indices in $\mathcal{I}$, say r_1 and $r_2 = r_1 + \dot{m}_T$, then this does not maximize either power or precision of the location estimate because one would need to compare two adjacent blocks exactly separated at T_1^0 but $T_1^0 \notin \mathcal{I}$ since $T_1^0 \in (r_1,\, r_2)$. Hence, 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 for each separation point compute $\mathrm{D}_{r^\diamond,T}(\omega)$ where $r^\diamond \in (r_1,\, r_2)$. We take the maximum value. 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. A natural question is why not take all integers between r_1 and r_2 and compute $\mathrm{D}_{r^\diamond,T}(\omega)$ for each. The reason is that in applications involving high frequency data (e.g., weakly, daily, and so on) that would be highly computationally intensive especially with multiple breaks as one wishes to change $\dot{m}_T$ when searching for an additional break. This procedure exploits the idea of bisection and combines it with a wild resampling technique similar to the one in Fryzlewicz (2014). The latter is characterized by using binary segmentation and drawing a large number of random intervals. Here the idea of using draws of random intervals is applied to the sequential top-down algorithm.

We are now ready to present the algorithm. Guidance as to a suitable choice of K will be given below. Let $v_T \to \infty$ with $v_T/T \to 0$ and $m_T/v_T \to 0$. Consider the test $\psi(\{X_{t,T}\}, \mathcal{I}) = 1$ if $\mathrm{S}_{\mathrm{Dmax},T}(\mathcal{I}) \geq 2D^* \sqrt{\log(M_T^*)/m_T^*}$ where

$$\mathrm{S}_{\mathrm{Dmax},T}(\mathcal{I}) \triangleq \max_{r\in\mathcal{I}} \max_{\omega_k\in\Pi'} \left| \frac{\sum_{j\in\mathbf{S}_{L,r}} f_{L,h,T}(j/T,\,\omega_k) - \sum_{j\in\mathbf{S}_{R,r}} f_{R,h,T}(j/T,\,\omega_k)}{\widehat{\sigma}_{L,r}(\omega_k)} \right|,$$

with D^*, m_T^* and M_T^* as defined in Section 4.

Algorithm 1. *Set* $\widehat{\mathcal{I}} = \{2m_T,\, 2m_T + \dot{m}_T, \ldots,\, 2m_T + \left(\dot{M}_T - j^*\right) \dot{m}_T\}$ *and* $\widehat{\mathcal{T}} = \emptyset$.
(1) For $r \in \widehat{\mathcal{I}} \backslash \{2m_T\}$ *uniformly draw (without replacement)* $K \in \{1, \ldots, \dot{m}_T\}$ *points* $r_k^\diamond$ *from* $\mathbf{I}(r) = \{r - \dot{m}_T + 1, \ldots, r\}$ *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 $\psi(\{X_{t,T}\}, \widehat{\mathcal{I}}) = 0$ *return* $\widehat{\mathcal{T}} = \emptyset$*. Otherwise proceed with step (3).*
(3) Estimate the change-point $T\widehat{\lambda}_T(\widehat{\mathcal{I}})$ *via (5.1) using* $\widehat{\mathcal{I}}$.
(4) Set $\widehat{\mathcal{I}} = \widehat{\mathcal{I}} \backslash \{T\widehat{\lambda}_T(\widehat{\mathcal{I}}) - v_T, \ldots, T\widehat{\lambda}_T(\widehat{\mathcal{I}}) + v_T\}$ *and* $\widehat{\mathcal{T}} = \widehat{\mathcal{T}} \cup \{T\widehat{\lambda}_T(\widehat{\mathcal{I}})\}$*. Return to step (1).*

Finally, arrange the estimated change-points $\widehat{\lambda}_{l,T}$ in $\widehat{\mathcal{T}}$ in chronological order and use the symbol $|\mathcal{S}|$ for the cardinality of a set $\mathcal{S}$. To each $\widehat{\lambda}_{l,T}$ the procedure can return the frequency $\widehat{\omega}_l$ at which the break is found.

Assumption 5.2. $\delta_{l,T} = \delta_l \neq 0$ *is fixed or* $\delta_{l,T} \to 0$ *with* $\inf_{1\leq l\leq m_0} \delta_{l,T} \geq 2D^* M_{S,T}^{-1/2} (\log(T))^{2/3}$*. For* $\nu_T \to \infty$ *with* $\nu_T = o(T/v_T)$*, it holds that* $\inf_{1\leq l\leq m_0-1} |\lambda_{l+1}^0 - \lambda_l^0| \geq \nu_T^{-1}$.

Assumption 5.2 allows for shrinking shifts and a possibly growing number of change-points as long as $m_0/\nu_T \to 0$. The following proposition presents the consistency result for the number of change-points m_0 and for the change-point locations λ_l^0 $(l = 1, \ldots, m_0)$, and the rate of convergence of their estimates.

Proposition 5.2. *Let Assumption 3.2-3.3, 3.6 and 4.1 with* $m_0 = 1$*. Then, under* $\mathcal{H}_1^{\mathrm{B},m_0}$ *we have (i)* $\mathbb{P}(|\widehat{\mathcal{T}}| = m_0) \to 1$ *and* $\sup_{1\leq l\leq m_0} |\widehat{\lambda}_{l,T} - \lambda_l^0| = o_{\mathbb{P}}(1)$*, and (ii)* $\sup_{1\leq l\leq m_0} |\widehat{\lambda}_{l,T} - \lambda_l^0| = O_{\mathbb{P}}(m_{S,T}\sqrt{M_{S,T}\log(T)}/(T \inf_{1\leq l\leq m_0} \delta_{l,T}))$*. Furthermore, if* $K = O(a_T \dot{m}_T)$ *with* $a_T \in (0,\,1]$ *such that* $a_T \to 1$*, then the breaks are detected in decreasing order of magnitude.*

The number of draws K may be fixed or increase with the sample size. However, the algorithm can return the change-point dates in decreasing order of the break magnitudes only if K is sufficiently large. Note that at each loop of the algorithm it is not possible to know to which

λ_l^0 $(l = 1, \ldots, m_0)$ the estimate $\widehat{\lambda}_{l,T}$ 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. The same procedure can be applied for the case of multiple smooth local changes, though the notation becomes cumbersome and so we omit it.

6 Implementation

In this section we explain how to choose the tuning parameters. The choice of m_T, n_T and $b_{W,T}$ could be based on a mean-squared error (MSE) criterion or cross-validation exploiting results derived for locally stationary series [e.g., data-dependent methods for bandwidths in the context of locally stationary processes were investigated by, among others, Casini (2023), Dahlhaus (2012), Dahlhaus and Giraitis (1998) and Richter and Dahlhaus (2019)]. The optimal amount of smoothing depends on the regularity exponent θ, on the boundness of the moments and on the extent of the dependence in $\{X_{t,T}\}$. Here we choose the order of the bandwidths, neglecting the constants, by following the restrictions in Assumption 3.6. In particular, we choose the largest possible values allowed by Assumption 3.6 in order to ensure the highest possible power. We conduct a sensitivity analysis based on simulations in the supplement. We relegate to future work a more detailed analysis of data-dependent methods for this problem with multiple smoothing directions.

For spectral densities satisfying Lipschitz continuity, $\theta = 1$ so that $m_T \propto T^{2/3-\epsilon}$ while for $\theta = 1/2$ we have $m_T \propto T^{1/2-\epsilon}$ where in both cases $\epsilon > 0$. In applied work, it is common to work under stationarity $(\theta > 1)$ or local stationarity with Lipschitz smoothness $(\theta = 1)$. Hence, we use the bandwidths corresponding to $\theta = 1$ which works for both cases. Of course, if one has prior knowledge about the smoothness of the parameters of the data-generating process, one can choose a suitable θ. Assuming $q = 8$ and γ large enough we have $\tau_T \propto T^{1/4}$ and so values that satisfy Assumption 3.6 are $m_T = T^{0.66}$, $n_T = T^{0.62}$ and $b_{W,T} = n_T^{-1/6}$. The scaling is normalized to 1, as our simulations show this to provide good finite-sample properties, see Section 7.

As for the tapering function $h(\cdot)$ and weight function $W(\cdot)$ we use a rectangular taper (i.e., $h(u) = 1$ for all u) and a rectangular kernel. The rectangular kernel is known as the Daniell kernel with parameter $n_T b_{W,T}$ (it is a centered moving average which creates a smoothed value at time Tu by averaging all values between $Tu - n_T b_{W,T}$ and $Tu + n_T b_{W,T}$). These are the simplest choices for $h(\cdot)$ and $W(\cdot)$. As for the bandwidth and kernel of the estimator $\widehat{\sigma}_{L,r}(\omega)$, we follow the results in Casini (2022, 2023) that suggest $b_{1,T} = M_{S,T}^{-1/3}$. This corresponds to the MSE-optimal bandwidth when $K_1(\cdot)$ is the Bartlett kernel. For the choice of the number and values of the frequencies, the theory does not suggest particular values. Thus, we tried several values $n_\omega = 15, 11, 7, 5$. Our default choice is $n_\omega = 7$. A sensitivity analysis suggests that different choices for n_ω lead to negligible differences in the results. For the selection of the set of frequencies we use the function linspace that generates a linearly spaced sequence, e.g., in MATLAB we used the command $\mathsf{linspace}(-\pi+\mathsf{1e}-\mathsf{3},\,\mathsf{0},\,(n_\omega+\mathsf{1})/\mathsf{2})$.

The regularity exponent θ also affects the test $\psi(\{X_{t,T}\},\mathcal{I})$ in Algorithm 1. It is possible to get an estimate of θ under the null as follows. Compute $\mathrm{S}_{\mathrm{Dmax},T}$ where the maximum is taken among the indices of the blocks such that the null hypothesis is not violated and label it s^*_{Dmax}. Solve $s^*_{\mathrm{Dmax}}=2\sqrt{\log\left(M^*_T\right)/m^*_T}$ for θ, where recall that m^*_T and M^*_T depend on θ. This yields a preliminary estimate of θ which can then be used for the test $\psi(\{X_{t,T}\},\mathcal{I})$. Similarly, b^{opt}_T depends on θ and θ'. Using the same approach, for a given θ' one can solve $s^*_{\mathrm{Dmax}}=b^{\mathrm{opt}}_T$ for θ as function of θ'. If one is interested in the alternative $\mathcal{H}^{\mathrm{B}}_1$, we have $\theta'=0$ and so this immediately yields an estimate for θ. If one is interested in the alternative $\mathcal{H}^{\mathrm{S}}_1$, then one can try a few values of θ' in the range $(0,\,\theta)$. However, note that in order to use Algorithm 1 only θ is needed. The knowledge of θ' under $\mathcal{H}^{\mathrm{S}}_1$ is only needed to obtain b^{opt}_T.

We set $v_T=T^{0.666}$ which satisfies $m_T/v_T\to 0$. Our default recommendation is $K=10$. Our simulations with different data-generating processes and sample sizes show that this choice strikes a good balance between the precision of the change-point estimates and computing time. For $T>1000$, we recommend setting $K=\lfloor \dot{m}_T/3\rfloor$.

The test statistics $\mathrm{S}_{\max,T}(\omega)$ and $\mathrm{R}_{\max,T}(\omega)$ depend on ω. The choice of ω is, of course, important as it involves different frequency components and hence different periodicities. If the user does not have a priori knowledge about the frequency at which the spectrum has a change-point, our recommendation is to run the tests for multiple values of $\omega\in[0,\,\pi]$. Even if the change-point occurs at some ω_0 and one selects a value of ω close but not equal to ω_0 the tests are still able to reject the null hypothesis given the differentiability of $f(u,\,\omega)$. Thus, one can select a few values of ω evenly spread on $[0,\,\pi]$.

7 Small-Sample Evaluations

In this section, we conduct a Monte Carlo analysis to evaluate the properties of the proposed methods. We first discuss the detection of the change-points and then their localization. We investigate different types of changes and consider the test statistics $\mathrm{S}_{\max,T}(\omega)$, $\mathrm{S}_{\mathrm{Dmax},T}$, $\mathrm{R}_{\max,T}(\omega)$, $\mathrm{R}_{\mathrm{Dmax},T}$ proposed here and the test statistic $\widehat{D}$ proposed by Last and Shumway (2008). The latter is included for comparison since it applies to the same problems. 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. $\mathscr{N}(0,\,1)$. Models 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) $X_t=\rho\left(t/T\right)X_{t-1}+e_t$ where $\rho\left(t/T\right)=$

$0.4\cos(0.8-\cos(2t/T))$. Note that $\rho(t/T)$ varies smoothly from 0.1389 to 0.3920. 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 specification in model M1 and M2, respectively, for the first regime and consider two additional regimes with different specifications. Hence, two breaks are present. In model M3,

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

while, for model M4

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

where $\rho(t/T)$ is as in model M2. In model M3, the second regime involves higher serial dependence while in the third regime the variance doubles relative to the second regime. In model M4, the second regime involves a stationary autoregressive process with strong serial dependence while in the third regime X_t assumes the same dynamics as in the first regime. Models M3-M4 feature alternative hypotheses in the forms of breaks in the spectrum.

We consider the alternative hypothesis of more rough variation without signifying a break (i.e., $\mathcal{H}_1^{\mathrm{S}}$ defined in Section 4) in model M5 given by $X_t = \sigma(t/T)\,e_t$ where $\sigma^2(t/T) = \max\{1.5, \overline{\sigma}^2 + \cos(1+\cos(10t/T))\}$ with $\overline{\sigma}^2 = 1$. Note that even though $\sigma^2(\cdot)$ is locally stationary, the degree of smoothness alternates throughout the sample. It starts from $\sigma^2(\cdot) = 1.5$ and maintains this value for some time, then within a short period it increases slowly to $\sigma^2(\cdot) = 2$ and falls slowly back to $\sigma^2(\cdot) = 1.5$. It keeps this value until the final part of the sample where it increases slowly to $\sigma^2(\cdot) = 2$ for a short period. Thus, $\sigma^2(\cdot)$ alternates between periods where it is constant (i.e., $\theta > 1$) and periods where it becomes non-constant but less smooth (i.e., $\theta = 1$). Importantly, no break occurs; only a change in the smoothness as specified in $\mathcal{H}_1^{\mathrm{S}}$. In unreported simulations we also considered the case where θ changes from Lipschitz continuity (i.e., $\theta = 1$) to the continuity-path of Wiener processes (i.e., $\theta \approx 1/2$) with results that are similar to those reported here. For the test statistic $\widehat{D}$ of Last and Shumway (2008), we obtain the critical value by simulations. As suggested by the authors we compute the finite-sample distribution of $\widehat{D}$ by simulating a white noise under the null hypotheses with a sample size $T = 1000$ and then obtain the critical value. We consider the three sample sizes $T = 250$, 500 and 1000. The significance level is $\alpha = 0.05$. For the test statistics $\mathrm{S}_{\max,T}(\omega)$ and $\mathrm{R}_{\max,T}(\omega)$, we use as a default value $\omega = 0$ given that the interest is often in low frequency analysis. We set $\lambda_1^0 = 0.33$ and $\lambda_2^0 = 0.66$ throughout. The number of simulations is 5,000 for all cases.

The results are reported in Table 1-2. We first discuss the size of the tests. The tests proposed in this paper have good empirical size for both models and all sample sizes. The test statistics $\mathrm{S}_{\mathrm{Dmax},T}$ and $\mathrm{R}_{\mathrm{Dmax},T}$ are slightly undersized for $T = 250$ but their empirical size improves for $T = 500$ and 1000. The test statistics $\mathrm{S}_{\max,T}$ and $\mathrm{R}_{\max,T}$ share accurate empirical sizes in all cases. In contrast, the test statistic $\widehat{D}$ of Last and Shumway (2008) is largely oversized for $T = 250$ and 500. For $T = 1000$ it works better but it is still oversized. This means that the finite-sample distribution of $\widehat{D}$ has high variance and changes substantially across different sample sizes. Since the simulated critical value is obtained with a sample size $T = 1000$ it works better for this sample size than for the others for which the size control is poor.

Turning to the power of the tests, we note that it is not fair to compare the proposed tests with the test $\widehat{D}$ when $T = 250$ and 500 since the latter is largely oversized in those cases. In model M3, all the proposed tests have good power which increases with the sample size. The tests $\mathrm{S}_{\mathrm{Dmax},T}$ and $\mathrm{R}_{\max,T}$ have the highest power, followed by $\mathrm{S}_{\max,T}$ and lastly $\mathrm{R}_{\mathrm{Dmax},T}$. The power differences are not large except those involving $\mathrm{R}_{\mathrm{Dmax},T}$ for $T = 250$ which has substantially lower power. It is important to note that for $T = 1000$ the proposed tests have higher power than the test $\widehat{D}$ of Last and Shumway (2008) even though the latter is oversized. For $T = 250$ and 500, where the $\widehat{D}$ test is largely oversized, the proposed tests only have slightly lower power. This confirms that the proposed tests have very good power. Similar comments apply to model M4.

Model M5 involves changes in the smoothness without involving a break. This constitutes a more challenging alternative hypothesis, and as expected, the power for each test is lower than in models M3-M4. The test with the highest power is $\mathrm{S}_{\mathrm{Dmax},T}$. For $T = 1000$, the test with the lowest power is $\widehat{D}$ (for $T = 250$ the $\widehat{D}$ test has higher power again due to its oversize problem). Overall, the results show that the proposed tests have accurate empirical size even for small sample sizes and have good power against different forms of breaks or changes in "roughness".

Next, we consider the estimation of the number of change-points (m_0) and their locations. We consider the following two models, both with $m_0 = 2$. The 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 $\lambda_2^0 = 0.66$ and $T = 1000$

throughout. Table 3 reports summary statistics for $\widehat{m} - m_0$. It displays the percentage of times with $\widehat{m} = m_0$, the median, and the 25% and 75% quantile of the distribution of $\widehat{m}$. We only consider Algorithm 1. We do not report the results for the corresponding procedure of Last and Shumway (2008) because it is based on $\widehat{D}$ which is oversized and so it finds many more breaks than m_0. Table 3 shows that $\widehat{m} = m_0$ occurs for about 86% of the simulations with model M6 and about 81% with model M7. This suggests that Algorithm 1 is quite precise. As expected it performs better in model M6 since the specification of the alternative is farther from the null. The quantiles of the empirical distribution also suggest that the change-point estimates $\widehat{T}_1$ and $\widehat{T}_2$ are accurate. For example the median is very close to the their respective true value $T_1^0 = 333$ and $T_2^0 = 666$. Similar conclusions arise from different models and sample sizes, in unreported simulations.

8 Empirical Application

We demonstrate how to use our change-point methods for studying the causal effects of monetary policy. A fast growing literature in macroeconomics uses high-frequency data to identify the effects of monetary policy on the real economy. The identifying assumption used by Nakamura and Steinsson (2018) is that the volatility of the daily change in the nominal 2-year Treasury yields, say Δi_t, is higher during days when the Federal Open Market Committee (FOMC) meets to make monetary policy announcements relative to regular Tuesdays and Wednesdays with no announcement. Let η_t denote a pure monetary shock and suppose that the policy instrument Δi_t, which is observed in the data, is governed by both monetary and non-monetary shocks:

$$\Delta i_t = \mu_i + \eta_t + \varepsilon_t,$$

where ε_t is a function of all other shocks that affect Δi_t and μ_i is a constant. We normalize the impact of η_t and ε_t on Δi_t to one. Δi_t is a measure of the monetary policy news revealed in the FOMC announcement. The idea is that changes in the policy instrument during days when there is a FOMC announcement are dominated by the information about future monetary policy contained in the announcement. Let Δs_t denote the change in the outcome variable which is the yield on a five year zero-coupon Treasury bond. We wish to estimate the effects of the monetary shock η_t on Δs_t. The latter is also affected by both the monetary and non-monetary shocks:

$$\Delta s_t = \mu_s + \beta\eta_t + \alpha\varepsilon_t,$$

where μ_s is a constant and α and β are two parameters. The parameter of interest is β which represents the impact of the pure monetary shock η_t on Δs_t relative to its impact on Δi_t.

The identifying assumption is that the variance of monetary shocks increases during days of FOMC announcements, while the variances of the other shocks are unchanged. Let T_P denote the number of days containing a FOMC announcement, and let T_C the number of days with no such announcements. The subscript "P" in T_P refers to the "policy or treatment" sample while the subscript "C" in T_C refers to the "control" sample. The days in the control sample are comparable on other dimensions since they are all Tuesdays and Wednesdays with no FOMC meeting. The identifying restriction can then be written as

$$\sigma_{\eta,P} > \sigma_{\eta,C} \qquad \text{and} \qquad \sigma_{\varepsilon,P} = \sigma_{\varepsilon,C}, \tag{8.1}$$

where $\sigma_{a,i}$ is the volatility of variable $a = \eta,\, \varepsilon$ in the sample $i = P,\, C$. One can show that

$$\beta = \frac{\mathrm{Cov}_P\left(\Delta i_t,\, \Delta s_t\right) - \mathrm{Cov}_C\left(\Delta i_t,\, \Delta s_t\right)}{\mathrm{Var}_P\left(\Delta i_t\right) - \mathrm{Var}_C\left(\Delta i_t\right)}, \tag{8.2}$$

where $\mathrm{Cov}_i(\cdot,\, \cdot)$ (resp. $\mathrm{Var}_i(\cdot,\, \cdot)$) denotes the population covariance (resp. variance) in the $i = P,\, C$ sample. The parameter β can be identified only if $\mathrm{Var}_P\left(\Delta i_t\right) \neq \mathrm{Var}_C\left(\Delta i_t\right)$. If the volatility of the policy instrument Δi_t does not change across treatment and control samples, then β is not identified. If the change in volatility is small, then β is weakly identified.

Subsequent developments in the literature employed robust weak identification tests to argue that β is not strongly identified when using daily data. In contrast, β can be strongly identified if one uses ultra high-frequency data based on a 30 minute window around the announcement time. For the case of daily data, we show that the volatility of Δi_t in the control sample varies substantially over time and that several change-points can be detected. Thus, the regimes in the control sample where the volatility is high contribute to an average (over the control sample) volatility that approaches the average volatility in the treatment sample, thereby violating $\mathrm{Var}_P\left(\Delta i_t\right) \neq \mathrm{Var}_C\left(\Delta i_t\right)$. This implies that β may be weakly identified and its estimates may be imprecise which supports the recent evidence in the literature.

We obtain the data from Nakamura's webpage. The sample of "treatment" days is all scheduled FOMC meeting day from 1/1/2000 to 3/19/2014. The sample of "control" days is all Tuesdays and Wednesdays that are not FOMC meeting days from 1/1/2000 to 12/31/2012. In both the treatment and control samples, the second half of 2008, the first half of 2009 and a 10 day period after 9/11/2001 are dropped in Nakamura and Steinsson (2018). We follow the same practice.

The plot of $\{\Delta i_t\}$ for the control sample is reported in Figure 1. The series displays substantial changes in volatility and some changes in persistence. The tests $\mathrm{R}_{\max,T}$ and $\mathrm{R}_{\mathrm{D}\max,T}$ strongly reject the null hypothesis of no change-points in the spectrum. Algorithm 1 detects three change-points.

The first change-point date, denoted $\widehat{T}_1$, corresponds to April 24, 2007. Thus, the first regime $[1, \widehat{T}_1]$ refers to the period prior to the beginning of the 2007-09 financial crisis. It is evident that a large change in volatility and possibly a change in persistence occurred. We note that the change in volatility does not occur abruptly, it is rather gradual. This means that the volatility path changes gradually and possibly becomes more rough. This supports the usefulness of our hypothesis testing framework with local stationarity under the null hypothesis and with changes in the smoothness of the parameters that govern the data-generating process under the alternative hypothesis. The first change-point $\widehat{T}_1$ is associated to the volatility path becoming more rough with the level of the volatility increasing gradually.

The second change-point date, denoted $\widehat{T}_2$, corresponds to July 28, 2009. Thus, the second regime $[\widehat{T}_1 + 1, \widehat{T}_2]$ includes roughly the 2007-09 financial crisis. In this regime the volatility of the series is remarkably high. After $\widehat{T}_2$ the series shares a pattern similar to that in the first regime $[1, \widehat{T}_1]$, both in terms of persistence and volatility. The second change-point date $\widehat{T}_2$ is associated to an abrupt fall in volatility.

The third change-point date, denoted $\widehat{T}_3$, corresponds to February 2, 2011. The third regime $[\widehat{T}_2 + 1, \widehat{T}_3]$ corresponds to a zero lower bound (ZLB) period when the FOMC announced the conduct of unconventional monetary policies to stimulate the economy. The ZLB refers to a situation in which the short-term nominal interest rate is at or near zero, limiting the central bank's capacity. Unconventional monetary policies refer to large-scale asset purchases and active use of communication (i.e., "forward guidance") to shape expectations about future monetary policies, which can mitigate the limitations imposed by the ZLB [see, e.g., Swanson (2021)].

The last regime, $[\widehat{T}_3 + 1, T_C]$, corresponds to the period when the economy was witnessing the first effects of the expansive monetary policy and of the stability of the unconventional monetary policies that were introduced previously. This is a regime where initially the economy started the recovery and then reached stable economic growth. In this regime, the volatility level of the series is the lowest of the sample.

Overall, the first change-point date, $\widehat{T}_1$, corresponds to a change in the smoothness while the second change-point date, $\widehat{T}_2$, corresponds to an abrupt break. For the third change-point date, it is more difficult to tell from the plot whether this corresponds to an abrupt or smooth break.

We now discuss how the results about the change-points can be useful for the identification issue based on (8.1)-(8.2). One should think of $\text{Var}_C(\Delta i_t)$ as the average variance of Δi_t in the control sample. Our results show that there is significant time variation in $\text{Var}_C(\Delta i_t)$. In the second regime, $[\widehat{T}_1 + 1, \widehat{T}_2]$, the volatility is the highest of the sample. During this period, it is lower but very close to the average volatility of Δi_t in the treatment sample, $\text{Var}_P(\Delta i_t)$.[3] This

[3] We refer to Nakamura and Steinsson (2018) for details about the series in the treatment sample. We do not

contributes to make $\text{Var}_P(\Delta i_t) - \text{Var}_C(\Delta i_t)$ (over the full sample) closer to zero which then would lead to weak identification. In fact, Nakamura and Steinsson (2018) found the estimate of β to be imprecise and not meaningful from an economic standpoint. Furthermore, it was quite different from that obtained using 30 minute data instead of daily data. Our change-point analysis is useful because it suggests which periods or sub-samples contribute to this identification problem and which sub-samples can be used to obtain consistent estimates for β.

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. 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 to estimate 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 hypothesis beyond the fact that the spectrum is differentiable and bounded. Furthermore, the method allows for a broader range of alternative hypotheses compared to time-domain methods which usually have power against a limited set of alternatives. Overall, our simulations and empirical results show the usefulness of our method.

test for change-points in the treatment sample because the sample size in the treatment sample is relatively small ($T_P = 74$) and so we treat it as a single regime.

References

Adak, S., 1998. Time-dependent spectral analysis of nonstationary time series. Journal of the American Statistical Association 93, 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, 821–56.

Aue, A., Hórvath, L., 2013. Structural breaks in time series. Journal of Time Series Analysis 34, 1–16.

Aue, A., Nourinho, D.D., Hormann, S., 2015. On the prediction of stationary functional time series. Journal of the American Statistical Association 110, 378–392.

Aue, A., Rice, G., Sönmez, O., 2018. Detecting and dating structural breaks in functional data without dimension reduction. Journal of the Royal Statistical Society. Series B 80, 509–529.

Aue, A., van Delft, A., 2020. Testing for stationarity of functional time series in the frequency domain. Annals of Statistics Forthcoming.

Bai, J., 1994. Least squares estimation of a shift in linear processes. Journal of Time Series Analysis 15, 453–472.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47–78.

Bandyopadhyay, S., Carsten, J., Subba Rao, S., 2017. A spectral domain test for stationarity of spatio-temporal data. Journal of Time Series Analysis 38, 326–351.

Barigozzi, M., Cho, H., Fryzlewicz, P., 2018. Simultaneous multiple change-point and factor analysis for high-dimensional time series. Journal of Econometrics 206, 187–225.

Berkes, I., Philipp, W., 1979. Approximation theorems for independent and weakly dependent random vectors. Annals of Probability 7, 29–54.

Bibinger, M., Jirak, M., Vetter, M., 2017. Nonparametric change-point analysis of volatility. Annals of Statistics 45, 1542–1578.

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

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

Casini, A., 2018. Tests for forecast instability and forecast failure under a continuous record asymptotic framework. arXiv preprint arXiv:1803.10883.

Casini, A., 2022. Comment on Andrews (1991) "Heteroskedasticity and autocorrelation consistent covariance matrix estimation". Econometrica 90, 1–2.

Casini, A., 2023. Theory of evolutionary spectra for heteroskedasticity and autocorrelation robust inference in possibly misspecified and nonstationary models. Journal of Econometrics, 235, 372–392.

Casini, A., Deng, T., Perron, P., 2023. Theory of low frequency contamination from nonstationarity and misspecification: consequences for HAR inference. arXiv preprint arXiv:2103.01604.

Casini, A., Perron, P., 2019. Structural breaks in time series. Oxford Research Encyclopedia of Economics and Finance, Oxford University Press.

Casini, A., Perron, P., 2020. Generalized Laplace inference in multiple change-points models. Econometric Theory 38, 35–65.

Casini, A., Perron, P., 2021a. Continuous record asymptotics for change-point models. arXiv preprint arXiv:1803.10881.

Casini, A., Perron, P., 2021b. Continuous record Laplace-based inference about the break date in structural change models. Juornal of Econometrics 224, 3–21.

Casini, A., Perron, P., 2023a. Prewhitened long-run variance estimation robust to nonstattionarity. arXiv preprint arXiv:2103.02235.

Casini, A., Perron, P., 2023b. Supplement to “Change-point analysis of time series with evolutionary spectra”. Journal of Econometrics, forthcoming.

Chen, L., Wang, W., Wu, W.B., 2022. Inference of break-points in high-dimensional time series. Journal of the American Statistical Association 117, 1951–1963.

Cho, H., Fryzlewicz, P., 2012. Multiscale and multilevel technique for consistent segmentation of nonstationary time series. Statistica Sinica 22, 207–229.

Cho, H., Fryzlewicz, P., 2017. Multiple change-point detection for nonstationary time series using wild binary segmentation. Statistica Sinica 27, 287–311.

Coates, D.S., Diggle, P.J., 1986. Tests for comparing two estimated spectral densities. Journal of Time Series A 7, 7–20.

Dahlhaus, R., 1996. On the Kullback-Leibler information divergence of locally stationary processes. Stochastic Processes and their Applications 62, 139–168.

Dahlhaus, R., 1997. Fitting time series models to nonstationary processes. Annals of Statistics 25, 1–37.

Dahlhaus, R., 2012. Locally stationary processes, in: Subba Rao, T., Subba Rao, S., Rao, C. (Eds.), Handbook of Statistics. Elsevier. volume 30, pp. 351–413.

Dahlhaus, R., Giraitis, L., 1998. On the optimal segment length for parameter estimates for locally stationary time series. Journal of Time Series Analysis 19, 629–655.

van Delft, A., Characiejus, V., Dette, H., 2018. A nonparametric test for stationarity in functional time series. Statistica Sinica, forthcoming .

van Delft, A., Eichler, M., 2018. Locally stationary functional time series. Electronic Journal of Statistics 12, 107–170.

Dette, H., Preuß, P., Vetter, M., 2011. A measure of stationarity in locally stationary processes with applications to testing. Journal of the American Statistical Association 106, 1113–1124.

Dette, H., Wu, W., Zhou, Z., 2019. Change-point analysis of correlation in nonstationary time series. Statistica Sinica 29, 611–643.

Dwivedi, Y., Subba Rao, S., 2010. A test for second-order stationarity of a time series based on the discrete Fourier transform. Journal of Time Series Analysis 32, 68–91.

Eichinger, B., Kirch, C., 2018. A MOSUM procedure for the estimation of multiple random change points. Bernoulli 24, 526–564.

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

Grenander, U., Rosenblatt, M., 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, 509–523.

Hoffmann, M., Nickl, R., 2011. On adaptive inference and confidence bands. Annals of Statistics 39, 2383–2409.

Hörmann, S., Kokoszka, P., 2010. Weakly dependent functional data. Annals of Statistics 38, 1845–1884.

Horváth, L., 1993. The maximum likelihood method for testing changes in the parameters of normal observations. Annals of Statistics 21, 671–680.

Horváth, L., Kokoszka, P., Rice, G., 2014. Testing stationarity of functional time series. Journal of Econometrics 179, 66–82.

Huang, H.Y., Ombao, H., Stoffer, D.S., 2004. Discrimination and classification of nonstationary time series using the SLEX model. Journal of the American Statistical Association 99, 763–774.

Ingster, Y.I., 1993. Asymptotically minimax hypothesis testing for nonparametric alternatives I, II, III. Mathematical Methods of Statistics 2, 85–114.

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

Jentsch, C., Subba Rao, S., 2015. A test for second order stationarity of a multivariate time series. Journal of Econometrics 185, 124–161.

Kakizawa, Y., Shumway, R.H., Taniguchi, M., 1998. Discrimination and clustering for multivariate time series. Journal of the American Statistical Association 93, 328–340.

Kirch, C., Muhsal, B., Ombao, H., 2015. Detection of changes in multivariate time series with application to EEG data. Journal of the American Statistical Association 110, 1197–1216.

Last, M., Shumway, R., 2008. Detecting abrupt changes in a piecewise locally stationary time series. Journal of Multivariate Analysis 99, 191–214.

Liu, H., Gao, C., Samworth, R.J., 2021. Minimax rates in sparse, high-dimensional change-point detection. Annals of Statistics 49, 1081–1112.

Loader, C.R., 1996. Change point estimation using nonparameteric regression. Annals of Statistics 24, 1667–1678.

Müller, H.G., 1992. Change-points in nonparametric regression analysis. Annals of Statistics 20, 737–761.

Müller, H.G., Stadtmuller, U., 1999. Discontinuous versus smooth regression. Annals of Statistics 27, 299–337.

Nakamura, E., Steinsson, J., 2018. High frequency identification of monetary non-neutrality: The information effect. Quarterly Journal of Economics 133, 1283–1330.

Neumann, M.H., von Sachs, R., 1997. Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. Annals of Statistics 25, 38–76.

Panaretos, V.M., Tavakoli, S., 2013. Fourier analysis of stationary time series in function space. Annals of Statistics 41, 568–603.

Paparoditis, E., 2009. Testing temporal constancy of the spectral structure of a time series. Bernoulli 15, 1190–1221.

Parzen, E., 1957. On consistent estimates of the spectrum of a stationary time series. Annals of Mathematical Statistics 28, 329–348.

Preuß, P., Puchstein, R., Dette, H., 2015. Detection of multiple structural breaks in multivariate time series. Journal of the American Statistical Association 110, 654–668.

Preuß, P., Vetter, M., Dette, H., 2013. A test for stationarity based on empirical processes. Bernoulli 19, 2715–2749.

Priestley, M.B., Subba Rao, T., 1969. A test for nonstationarity of time series. Journal of the Royal Statistical Society: Series B 31, 140–149.

Priestley, M.B., 1965. Evolutionary spectra and non-stationary processes. Journal of the Royal Statistical Society: Series B 27, 204–237.

Richter, S., Dahlhaus, R., 2019. Cross validation for locally stationary processes. Annals of Statistics 47, 2145–2173.

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

Schröder, A.L., Ombao, H., 2019. FreSpeD: frequency-specific change-point detection in epileptic seizure multi-channel EEG data. Journal of the American Statistical Association 114, 115–128.

Shao, X., Wu, W.B., 2007. Local Whittle estimation of fractional integration for nonlinear processes. Econometric Theory 23, 899–929.

Spokoiny, V.G., 1998. Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice. Annals of Statistics 26, 1356–1378.

Swanson, E.T., 2021. Measuring the effects of federal reserve forward guidance and asset purchases on financial markets. Journal of Monetary Economics 118, 32–53.

Verzelen, N., Fromont, M., Lerasle, M., Reynaud-Bouret, P., 2020. Optimal change-point detection and localization. arXiv preprint arXiv:2010.11470 .

Vogt, M., Dette, H., 2015. Detecting gradual changes in locally stationary processes. Annals of Statistics 43, 713–740.

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, 14150–14154.

Wu, W.B., 2007. Strong invariance principles for dependent random variables. Annals of Probability 35, 2294–2320.

Wu, W.B., Zhao, Z., 2007. Inference of trends in time series. Journal of the Royal Statistical Society. Series B 69, 391–410.

Wu, W.B., Zhou, Z., 2011. Gaussian approximation for non-stationary multiple time series. Statistica Sinica 21, 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.

Yu, Y., Chatterjee, S., Xu, H., 2022. Localising change points in piecewise polynomials of general degrees. Electronic Journal of Statistics 16, 1855–1890.

Zhang, T., 2016. Testing for Jumps in the Presence of Smooth Changes in Trends of Nonstationary Time Series. Electronic Journal of Statistics 10, 706–735.

Zhang, T., Wu, W.B., 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, 726–740.

10 Appendix

10.1 Tables

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

	Model M1		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}(0)$	0.039	0.043	0.053
$\mathrm{S}_{\mathrm{Dmax},T}$	0.029	0.049	0.047
$\mathrm{R}_{\max,T}(0)$	0.040	0.054	0.042
$\mathrm{R}_{\mathrm{Dmax},T}$	0.025	0.032	0.038
$\widehat{D}$ statistic	0.581	0.471	0.068
	Model M2		
	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}(0)$	0.061	0.059	0.057
$\mathrm{S}_{\mathrm{Dmax},T}$	0.035	0.055	0.058
$\mathrm{R}_{\max,T}(0)$	0.036	0.035	0.039
$\mathrm{R}_{\mathrm{Dmax},T}$	0.025	0.032	0.035
$\widehat{D}$ statistic	0.731	0.583	0.102

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

	Model M3		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}(0)$	0.694	0.850	0.889
$\mathrm{S}_{\mathrm{Dmax},T}$	0.734	0.890	0.921
$\mathrm{R}_{\max,T}(0)$	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		
	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}(0)$	0.868	0.964	0.973
$\mathrm{S}_{\mathrm{Dmax},T}$	0.938	0.988	0.996
$\mathrm{R}_{\max,T}(0)$	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
$\mathrm{R}_{\mathrm{Dmax},T}$	0.025	0.189	0.304
$\widehat{D}$ statistic	0.834	0.695	0.172

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

Percent time $\widehat{m} = m_0$		$Q_{0.25}$	Median	$Q_{0.75}$
Model M6				
85.75	$\widehat{T}_1$	299	333	352
	$\widehat{T}_2$	632	663	688
Model M7				
80.88	$\widehat{T}_1$	317	336	359
	$\widehat{T}_2$	623	655	685

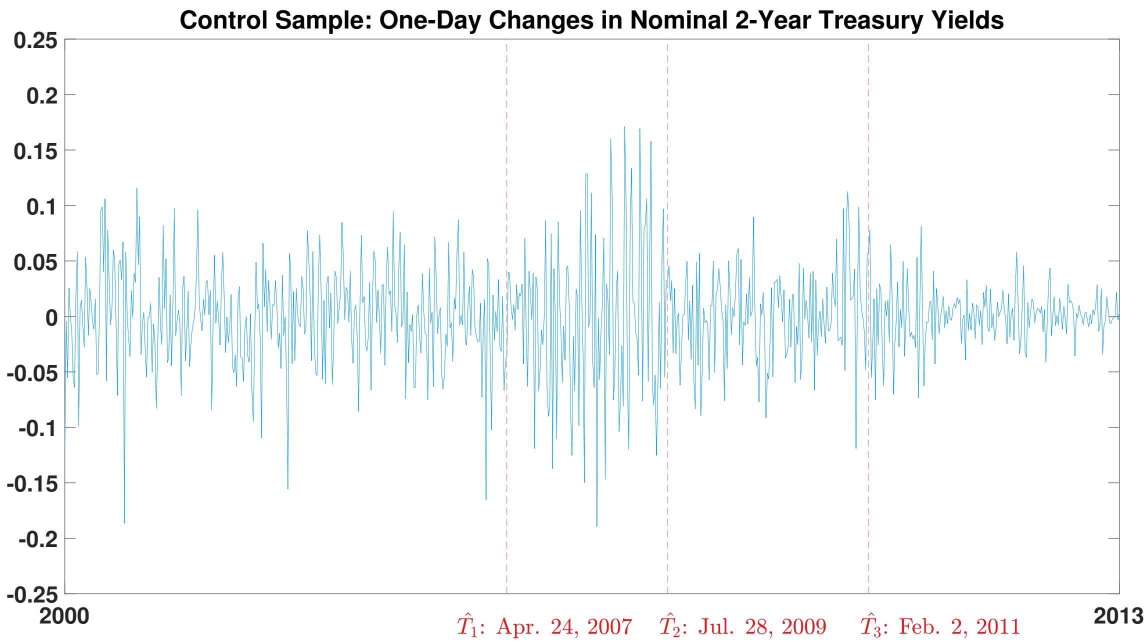

Figure 1: Plot of one-day changes in the nominal Treasury yields (Δi_t) in the control sample. The sample size is $T_C = 762$ which corresponds to all Tuesdays and Wednesdays that are not FOMC meeting days from 1/1/2000 to 12/31/2012. Following Nakamura and Steinsson (2018) we drop the second half of 2008, the first half of 2009 and a 10 day period after 9/11/2001. The red dashed lines are change-point dates estimated using Algorithm 1.

Supplemental Material to

Change-Point Analysis of Time Series with Evolutionary Spectra

Alessandro Casini — University of Rome Tor Vergata
Pierre Perron — Boston University

7th October 2023

Abstract

This supplemental material is structured as follows. Section S.A develops asymptotic results about high-order cumulants and spectra for locally stationary series that are needed in the proofs of the main results and are also of independent interest. Section S.B presents the Mathematical Appendix, which includes the proofs of the results of the paper and of Section S.A. Section S.C includes a sensitivity analysis for the tuning parameter choices. In Section S.D we present additional simulation results.

S.A Results About High-Order Cumulants and Spectra of Locally Stationary Series

This section establishes asymptotic results about high-order cumulants and spectra for locally stationary series. These are used to derive the limiting distributions of the test statistics introduced in Section 3. They are also of independent interest in the literature related to locally stationary and nonstationary processes more generally. We consider the tapered finite Fourier transform, the local and the smoothed local periodogram. Let

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

where $I_{h,T}(u,\omega)$ is the periodogram over a segment of length n_T with midpoint $\lfloor Tu\rfloor$. The smoothed local periodogram is defined as

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

where $W_T(\omega)$ and $b_{W,T}$ are defined in Section 3. Note that $d_{L,h,T}(u,\omega)$, $I_{L,h,T}(u,\omega)$ and $f_{L,h,T}(u,\omega)$ considered in Section 3 are asymptotically equivalent to $d_{h,T}(u,\omega)$, $I_{h,T}(u,\omega)$ and $f_{h,T}(u,\omega)$, respectively. Let $\mathbf{X}_{t,T}=(X_{t,T}^{(a_1)},\ldots,X_{t,T}^{(a_p)})$ with finite $p\geq 1$. Denote by $\kappa_{\mathbf{X},t}^{(a_1,\ldots,a_r)}(k_1,\ldots,k_{r-1})$ the time-t cumulant of order r of $(X_{t+k_1,T}^{(a_1)},\ldots,X_{t+k_{r-1},T}^{(a_{r-1})},X_{t,T}^{(a_r)})$ with $r\leq p$. If (3.6) holds for $l=0$, then we can define the rth order cumulant spectrum at the rescale time $u\in(0,1)$,

$$f_{\mathbf{X}}^{(a_1,\ldots,a_r)}(u,\omega_1,\ldots,\omega_{r-1})=(2\pi)^{r-1}\sum_{k_1,\ldots,k_{r-1}=-\infty}^{\infty}\kappa_{\mathbf{X},Tu}^{(a_1,\ldots,a_r)}(k_1,\ldots,k_{r-1})\exp\left(-i\sum_{j=1}^{k-1}\omega_jk_j\right),\qquad\text{(S.1)}$$

for any r tuple $a_1,\ldots,a_r$ with $r=2,3,\ldots$ When we assume that Assumption 3.1 holds, we mean that Assumption 3.1 holds for each element of $\mathbf{X}_{t,T}$.

S.A.1 Local Finite Fourier Transform

We first present the asymptotic expression for the joint cumulants of the finite Fourier transform. Next, we use this result to obtain the limit distribution of the transform. This result is subsequently used to derive the second-order properties of the local periodogram and smoothed local periodogram in the next subsections. Corresponding results for a stationary series can be found in Brillinger (1975) and references therein. Let $\mathbf{d}_{h,T}(u,\omega)=[d_{h,T}^{(a_j)}(u,\omega)]$ $(j=1,\ldots,r)$,

$$H_{n_T}^{(a_1,\ldots,a_r)}(\omega)=\sum_{s=0}^{n_T-1}\left(\prod_{j=1}^{r}h_{a_j}(s/n_T)\right)\exp(-i\omega s),\quad\text{and}$$
$$H^{(a_1,\ldots,a_r)}(\omega)=\int\left(\prod_{j=1}^{r}h_{a_j}(t)\right)\exp(-i\omega t)\,dt.$$

Let $\mathscr{N}_p^{\mathrm{C}}(\mathbf{c},\,\Sigma)$ denote the complex normal distribution for some p-dimensional vector $\mathbf{c}$ and $p\times p$ Hermitian positive semidefinite matrix Σ.

Theorem S.A.1. *Let Assumption 3.1, 3.2 with $l=0$ and Assumption 3.3-(ii) hold. Let $h_{a_j}(x)$ satisfy Assumption 3.3-(i) for all $j=1,\ldots,p$. We have*

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

where $\varepsilon_T=o(n_T)$ uniformly 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 $(j=1,\ldots,\,r)$. Furthermore,

$$f_{\mathbf{X}}^{(a_1,\ldots,a_r)}(u,\,\omega_1,\ldots,\,\omega_{r-1})=A^{(a_1)}(\lfloor Tu\rfloor,\,\omega_1)\cdots A^{(a_p)}(\lfloor Tu\rfloor,\,\omega_p)\,g_p(\omega_1,\ldots,\,\omega_{p-1}),$$

i.e., the spectrum that corresponds to the spectral representation (2.1) with $m_0=0$.

Theorem S.A.2. *Let Assumption 3.1, 3.2 with $l=0$ and Assumption 3.3-(ii) hold. Let $h_{a_j}(x)$ satisfy Assumption 3.3-(i) for all $j=1,\ldots,p$. We have: (i) If $2\omega_j,\,\omega_j\pm\omega_k\not\equiv 0\,(\mathrm{mod}\,2\pi)$ for $1\le j<k\le J_\omega$ with $1\le J_\omega<\infty$, $\mathbf{d}_{h,T}(u,\,\omega_j)$ $(j=1,\ldots,\,J_\omega)$ are asymptotically independent $\mathscr{N}_p^{\mathrm{C}}(0,2\pi n_T[H^{(a_l,a_r)}(0)\,f_{\mathbf{X}}^{(a_l,a_r)}(u,\,\omega_j)])$ $(l,\,r=1,\ldots,\,p)$ variables; (ii) If $\omega=0,\,\pm\pi,\,\pm2\pi,\,\pm3\pi,\,\ldots,\,\mathbf{d}_{h,T}(u,\,\omega)$ is asymptotically $\mathscr{N}_p(0,2\pi n_T[H^{(a_l,a_r)}(0)\,f^{(a_l,a_r)}(u,\,\omega)])$ $(l,\,r=1,\ldots,\,p)$ independently from the previous variates.*

When the series is stationary, (i.e., $f_{\mathbf{X}}^{(a_1,\ldots,a_r)}(u,\,\omega_1,\ldots,\,\omega_{r-1})$ does not depend on u), the results in Theorem S.A.1-S.A.2 reduce to the well-known results on the cumulants and asymptotic distribution of the Fourier transform when the latter is constructed using a segment of length n_T.

S.A.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 about its 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. This result is exploited when deriving the limit distribution of the test statistics that do not require knowledge of the frequency at which the change-point occurs.

Theorem S.A.3. *Let Assumption 3.1, 3.2 with $l=0$ and Assumption 3.3-(i,ii) hold. We have for $-\infty<\omega<\infty$,*

$$\begin{aligned}\mathbb{E}(I_{h,T}(u,\,\omega))&=\left(\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2\,d\alpha\right)^{-1}\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2\,f_{\mathbf{X}}(u,\,\omega-\alpha)\,d\alpha+O\left(\frac{\log(n_T)}{n_T}\right)\\ &=f_{\mathbf{X}}(u,\,\omega)+\frac{1}{2}\left(\frac{n_T}{T}\right)^2\left(\int_0^1 h^2(x)\,dx\right)^{-1}\int_0^1 x^2h^2(x)\,dx\frac{\partial^2}{\partial u}f_{\mathbf{X}}(u,\,\omega)\\ &\quad+o\left(\left(\frac{n_T}{T}\right)^2\right)+O\left(\frac{\log(n_T)}{n_T}\right).\end{aligned}\tag{S.2}$$

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 a neighborhood of ω and relative weights determined by the taper. The second equality shows that $I_{h,T}(u,\omega)$ is asymptotically unbiased for $f_{\mathbf{X}}(u,\omega)$ and provides a bound on the asymptotic bias.

Theorem S.A.4. *Let Assumption 3.1, 3.2 with $l=0$ and Assumption 3.3-(i,ii) hold. We have: (i) For $-\infty<\omega_j,\,\omega_k<\infty$,*

$$\begin{aligned}&\mathrm{Cov}\left\{I_{h,T}(u,\omega_j),\,I_{h,T}(u,\omega_k)\right\}\\&\quad=\left|H_{2,n_T}(0)\right|^{-2}\left(\left|H_{2,n_T}(\omega_j-\omega_k)\right|^2+\left|H_{2,n_T}(\omega_j+\omega_k)\right|^2\right)f_{\mathbf{X}}(u,\omega_j)^2+O\left(n_T^{-1}\right),\end{aligned}\tag{S.3}$$

where $O(n_T^{-1})$ is uniform in ω_j and ω_k; (ii) If $2\omega_j,\,\omega_j\pm\omega_k\not\equiv 0\,(\mathrm{mod}\,2\pi)$ with $1\le j<k\le J_\omega$, the variables $I_{h,T}(u,\omega_j)$ $(j=1,\ldots,J_\omega)$ are asymptotically independent $f_{\mathbf{X}}(u,\omega_j)\,\chi_2^2/2$ variates. Also, if $\omega=\pm\pi,\,\pm3\pi,\ldots,$ $I_{h,T}(u,\omega)$ is asymptotically $f_{\mathbf{X}}(u,\omega)\,\chi_1^2$, independent of the previous variates.

The variance expression for the tapered local periodogram follows as a special case of (S.3).

S.A.3 Smoothed Local Periodogram

We now extend Theorem S.A.3-S.A.4 to the smoothed local periodogram. Since our test statistics are based on it, these results are directly employed to derive their limiting null distributions.

Theorem S.A.5. *Let Assumption 3.1, 3.2 with $l=0$ and Assumption 3.3 hold. Let $b_{W,T}\to0$ as $T\to\infty$ with $b_{W,T}n_T\to\infty$. Then,*

$$\begin{aligned}\mathbb{E}\left(f_{h,T}(u,\omega)\right)&=\int_{-\infty}^{\infty}W(\beta)\,f_{\mathbf{X}}(u,\omega-b_{W,T}\beta)\,d\beta+O\left((n_Tb_{W,T})^{-1}\right)+O\left(\log(n_T)\,n_T^{-1}\right)\\&=f_{\mathbf{X}}(u,\omega)+\frac{1}{2}\left(\frac{n_T}{T}\right)^2\left(\int_0^1h^2(x)\,dx\right)^{-1}\int_0^1x^2h^2(x)\,dx\frac{\partial^2}{\partial u}f_{\mathbf{X}}(u,\omega)\\&\quad+\frac{1}{2}b_{W,T}^2\int_0^1x^2W(x)\,dx\frac{\partial^2}{\partial\omega}f_{\mathbf{X}}(u,\omega)+O\left((n_T/T)^{-2}\right)+O\left(\log(n_T)\,n_T^{-1}\right)+o\left(b_{W,T}^2\right).\end{aligned}\tag{S.4}$$

The error terms are uniform in ω.

Theorem S.A.6. *Let Assumption 3.1, 3.2 with $l=0$ and Assumption 3.3-(i,ii) hold. Let $b_{W,T}\to0$ as $T\to\infty$ with $b_{W,T}n_T\to\infty$. Then, $f_{h,T}(u,\omega_1),\ldots,f_{h,T}(u,\omega_{J_\omega})$ are asymptotically jointly normal satisfying*

$$\begin{aligned}&\lim_{T\to\infty}n_Tb_{W,T}\mathrm{Cov}\left(f_{h,T}(u,\omega_j),\,f_{h,T}(u,\omega_k)\right)\\&\quad=2\pi\left[\eta\{\omega_j-\omega_k\}+\eta\{\omega_j+\omega_k\}\right]\int h(t)^4\,dt\left[\int h(t)^2\,dt\right]^{-2}\int W(\alpha)^2\,d\alpha\,f_{\mathbf{X}}(u,\omega_j)^2.\end{aligned}\tag{S.5}$$

The variance expression for $f_{h,T}(u,\omega_j)$ follows immediately as a special case of (S.5). Consistency of the spectral density estimates of a stationary time series was 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 S.A.6 presents corresponding results for the locally stationary case which highlight the effect of the time-smoothing in addition to the smoothing 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. Paparoditis (2009) established similar results for linear locally stationary processes.

S.B Mathematical Appendix

S.B.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 w define $H_{n_T}(w(\cdot),\omega)=\sum_{s=0}^{n_T-1}w(s)\exp(-i\omega s)$, 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.B.1. *Let $\varPi\triangleq(-\pi,\pi]$. With a constant K independent of T the following properties hold: (i) $L_T(\omega)$ is monotone increasing in T and decreasing in $\omega\in[0,\pi]$; (ii) $\int_\varPi L_T(\alpha)d\alpha\leq K\ln T$ for $T>1$; (iii) $\int_\varPi L_T(\alpha)^k d\alpha\leq KT^{k-1}$ for $k\geq 2$.*

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

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

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

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

Proof of Lemma S.B.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 applications of Abel's transformation,

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

$$= \sum_{t=0}^{s-1} \left(h\left(\frac{t}{n_T}\right) - h\left(\frac{t-1}{n_T}\right)\right) H_t\left(1,\, \omega\right) + 0,$$

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

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

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

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

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

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

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

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

Proof of Lemma S.B.3. See Lemma P4.1 in Brillinger (1975). □

Lemma S.B.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 $(Y_T^{(a_1)},\, \overline{Y}_T^{(a_1)}, \ldots,\, Y_T^{(a_p)},\, \overline{Y}_T^{(a_p)})$ exist and tend to the corresponding cumulants of a variate $(Y^{(a_1)},\, \overline{Y}^{(a_1)}, \ldots,\, Y^{(a_p)},\, \overline{Y}^{(a_p)})$ 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.B.4. It follows from Lemma P4.5 in Brillinger (1975). □

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

S.B.2.1 Proof of Theorem S.A.1

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

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

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

$$\begin{aligned}&\text{cum}\left(X_{t_1,T}, \ldots, X_{t_p,T}\right)\\&\quad= \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(i\left(\left(\omega_1 + \cdots + \omega_{p-1}\right)t_p + \omega_1 k_1 + \cdots + \omega_{p-1}k_{p-1} + t_p\omega_p\right)\right)\\&\quad\quad \times A\left(\lfloor Tu \rfloor, \omega_1\right) \cdots A\left(\lfloor Tu \rfloor, \omega_p\right) \eta\left(\sum_{j=1}^{p} \omega_j\right) g_p\left(\omega_1, \ldots, \omega_{p-1}\right) d\omega_1 \cdots d\omega_p + O\left(n_T/T\right)\\&\quad= \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(i\left(\left(\omega_1 + \cdots + \omega_{p-1} + \omega_p\right)t_p + \omega_1 k_1 + \cdots + \omega_{p-1}k_{p-1}\right)\right)\\&\quad\quad \times A\left(\lfloor Tu \rfloor, \omega_1\right) \cdots A\left(\lfloor Tu \rfloor, \omega_p\right) \eta\left(\sum_{j=1}^{p} \omega_j\right) g_p\left(\omega_1, \ldots, \omega_{p-1}\right) d\omega_1 \cdots d\omega_p + O\left(n_T/T\right)\\&\quad\triangleq \kappa_{Tu,t_p}\left(k_1 \ldots, k_{p-1}\right) + O\left(n_T/T\right).\end{aligned} \tag{S.2}$$

This shows that $\text{cum}\left(X_{t_1,T}, \ldots, X_{t_p,T}\right)$ depends on t_p only through $\exp\left(i\left(\omega_1 + \cdots + \omega_{p-1} + \omega_p\right)t_p\right)$. The cumulant of interest in Theorem S.A.1 has the following form,

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

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

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

plus a remainder term R_u with

$$\begin{aligned}|R_u|&\leq C\frac{n_T}{T}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}L_{n_T}(\omega_1-\gamma_1)\cdots L_{n_T}(\omega_p-\gamma_p)\exp\left(i\left((\gamma_1+\cdots+\gamma_p)\lfloor Tu\rfloor\right)\right)\\&\quad\times\eta\left(\sum_{j=1}^{p}\gamma_j\right)g_p(\gamma_1,\ldots,\,\gamma_{p-1})\,d\gamma_1\cdots d\gamma_p\\&\leq C\frac{n_T}{T}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}L_{n_T}(\omega_1-\gamma_1)\cdots L_{n_T}(\omega_p-\gamma_p)\,d\gamma_1\cdots d\gamma_p\\&\leq C\frac{n_T}{T}(\ln n_T)^p,\end{aligned}\tag{S.4}$$

where we have used $g_p(\gamma_1,\ldots,\,\gamma_{p-1})\leq\text{const}_p$, the fact that $\int_{-\pi}^{\pi}\exp\{i(\gamma\lfloor Tu\rfloor)\}\,d\gamma=2\sin(\pi\lfloor Tu\rfloor)/\lfloor Tu\rfloor$, and the third inequality follows from Lemma S.B.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 have that (S.3) is equal to

$$\begin{aligned}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}&A^{(a_1)}(u,\,\omega_1)\cdots A^{(a_p)}(u,\,\omega_p)\\&\times H_{n_T}^{(a_1)}(\omega_1-\gamma_1)\cdots H_{n_T}^{(a_p)}(\omega_p-\gamma_p)\\&\times\exp\left(i\left((\omega_1+\cdots+\omega_p)\lfloor Tu\rfloor\right)\right)\eta\left(\sum_{j=1}^{p}\omega_j\right)g_p(\omega_1,\ldots,\,\omega_{p-1})\,d\gamma_1\cdots d\gamma_p.\end{aligned}\tag{S.5}$$

By Lemma S.B.3,

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

Thus, (S.5) is equal to

$$\begin{aligned}\sum_{k_1=-n_T}^{n_T}\cdots\sum_{k_{p-1}=-n_T}^{n_T}&\exp\left(-i\sum_{j=1}^{p-1}\omega_j k_j\right)\\&\times\left(\kappa_{Tu,t_p}^{(a_1,\ldots,\,a_p)}(k_1\ldots,\,k_{p-1})\,H_T^{(a_1,\ldots,\,a_p)}\left(\sum_{j=1}^{p}\omega_j\right)+O(n_T/T)\right)+\varepsilon_T,\end{aligned}\tag{S.6}$$

where $\kappa_{Tu,t_p}^{(a_1,\ldots,a_p)}(k_1\ldots,k_{p-1})=\mathrm{cum}(X_{t_1,T}^{(a_1)},\ldots,X_{t_p,T}^{(a_p)})+(n_T/T)$ and

$$|\varepsilon_T|\leq C\sum_{k_1=-n_T}^{n_T}\cdots\sum_{k_{p-1}=-n_T}^{n_T}\kappa_{Tu,t_p}^{(a_1,\ldots,a_p)}(k_1\ldots,k_{p-1})(|k_1|+\cdots+|k_p|)<\infty.$$

Note that $|\varepsilon_T|/n_T\to0$ since $(|k_1|+\cdots+|k_p|)/n_T\to0$. Thus, $\varepsilon_T=o(n_T)$ uniformly in ω_j $(j=1,\ldots,p)$. Altogether we have

$$\begin{aligned}&\mathrm{cum}\left(d_{h,T}^{(a_1)}(u,\omega_1),\ldots,d_{h,T}^{(a_p)}(u,\omega_p)\right)\\&\qquad=(2\pi)^{r-1}H_{n_T}^{(a_1,\ldots,a_p)}\left(\sum_{j=1}^{p}\omega_j\right)f_{\mathbf{X}}^{(a_1,\ldots,a_p)}(u,\omega_1,\ldots,\omega_{p-1})+\varepsilon_T,\end{aligned}$$

where $f_{\mathbf{X}}^{(a_1,\ldots,a_p)}(u,\omega_1,\ldots,\omega_{p-1})$ is given in (2.1). The proof for the rth cumulant of $d_{h,T}^{(a_j)}(u,\omega_1)$ $(j=1,\ldots,r)$ with $r<p$ is the same as for the pth cumulant.

Note that from (S.6) we have

$$\begin{aligned}&\sum_{k_1=-n_T}^{n_T}\cdots\sum_{k_{p-1}=-n_T}^{n_T}\exp\left(-i\sum_{j=1}^{p-1}\omega_jk_j\right)\kappa_{Tu,t_p}(k_1\ldots,k_{p-1})\\&\quad=\sum_{k_1=-n_T}^{n_T}\cdots\sum_{k_{p-1}=-n_T}^{n_T}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}\exp(i((\gamma_1+\cdots+\gamma_{p-1}+\gamma_p)t_p\\&\quad+(\omega_1-\gamma_1)k_1+\cdots+(\omega_{p-1}-\gamma_{p-1})k_{p-1}))\\&\quad\times A^{(a_1)}(\lfloor Tu\rfloor,\gamma_1)\cdots A^{(a_p)}(\lfloor Tu\rfloor,\gamma_p)\,\eta\left(\sum_{j=1}^{p}\gamma_j\right)g_p(\gamma_1,\ldots,\gamma_{p-1})\,d\gamma_1\cdots d\gamma_p.\end{aligned}$$

Since $\sum_{j=1}^{p}\gamma_j\equiv0\,(\mathrm{mod}\,2\pi)$, γ_p is normalized and so the latter is equivalent to

$$\begin{aligned}&\sum_{k_1=-n_T}^{n_T}\cdots\sum_{k_{p-1}=-n_T}^{n_T}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}\exp(i(\gamma_1k_1+\cdots+\gamma_{p-1}k_{p-1}))\\&\qquad\times A^{(a_1)}(\lfloor Tu\rfloor,\omega_1)\cdots A^{(a_p)}(\lfloor Tu\rfloor,\omega_p)\,g_p(\omega_1,\ldots,\omega_{p-1})\,d\gamma_1\cdots d\gamma_{p-1},\end{aligned}$$

where we have used the continuity of $A(\cdot,\omega)$ and g_p. Then,

$$A^{(a_1)}(\lfloor Tu\rfloor,\omega_1)\cdots A^{(a_p)}(\lfloor Tu\rfloor,\omega_p)\,g_p(\omega_1,\ldots,\omega_{p-1})=f_{\mathbf{X}}^{(a_1,\ldots,a_p)}(u,\omega_1,\ldots,\omega_{p-1})\qquad\text{(S.7)}$$

is the spectrum that corresponds to the spectral representation (2.1) with $m_0=0$. In view of the following identities [see e.g., Exercise 1.7.5-(c,d) in Brillinger (1975)],

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

we have

$$\mathrm{cum}\left(d_T^{(a_1)}(u,\omega_1),\ldots,d_T^{(a_p)}(u,\omega_p)\right)$$

$$= (2\pi)^{p-1} H_T^{(a_1,\ldots,a_p)} \left(\sum_{j=1}^{p} \omega_j \right) f^{(a_1,\ldots,a_p)} (u, \omega_1, \ldots, \omega_{p-1}) + \varepsilon_T,$$

which verifies (S.7). □

S.B.2.2 Proof of Theorem S.A.2

We have,

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

By Theorem S.A.1 we deduce

$$\begin{aligned} n_T^{-1} \text{Cov} \left(d_{h,T}^{(a_l)} (u, \pm\omega_j), d_{h,T}^{(a_r)} (u, \pm\omega_k) \right) \\ = n_T^{-1} 2\pi H_{n_T}^{(a_l, a_r)} (\pm\omega_j \mp \omega_k) f_{\mathbf{X}}^{(a_l, a_r)} (u, \pm\omega_j (n_T)) + o(1) + O\left(n_T^{-1} \right). \end{aligned} \tag{S.8}$$

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

$$\left| H_{n_T}^{(a_1,\ldots,a_p)} (\omega) \right| \leq \frac{C}{|\sin(\omega/2)|}, \tag{S.9}$$

where C is a constant with $0 < C < \infty$. If $\omega_j \pm \omega_k \not\equiv 0 \,(\text{mod}\, 2\pi)$ the first term on the right-hand side of (S.8) tends to zero using (S.9). If $\pm\omega_j \mp \omega_k \equiv 0 \,(\text{mod}\, 2\pi)$ the right-hand side of (S.8) tends to

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

This shows that the second-order cumulants behave as indicated by the theorem. By Theorem S.A.1 for $r > 2$,

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

The latter tends to 0 as $n_T \to \infty$ if $r > 2$ because $H_{n_T}^{(a_1,\ldots,a_r)} (\omega) = O(n_T)$. 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.B.4. The second part of the theorem follows from the fact that $\sin(\omega) = 0$ for $\omega = 0, \pm\pi, \pm 2\pi, \pm 2\pi, \ldots$ □

S.B.2.3 Proof of Theorem S.A.3

The proof of the second equality in (S.2) is similar to Dahlhaus (1997) who proved the result under stronger assumptions on the data taper. Using the spectral representation (2.1),

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

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

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

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

$$\left|H_t\left(h\left(\frac{\cdot}{n_T},\,\omega-\eta\right)\right)\right|\leq L_t\left(\omega-\eta\right)\leq L_{n_T}\left(\omega-\eta\right). \tag{S.10}$$

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

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

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

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

Using Lemma S.B.2 and (S.10), we have

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

Thus, after integration over η we obtain that the 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

$$A(u,\,\omega)\,A(u,\,-\omega)=|A(u,\,\omega)|^2=f_{\mathbf{X}}(u,\,\omega).$$

In view of

$$\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2\,d\alpha=2\pi\sum_{t=0}^{n_T-1}\left(\frac{t}{n_T}\right)^2, \tag{S.11}$$

we then have

$$\mathbb{E}(I_{h,T}(u,\,\omega))=\frac{1}{2\pi H_{2,n_T}(0)}\sum_{t=0}^{n_T-1}\sum_{s=0}^{n_T-1}h\left(\frac{t}{T}\right)\left(\frac{s}{T}\right)\int_{-\pi}^{\pi}\exp(-i(\omega-\alpha)(s-t))\,f_{\mathbf{X}}(u,\,\alpha)\,d\alpha+O\left(\frac{\log n_T}{n_T}\right) \tag{S.12}$$

$$\begin{aligned}&=\frac{1}{\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2\,d\alpha}\int_{-\pi}^{\pi}|H_{n_T}(\omega-\alpha)|^2\,f_{\mathbf{X}}(u,\,\alpha)\,d\alpha+O\left(\frac{\log n_T}{n_T}\right)\\&=\frac{1}{\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2\,d\alpha}\int_{-\pi}^{\pi}|H_{n_T}(\alpha)|^2\,f_{\mathbf{X}}(u,\,\omega-\alpha)\,d\alpha+O\left(\frac{\log n_T}{n_T}\right).\end{aligned}$$

This shows the first equality of (S.2). For the second equality replace $A^0_{\lfloor Tu\rfloor-n_T/2+t}(\eta)$ by $A(u+(t-n_T/2)/T,\,\omega)$ and $A^0_{\lfloor Tu\rfloor-n_T/2+t}(-\eta)$ by $A(u+(t-n_T/2)/T,\,-\omega)$ so that (S.12) holds with $f_{\mathbf{X}}(u+(t-n_T/2)/T,\,\omega)$ in place of $f_{\mathbf{X}}(u,\,\alpha)$. Then take a second-order Taylor expansion of $f_{\mathbf{X}}$ around around u to obtain

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

S.B.2.4 Proof of Theorem S.A.4

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

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

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 $|P_i|=\mathrm{card}(P_i)\geq 2$ of the two-way table,

$$\begin{vmatrix}a_{j,1} & a_{j,2}\\ a_{k,1} & a_{k,2}\end{vmatrix},$$

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

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

Then, by Theorem S.A.1,

$$\begin{aligned}
&\mathrm{cum}\left(d_{h,T}(u,\,\omega_j)\,d_{h,T}(u,\,-\omega_j)\,,\,d_{h,T}(u,\,\omega_k)\,d_{h,T}(u,\,-\omega_k)\right)\\
&\quad=(2\pi)^3\,H_{4,n_T}(0)\,f_{\mathbf{X}}(u,\,\omega_j,\,-\omega_j,\,\omega_k)+O(1)\\
&\qquad+\left[2\pi H_{2,n_T}(\omega_j+\omega_k)\,f_{\mathbf{X}}(u,\,\omega_j)+O(1)\right]\left[2\pi H_{2,n_T}(-\omega_j-\omega_k)\,f_{\mathbf{X}}(u,\,\omega_j)+O(1)\right]\\
&\qquad+\left[2\pi H_{2,n_T}(\omega_j-\omega_k)\,f_{\mathbf{X}}(u,\,\omega_j)+O(1)\right]\left[2\pi H_{2,n_T}(-\omega_j+\omega_k)\,f_{\mathbf{X}}(u,\,\omega_j)+O(1)\right].
\end{aligned}\tag{S.13}$$

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

$$H_{2,n_T}(\omega_j-\omega_k)\,H_{2,n_T}(-\omega_j+\omega_k)=\left|H_{2,n_T}(\omega_j-\omega_k)\right|,^2$$

the result of the theorem follows because

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

and because the $O(1)$ terms on the right-hand side of (S.13) become negligible when multiplied by $H_{2,n_T}^{-2}(0)$.

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

$$\begin{bmatrix}\mathrm{Re}\,\mathbf{z}\\ \mathrm{Im}\,\mathbf{z}\end{bmatrix}$$

is distributed as

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

where Σ_z is a $p\times p$ hermitian positive semidefinite matrix. By Theorem S.A.2 we know that $\mathrm{Re}\,\mathbf{d}_{h,T}(\omega_j)$ and $\mathrm{Im}\,\mathbf{d}_{h,T}(\omega_j)$ are asymptotically independent $\mathscr{N}(0,\,\pi n_T f_{\mathbf{X}}(u,\,\omega_j))$ variates. Hence, by the Mann-

Wald Theorem,

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

is asymptotically distributed as $f_{\mathbf{X}}(u,\,\omega_j)\,\chi_2^2/2$ if $2\omega_j \not\equiv 0\,(\mathrm{mod}\,2\pi)$. This proves part (i). For part (ii), if $\omega = \pm\pi,\,\pm 3\pi, \ldots$, then $I_{h,T}(u,\,\omega)$ is asymptotically distributed as $f_{\mathbf{X}}(u,\,\omega)\,\chi_1^2$, independently from the previous variates. □

S.B.2.5 Proof of Theorem S.A.5

Using Theorem S.A.3, we have

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

The first term on the right-hand side is

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

where the last equality follows from the change in variable $\beta = b_T^{-1}(\omega-\alpha)$. This yields the first equality of (S.4). The second equality follows from the first and Theorem S.A.3 along with a Taylor expansion. □

S.B.2.6 Proof of Theorem S.A.6

Let

$$c_T(u,\,k) = H_{2,T}(0)^{-1}\sum_{s=0}^{n_T-1} h\left(\frac{s+k}{T}\right) h\left(\frac{s}{T}\right) X_{\lfloor Tu\rfloor - n_T/2+s+k+1,T}\, X_{\lfloor Tu\rfloor - n_T/2+s+1,T}.$$

We can rewrite $I_{h,T}(u,\,\omega)$ using $c_T(u,\,k)$ as follows,

$$I_{h,T}(u,\,\omega) = (2\pi)^{-1}\sum_{k=-\infty}^{\infty}\exp(-i\omega k)\, c_T(u,\,k).$$

Note that

$$f_{h,T}(u,\,\omega) = \int_0^{2\pi} W_{2,T}(\omega-\alpha)\, I_{h,T}(u,\,\alpha)\, d\alpha + O\left((n_T b_{W,T})^{-1}\right),$$

where $W_{2,T}(\omega) = \sum_{k=-\infty}^{\infty} w(b_{W,T}k)\exp(-i\omega k)$ and $w(k) = \int_{-\infty}^{\infty} W_{2,T}(\alpha)\exp(i\alpha k)\,d\alpha$ for $k \in \mathbb{R}$. From Theorem S.A.4,

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

We now show that

$$
\begin{aligned}
&\int_0^{2\pi} W_{2,T}(\omega_k-\beta)\,|H_{2,n_T}(\alpha-\beta)|^2\,d\beta \qquad \text{(S.14)}\\
&\qquad\qquad = 2\pi W_{2,T}(\omega_k-\alpha)\sum_{s=0}^{n_T-1} h^4(s) + O\left(b_{W,T}^{-2}\right),
\end{aligned}
$$

uniformly in α. We can expand (S.14) as follows,

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

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

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

and

$$
|R_T| \leq C\sum_{k=-\infty}^{\infty} |w(b_{W,T}k)|\,|k| \sim Cb_{W,T}^{-2}\int |x|\,|w(x)|\,dx,
$$

for $0 < C < \infty$. The latter result follows because

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

for a finite $0 < C < \infty$. A similar result holds for the second term involving $\left|H_{2,n_T}\left(\alpha+\beta\right)\right|^2$. Overall, we have

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

Equation (S.5) follows from

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

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

$$
\begin{aligned}
& \mathrm{cum}\left(f_{h,T}\left(u,\, \omega_1\right), \ldots,\, f_{h,T}\left(u,\, \omega_r\right)\right) \\
& \quad = 2\pi \left\{H_{2,n_T}\left(0\right)\right\}^{-r}
\end{aligned}
\tag{S.15}
$$

$$
\begin{aligned}
&\times \sum_{t_1=0}^{n_T-1} \cdots \sum_{t_{2r}=0}^{n_T-1} w\left(b_T\left(t_1-t_2\right)\right) \cdots w\left(b_T\left(t_{2r-1}-t_{2r}\right)\right) \\
&\times \exp\left(-i\omega_1\left(t_1-t_2\right)-\ldots-i\omega_r\left(t_{2r-1}-t_{2r}\right)\right) h_{n_T}\left(t_1\right) \cdots h_{n_T}\left(t_{2r}\right) \\
&\times \operatorname{cum}\left(X_{\lfloor Tu\rfloor-n_T/2+t_1+1,T} X_{\lfloor Tu\rfloor-n_T/2+t_2+1,T}, \ldots,\right. \\
&\quad \left. X_{\lfloor Tu\rfloor-n_T/2+t_{2r-1}+1,T} X_{\lfloor Tu\rfloor-n_T/2+t_{2r}+1,T}\right).
\end{aligned}
$$

Note that

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

where $c_{X\cdots X}\left(u;\, t_j,\, j\in v_1\right)$ is the time-Tu cumulant involving the variables X_{t_j} for $j\in v_1$ and where the summation is over all indecomposable partitions $\mathbf{v}=(v_1,\ldots,\, v_{\mathrm{P}})$ of the table

$$
\begin{vmatrix}
1 & 2 \\
3 & 4 \\
\vdots & \vdots \\
2r-1 & 2r
\end{vmatrix}.
$$

As the partition is indecomposable, in each set v_{p} of the partition we may find an element t^*_{p} such that none of $t_j-t^*_{\mathrm{p}}$, $j\in v_{\mathrm{p}}$ ($\mathrm{p}=1,\ldots,\mathrm{P}$) is $t_{2l-1}-t_{2l}$, $l=1,2,\ldots,r$. Define $2r-\mathrm{P}$ new variables $k_1,\ldots,\, k_{2r-\mathrm{P}}$ as the nonzero $t_j-t^*_{\mathrm{p}}$. Eq. (S.15) is now bounded by

$$
\begin{aligned}
C^r n_T^{-r} \sum_{\mathbf{v}} \sum_{t^*_1} \cdots \sum_{t^*_{\mathrm{P}}} \sum_{k_1} \cdots \sum_{k_{2r-\mathrm{P}}} \Big| & w\left(b_{W,T}\left(k_{\alpha_1}+t^*_{\beta_1}-k_{\alpha_1}-t^*_{\beta_2}\right)\right) \\
& \cdots \times w\left(b_{W,T}\left(k_{\alpha_{2r-1}}+t^*_{\beta_{2r-1}}-k_{\alpha_{2r}}-t^*_{\beta_{2r}}\right)\right)\Big| \\
& \times \left|h\left(t^*_1/n_T\right)\right|^{2r} \left|c_{X\cdots X}\left(u;\, k_1,\ldots\right) \cdots c_{X\cdots X}\left(u;\, \ldots,\, k_{2r-\mathrm{P}}\right)\right|,
\end{aligned}
$$

for some finite C, where $\alpha_1,\ldots,\, \alpha_{2r}$ are selected from $1,\ldots,\, 2r$ and $\beta_1,\ldots,\, \beta_{2r}$ from $1,\ldots,\, \mathrm{P}$. By Lemma 2.3.1 in Brillinger (1975), there are $\mathrm{P}-1$ linearly independent differences among the $t_{\beta^*_1}-t_{\beta^*_2},\ldots,\, t_{\beta^*_{2r-1}}-t_{\beta^*_{2r}}$. Suppose these are $t_{\beta^*_1}-t_{\beta^*_2},\ldots,\, t_{\beta^*_{2r-2}}-t_{\beta^*_{2r-1}}$. Making the change of variables

$$
\begin{aligned}
s_1 &= k_{\alpha_1}+t^*_{\beta_1}-k_{\alpha_1}-t^*_{\beta_2} \\
&\vdots \\
s_{\mathrm{P}-1} &= k_{\alpha_{2\mathrm{P}-3}}+t^*_{\beta_{2\mathrm{P}-3}}-k_{\alpha_{2\mathrm{P}-2}}-t^*_{\beta_{2\mathrm{P}-2}},
\end{aligned}
$$

the cumulant (S.15) is bounded by

$$
\begin{aligned}
C^r n_T^{-r} \sum_{\mathbf{v}} \sum_{t^*_1} \sum_{s_1} \cdots \sum_{s_{\mathrm{P}-1}} \sum_{k_1} \cdots \sum_{k_{2r-\mathrm{P}}} & \left|w\left(b_{W,T}s_1\right) \cdots w\left(b_{W,T}s_{P-1}\right)\right| \\
& \left|h\left(t^*_1/n_T\right)\right|^{2r} \left|c_{X\cdots X}\left(u;\, k_1,\ldots\right) \cdots c_{X\cdots X}\left(u;\, \ldots,\, k_{2r-\mathrm{P}}\right)\right| \\
\leq C^r n_T^{-r+1} b_{W,T}^{-(\mathrm{P}-1)} & \sum_{\mathbf{v}} C_{n_{2,1}} \cdots C_{n_{2,\mathrm{P}}}
\end{aligned}
$$

$$= O\left(n_T^{-r+1} b_{W,T}^{-(\mathrm{P}-1)}\right),$$

where $\mathrm{P} \leq r$ and $C_{n_{2},j} = \sup_{u \in [0,1]} \sum_{t_1,\ldots,t_{n_{2,j}}} |c_{X \cdots X}(u;\, t_1, \ldots, t_{n_{2,j}})|$ with $n_{2,j}$ denoting the number of elements in the jth set of the partition $\mathbf{v}$. It follows that for $r > 2$,

$$\mathrm{cum}\left(\left(n_T b_{W,T}\right)^{1/2} f_{h,T}(u,\, \omega_1), \ldots, \left(n_T b_{W,T}\right)^{1/2} f_{h,T}(u,\, \omega_r)\right) \to 0.$$

Thus, the variates $f_{h,T}(u,\, \omega_1), \ldots, f_{h,T}(u,\, \omega_r)$ are asymptotically normal with the moment structure given in the theorem. □

S.B.3 Proof of the Results of Section 3

S.B.3.1 Preliminary Lemmas

Let $I_T^*(j/T,\, \omega) = I_{L,h,T}(j/T,\, \omega) - \mathbb{E}(I_{L,h,T}(j/T,\, \omega))$. For $w \geq 0$ consider the dependence measure,

$$\phi_{I,w,q} = \sup_{j \in \{\mathbf{S}_r;\, r=1,\ldots,\, M_T-2\}} \left\| I_{h,T}^*(j/T,\, \omega) - I_{h,T,\{w\}}^*(j/T,\, \omega) \right\|_q, \tag{S.16}$$

where $X_{t,T}$ in $I_{h,T,\{w\}}^*(j/T,\, \omega)$ is replaced by $X_{t,T,\{w\}}$. Let $\Upsilon_{n,q} = \sum_{j=n}^{\infty} \phi_{I,j,q}$.

Lemma S.B.5. *Let Assumption 3.1-3.2 hold. We have $I_{h,T}^*(j/T,\, \omega) \in \mathscr{L}^q$ and, for $q > 2$, $\Upsilon_{n,q} = O(n^{-\gamma})$ for some $\gamma > 0$.*

Proof of Lemma S.B.5. We have

$$\phi_{I,w,q} = \sup_{j \in \{\mathbf{S}_r;\, r=1,\ldots,\, M_T-2\}} \Bigg\| \frac{1}{2\pi H_{2,n_T}(0)} \sum_{s=0}^{n_T-1} \sum_{t=0}^{n_T-1} h\left(\frac{s}{n_T}\right) h\left(\frac{t}{n_T}\right)$$
$$\left(X_{j-n_T+s+1,T} X_{j-n_T+t+1,T} - X_{j-n_T+s+1,T,\{w\}} X_{j-n_T+t+1,T,\{w\}}\right) \exp\left(-i\omega(s-t)\right) \Bigg\|_q .$$

Note that

$$\left\| \frac{1}{2\pi H_{2,n_T}(0)} \sum_{s=0}^{n_T-1} \sum_{t=0}^{n_T-1} h\left(\frac{s}{n_T}\right) h\left(\frac{t}{n_T}\right) \left(X_{j-n_T+s+1,T}\left(X_{j-n_T+t+1,T} - X_{j-n_T+t+1,T,\{w\}}\right)\right) \right\|_q$$
$$\leq \frac{1}{2\pi H_{2,n_T}(0)} \sum_{s=0}^{n_T-1} \sum_{t=0}^{n_T-1} \int_{-\pi}^{\pi} \left| h\left(\frac{s}{n_T}\right) h\left(\frac{t}{n_T}\right) \exp\left(-i\omega(s-t)\right) \right|$$
$$\times \left\| X_{j-n_T+s+1,T} \right\|_q \left\| X_{j-n_T+t+1,T} - X_{j-n_T+t+1,T,\{w\}} \right\|_q$$
$$\leq C \frac{1}{2\pi H_{2,n_T}(0)} \sum_{s=0}^{n_T-1} \sum_{t=0}^{n_T-1} \int_{-\pi}^{\pi} \left| h\left(\frac{s}{n_T}\right) h\left(\frac{t}{n_T}\right) \exp\left(-i\omega(s-t)\right) \right|$$
$$\times \sup_j \left\| X_{j-n_T+t+1,T} - X_{j-n_T+t+1,T,\{w\}} \right\|_q$$
$$\leq C \frac{1}{2\pi H_{2,n_T}(0)} \sum_{s=0}^{n_T-1} \sum_{t=0}^{n_T-1} \int_{-\pi}^{\pi} \left| h\left(\frac{s}{n_T}\right) h\left(\frac{t}{n_T}\right) \exp\left(-i\omega(s-t)\right) \right| \phi_{w,q}. \tag{S.17}$$

By Lemma A.7 in Dahlhaus (1997), we have $\sum_{s=0}^{n_T-1} |h(s/n_T)\exp(-i\omega s)| \leq Cn_T^{-1}L_{n_T}(\omega)^2$ for some $C<\infty$. Using Lemma S.B.1-(iii), the right-hand side of (S.17) is less than or equal to

$$C\frac{K^2 n_T^{-2} n_T^3}{2\pi H_{2,n_T}(0)}\phi_{w,q} \leq C_2\phi_{w,q},$$

for some $C_2<\infty$. Overall, we obtain $\phi_{I,w,q} \leq C_2\phi_{w,q}$ and so $\Upsilon_{n,q} \leq C_2\sum_{j=n}^{\infty}\phi_{j,q}$. Using Assumption 3.2 we have $\Upsilon_{n,q} = C_2 O(n^{-q+1})$ since $(\sum_{j=n}^{\infty}\phi_{j,q}^2)^{1/2} \leq \sum_{j=n}^{\infty}\phi_{j,q}$ and $\sum_{n=0}^{\infty} n^{q-1}\phi_{n,q}<\infty$. □

The result in Lemma S.B.5 also holds for $I_T^*(j/T,\omega)$ constructed using $I_{R,h,T}(j/T,\omega)$ in place of $I_{L,h,T}(j/T,\omega)$. 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 $\gamma>0$.

The proofs below can be simplified by noting that $\widehat{\sigma}_{L,r}^2(\omega)$ is a consistent estimate of $\sigma_{L,r}^2(\omega) = \mathrm{Var}(\sqrt{M_{S,T}}\widetilde{f}_{L,r,T}^*(\omega))$ where

$$\widetilde{f}_{L,r,T}^*(\omega) = M_{S,T}^{-1}\sum_{j\in\mathbf{S}_r} f_{L,h,T}^*(j/T,\omega)$$

with $f_{L,h,T}^*(j/T,\omega) = f_{L,h,T}(j/T,\omega) - \mathbb{E}(f_{L,h,T}(j/T,\omega))$. The consistency result follows from results in Casini (2023). The rate of convergence of $\widehat{\sigma}_{L,r}^2(\omega)$ is $O(\sqrt{\widetilde{M}_{S,T}b_{1,T}})$. Given Assumption 3.3-(iv), $(\widetilde{M}_{S,T}b_{1,T})^{-1/2}M_{S,T}^{1/2}\log T \to 0$ and so one can replace $\widehat{\sigma}_{L,r}(\omega)$ by $\sigma_{L,r}(\omega)$ in the definition of $\mathrm{S}_{\max,T}(\omega)$ and $\mathrm{S}_{\mathrm{Dmax},T}$ throughout the proofs of Lemma S.B.6 and of Theorem 3.1-3.2.

Let $\widetilde{f}_{r,T}(\omega)$ and $\sigma_{f,r}(\omega)$ be defined as $\widetilde{f}_{L,r,T}(\omega)$ and $\sigma_{L,r}(\omega)$, respectively, with $f_{L,h,T}(j/T,\omega)$ replaced by $f_{h,T}(j/T,\omega)$.

Lemma S.B.6. *Let Assumption 3.1-3.4 and 3.6 hold. Under $\mathcal{H}_0$, $\sqrt{\log(M_T)}M_{S,T}^{1/2}(\mathrm{S}_{\max,T}(\omega)-\widetilde{\mathrm{S}}_{\max,T}(\omega)) \overset{\mathbb{P}}{\to} 0$ for any $\omega\in[-\pi,\pi]$ where*

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

Proof of Lemma S.B.6. Note that for arbitrary sequences of numbers $(a_i)_{i=1,\ldots,N}$ and $(b_i)_{i=1,\ldots,N}$ with $N\geq 1$, we have for any i,

$$|a_i| \leq |a_i-b_i|+|b_i| \leq \max_{i=1,\ldots,N}|a_i-b_i| + \max_{i=1,\ldots,N}|b_i|. \tag{S.18}$$

The inequality still holds if on the left-hand side we replace $|a_i|$ by $\max_{i=1,\ldots,N}|a_i|$. We then have

$$\begin{aligned}\mathrm{S}_{\max,T}(\omega) &- \widetilde{\mathrm{S}}_{\max,T}(\omega)\\ &= \max_{r=1,\ldots,M_T-2}\left|\frac{\widetilde{f}_{L,r,T}(\omega)-\widetilde{f}_{R,r+1,T}(\omega)}{\sigma_{L,r}(\omega)}\right| - \max_{r=1,\ldots,M_T-2}\left|\frac{\widetilde{f}_{r,T}(\omega)-\widetilde{f}_{r+1,T}(\omega)}{\sigma_{f,r}(\omega)}\right|.\end{aligned} \tag{S.19}$$

Using (S.18) the right-hand side of (S.19) is less than or equal to

$$\max_{r=1,\ldots,M_T-2}\left|\frac{\widetilde{f}_{L,r,T}(\omega)-\widetilde{f}_{R,r+1,T}(\omega)}{\sigma_{L,r}(\omega)} - \frac{\widetilde{f}_{r,T}(\omega)-\widetilde{f}_{r+1,T}(\omega)}{\sigma_{L,r}(\omega)}\right|$$

$$+ \max_{r=1,\dots,\,M_T-2} \left| \frac{\widetilde{f}_{r,T}(\omega) - \widetilde{f}_{r+1,T}(\omega)}{\sigma_{L,r}(\omega)} \right| - \max_{r=1,\dots,\,M_T-2} \left| \frac{\widetilde{f}_{r,T}(\omega) - \widetilde{f}_{r+1,T}(\omega)}{\sigma_{f,r}(\omega)} \right|.$$

The second line converges to zero in probability given the uniform asymptotic equivalence of $\sigma_{L,r}(\omega)$ and $\sigma_{f,r}(\omega)$ with an error $O(T^{-1})$. Thus, it is sufficient to show

$$\max_{r=1,\dots,\,M_T-2} \sqrt{\log(M_T)} M_{S,T}^{1/2} \left| \frac{\widetilde{f}_{L,r,T}(\omega) - \widetilde{f}_{R,r+1,T}(\omega) - \left(\widetilde{f}_{r,T}(\omega) - \widetilde{f}_{r+1,T}(\omega)\right)}{\sigma_{L,r}(\omega)} \right| \xrightarrow{\mathbb{P}} 0.$$

We use the following decomposition,

$$\begin{aligned} \sqrt{\log(M_T)} M_{S,T}^{1/2} & \left| \frac{\widetilde{f}_{L,r,T}(\omega) - \widetilde{f}_{R,r+1,T}(\omega) - \left(\widetilde{f}_{r,T}(\omega) - \widetilde{f}_{r+1,T}(\omega)\right)}{\sigma_{L,r}(\omega)} \right| \\ & \leq \sqrt{\log(M_T)} M_{S,T}^{1/2} \left| \frac{\widetilde{f}_{L,r,T}(\omega) - \widetilde{f}_{r,T}(\omega)}{\sigma_{L,r}(\omega)} \right| \\ & \quad + \sqrt{\log(M_T)} M_{S,T}^{1/2} \left| \frac{\left(\widetilde{f}_{R,r+1,T}(\omega) - \widetilde{f}_{r+1,T}(\omega)\right)}{\sigma_{L,r}(\omega)} \right|. \end{aligned} \tag{S.20}$$

Let us consider the first term on the right-hand side of (S.20). Note that $\widetilde{f}_{L,r,T}(\omega)$ and $\widetilde{f}_{r,T}(\omega)$ are weighted averages of stochastic θ-Hölder continuous variables each standardized by $\sigma_{L,r}(\omega)$. These variables standardized by $\sigma_{L,r}(\omega)$ belong to the same block r of the sample. Thus, their weighted average is asymptotically approximated by a stochastic variable that satisfies stochastic θ-Hölder continuity as for $X_{t,T}$ in (3.5), where this property holds uniformly in r. Thus, it follows that

$$\begin{aligned} \sigma_{L,r}^{-1}(\omega) \sqrt{\log(M_T)} M_{S,T}^{1/2} & \left| \widetilde{f}_{L,r,T}(\omega) - \widetilde{f}_{r,T}(\omega) \right| \\ & = \sqrt{\log(M_T)} M_{S,T}^{1/2} \left(O\left((m_T/T)^{\theta}\right) + O_{\mathbb{P}}\left((n_T b_{W,T})^{-1/2}\right)\right) \\ & = o_{\mathbb{P}}(1), \end{aligned}$$

where the $O_{\mathbb{P}}((n_T b_{W,T})^{-1/2})$ rate follows from Theorem S.A.6 and the last equality uses Assumption 3.6. Using Markov's inequality, this shows that for all $\epsilon > 0$,

$$\mathbb{P}\left(\max_{r=1,\dots,\,M_T-2} \sqrt{\log(M_T)} M_{S,T}^{1/2} \left| \widetilde{f}_{L,r,T}(\omega) - \widetilde{f}_{r,T}(\omega) \right| > \epsilon \right) \to 0.$$

The argument for the second term of (S.20) is analogous. □

Lemma S.B.7. *Let Assumption 3.1-3.4 and 3.6 hold. Under $\mathcal{H}_0$, we have $\sqrt{\log(M_T)} M_{S,T}^{1/2}(\mathrm{R}_{\max,T}(\omega) - \widetilde{\mathrm{R}}_{\max,T}(\omega)) \xrightarrow{\mathbb{P}} 0$ for any $\omega \in [-\pi, \pi]$, where*

$$\widetilde{\mathrm{R}}_{\max,T}(\omega) \triangleq \max_{r=1,\dots,\,M_T-2} \left| \frac{\widetilde{f}_{r,T}(\omega)}{\widetilde{f}_{r+1,T}(\omega)} - 1 \right|.$$

Proof of Lemma S.B.7. Using (S.18), we have

$$\begin{aligned}\left|\mathrm{R}_{\max,T}(\omega)-\widetilde{\mathrm{R}}_{\max,T}(\omega)\right| &\leq \max_{r=1,\ldots,M_T-2}\left|\frac{\widetilde{f}_{L,r,T}(\omega)}{\widetilde{f}_{R,r+1,T}(\omega)}-1-\left(\frac{\widetilde{f}_{r,T}(\omega)}{\widetilde{f}_{r+1,T}(\omega)}-1\right)\right| \\ &\leq \max_{r=1,\ldots,M_T-2}\left|\widetilde{f}_{L,r,T}(\omega)\left(\frac{1}{\widetilde{f}_{R,r+1,T}(\omega)}-\frac{1}{\widetilde{f}_{r+1,T}(\omega)}\right)\right| \\ &\quad+\max_{r=1,\ldots,M_T-2}\left|\frac{\widetilde{f}_{L,r,T}(\omega)-\widetilde{f}_{r,T}(\omega)}{\widetilde{f}_{r+1,T}(\omega)}\right|.\end{aligned} \tag{S.21}$$

Let us consider the second term on the right-hand side of (S.21). Note that for all $\epsilon>0$ and all constants $C>0$, we have

$$\begin{aligned}&\mathbb{P}\left(\max_{r=1,\ldots,M_T-2}\left|\sqrt{\log(M_T)}M_{S,T}^{1/2}\frac{\widetilde{f}_{L,r,T}(\omega)-\widetilde{f}_{r,T}(\omega)}{\widetilde{f}_{r+1,T}(\omega)}\right|>\epsilon\right) \\ &\leq \mathbb{P}\left(\max_{r=1,\ldots,M_T-2}\sqrt{\log(M_T)}M_{S,T}^{1/2}\left|\widetilde{f}_{L,r,T}(\omega)-\widetilde{f}_{r,T}(\omega)\right|\cdot\max_{r=1,\ldots,M_T-2}\left|\frac{1}{\widetilde{f}_{r+1,T}(\omega)}\right|>\epsilon\right) \\ &\leq \mathbb{P}\left(\max_{r=1,\ldots,M_T-2}\sqrt{\log(M_T)}M_{S,T}^{1/2}\left|\widetilde{f}_{L,r,T}(\omega)-\widetilde{f}_{r,T}(\omega)\right|>\frac{\epsilon}{C}\right) \\ &\quad+\leq \mathbb{P}\left(\max_{r=1,\ldots,M_T-2}\left|\frac{1}{\widetilde{f}_{r+1,T}(\omega)}\right|>C\right).\end{aligned} \tag{S.22}$$

Theorem S.A.5 implies that

$$\mathbb{E}\left(f_{h,T}(u,\omega)\right)=f(u,\omega)+O\left((n_T/T)^{-2}\right)+O\left(\log(n_T)\,n_T^{-1}\right)+o\left(b_{W,T}^2\right).$$

The same result holds for $f_{L,h,T}(u,\omega)$. By Assumption 4.1 we have

$$f\left(\left((r+1)m_T+j\right)/T,\omega\right)-f\left((rm_T+j)/T,\omega\right)=O\left((m_T/T)^{\theta}\right),\qquad \text{uniformly in } r \text{ and } j. \tag{S.23}$$

Thus, using the bound for the variance of $f_{h,T}(u,\omega)$ in Theorem S.A.6 and Assumption 3.6, we have

$$\begin{aligned}&\max_{r=1,\ldots,M_T-2}\sqrt{\log(M_T)}M_{S,T}^{1/2}\left|\widetilde{f}_{L,r,T}(\omega)-\widetilde{f}_{r,T}(\omega)\right| \\ &\qquad=\sqrt{\log(M_T)}M_{S,T}^{1/2}\left(O\left((m_T/T)^{\theta}\right)+O_{\mathbb{P}}\left((n_Tb_{W,T})^{-1/2}\right)\right) \\ &\qquad=o_{\mathbb{P}}(1).\end{aligned}$$

By using Markov's inequality, this shows that

$$\mathbb{P}\left(\max_{r=1,\ldots,M_T-2}\sqrt{\log(M_T)}M_{S,T}^{1/2}\left|\widetilde{f}_{L,r,T}(\omega)-\widetilde{f}_{r,T}(\omega)\right|>\frac{\epsilon}{C}\right)\to 0. \tag{S.24}$$

By Theorem S.A.6, $\widetilde{f}_{r+1,T}(\omega)=f\left((r+1)m_T,\omega\right)+o_{\mathbb{P}}(1)$. Thus, the second term of (S.22) also converges to zero for example by choosing $C=3/f_{-}$. Altogether we obtain that the right-hand side of (S.22) converges to zero. Next, we consider the first term of (S.21). For any $\epsilon > 0$ and any $C > 0$, we have

$$\begin{aligned} &\mathbb{P}\left(\max_{r=1,\ldots,M_T-2} \sqrt{\log(M_T)} M_{S,T}^{1/2} \left| \widetilde{f}_{L,r,T}(\omega) \left(\frac{1}{\widetilde{f}_{R,r+1,T}(\omega)} - \frac{1}{\widetilde{f}_{r+1,T}(\omega)} \right) \right| > \epsilon \right) \\ &\leq \mathbb{P}\left(\max_{r=1,\ldots,M_T-2} \sqrt{\log(M_T)} M_{S,T}^{1/2} \widetilde{f}_{L,r,T}(\omega) \left| \widetilde{f}_{r+1,T}(\omega) - \widetilde{f}_{R,r+1,T}(\omega) \right| > \frac{\epsilon}{C} \right) \\ &\quad + \mathbb{P}\left(\max_{r=1,\ldots,M_T-2} \left| \frac{1}{\widetilde{f}_{r+1,T}(\omega) \widetilde{f}_{R,r+1,T}(\omega)} \right| > C \right). \end{aligned} \tag{S.25}$$

The first term on the right-hand side above is less than or equal to,

$$\begin{aligned} &\mathbb{P}\left(\max_{r=1,\ldots,M_T-2} \left| \widetilde{f}_{L,r,T}(\omega) \right| > C_2 \right) \\ &\quad + \mathbb{P}\left(\max_{r=1,\ldots,M_T-2} \sqrt{\log(M_T)} M_{S,T}^{1/2} \left| \widetilde{f}_{r+1,T}(\omega) - \widetilde{f}_{R,r+1,T}(\omega) \right| > \frac{\epsilon}{C \cdot C_2} \right), \end{aligned}$$

for all $C_2 > 0$. We can choose C_2 large enough such that the first term above converges to zero. The second term above converges to zero by the same argument as in (S.24). The second term on the right-hand side of (S.25) can be expanded as follows,

$$\begin{aligned} &\mathbb{P}\left(\max_{r=1,\ldots,M_T-2} \left| \frac{1}{\widetilde{f}_{r+1,T}(\omega) \widetilde{f}_{R,r+1,T}(\omega)} \right| > C \right) \\ &\leq \mathbb{P}\left(\min_{r=1,\ldots,M_T-2} \left| \widetilde{f}_{R,r+1,T}(\omega) \right| < C^{-1/2} \right) + \mathbb{P}\left(\min_{r=1,\ldots,M_T-2} \left| \widetilde{f}_{r+1,T}(\omega) \right| < C^{-1/2} \right) \\ &\leq \mathbb{P}\left(\min_{r=1,\ldots,M_T-2} \left| \widetilde{f}_{R,r+1,T}(\omega) \right| < C^{-1/2} \right) + \mathbb{P}\left(\min_{r=1,\ldots,M_T-2} \left| \widetilde{f}_{R,r+1,T}(\omega) \right| < 2C^{-1/2} \right) \\ &\quad + \mathbb{P}\left(\max_{r=1,\ldots,M_T-2} \left| \widetilde{f}_{R,r+1,T}(\omega) - \widetilde{f}_{r+1,T}(\omega) \right| > 2C^{-1/2} \right). \end{aligned}$$

The first two terms on the right-hand side have already been discussed above. The third term has also been discussed above with the multiplicative factor $\sqrt{\log(M_T)} M_{S,T}^{1/2}$. □

S.B.3.2 Proof of Lemma 3.1

Let $J_t = \kappa_{\mathbf{X},t}(k_1, \cdots, k_{r-1})$ where $t \leq t + k_1 \leq \ldots \leq t + k_{r-1}$. For $1 \leq l \leq r-1$, let $n_l = k_l - k_{l-1}$. Define the vector $y_0 = y_{0,l} = (k_1 - k_{l-1}, \cdots, k_{l-2} - k_{l-1}, 0)$. Let $\mathcal{G}_n = (\ldots, e'_{n-1}, e'_n)$ where $n \in \mathbb{N}$ and for $t \geq 0$ define $X^*_{t,T} = H(t/T, \{\mathcal{G}_0, e_1, \cdots, e_t\})$. Following Proposition 2 in Wu and Shao (2004) and by the additivity of cumulants,

$$\begin{aligned} \sup_t |J_t| &= \sup_{k_{l-1}, l \in \{1,\ldots,r-1\}} \left| \kappa_{\mathbf{X},-k_{l-1}}(y_0, k_l - k_{l-1}, k_{l+1} - k_{l-1}, \cdots, k_{r-1} - k_{l-1}) \right| \\ &= \left| \kappa_{\mathbf{X},-k_{l_*-1}}(y_0^*, k_{l_*} - k_{l_*-1}, k_{l_*+1} - k_{l_*-1}, \cdots, k_{r-1} - k_{l_*-1}) \right| \\ &\leq \sum_{j=0}^{r-l_*-1} \left| \mathrm{cum}\Big(Y_0, X^*_{k_{l_*} - k_{l_*-1},T}, \cdots, X^*_{k_{l_*+j-1} - k_{l_*-1},T}, \right. \end{aligned} \tag{S.26}$$

$$X_{k_{l_*+j}-k_{l_*-1},T}-X^*_{k_{l_*+j}-k_{l_*-1},T},\,X_{k_{l_*+j+1}-k_{l_*-1},T},\,\cdots,\,X_{k_{k-1}-k_{l_*-1},T}\Big)\Big|$$
$$=\sum_{j=0}^{r-l_*-1}B_j\left(l_*\right),$$

where $y_0^*=y_{0,l_*}^*=(k_1-k_{l_*-1},\cdots,\,k_{l_*-2}-k_{l_*-1},\,0)$, and $Y_0=Y_{0,l_*}=(X_{k_1-k_{l_*-1},T},\cdots,\,X_{k_{l_*-2}-k_{l_*-1},T},\,X_{0,T})$. Let $\zeta_t=||X_{t,T}-X^*_{j,T}||_r$ and $S_j=S_{j,r}=\sum_{i=j}^{\infty}\phi_{i,r}^2$. By Proposition 2 in Wu and Shao (2004) we have $|B_j\left(l_*\right)|\leq C_1\zeta_{k_{l_*+j}-k_{l_*-1}}$, where C_1 only depends on r and on $\sup_t\mathbb{E}|X_{t,T}|^m$ with $1\leq m\leq r$. Thus, using (S.26) and Proposition 2 of Wu (2007), we have

$$\sup_t|J_t|\leq C_1\sum_{j=0}^{r-l_*-1}\zeta_{k_{l_*+j}-k_{j-1}}\leq C_2\sum_{j=0}^{r-l_*-1}S^{1/2}_{k_{l_*+j}-k_{j-1}}\leq C_3S^{1/2}_{n_{l_*}},$$

where $C_2=18r^{3/2}\left(r-1\right)^{-1/2}C_1$ and $C_3=C_2r$. Since $1\leq l_*\leq r-1$, we have $\sup_t|J_t|\leq C_3S^{1/2}_{n_{l_*}}$. Then,

$$\begin{aligned}\sum_{k_1,\dots,\,k_{r-1}=-\infty}^{\infty}\left(1+|k_j|^l\right)\sup_t|\kappa_{\mathbf{X},t}\left(k_1,\cdots,\,k_{r-1}\right)|&\leq 2r!\sum_{k_1,\dots,\,k_{r-1}=0}^{\infty}|k_{r-1}|^l\sup_t|\kappa_{\mathbf{X},t}\left(k_1,\cdots,\,k_{r-1}\right)|\\&\leq 2r!\sum_{k_1,\dots,\,k_{r-1}=0}^{\infty}|k_{r-1}|^lC_3S^{1/2}_{n_{l_*}}\\&\leq 2C_3r!\sum_{n=0}^{\infty}n^{l+r-2}S_n^{1/2}\\&<\infty,\end{aligned}$$

where the last inequality follows from Assumption 3.2. □

S.B.3.3 Proof of Theorem 3.1

From Lemma S.B.6 it is sufficient to show the result for $\widetilde{\mathrm{S}}_{\max,T}\left(\omega\right)$ since the latter is asymptotically equivalent to $\mathrm{S}_{\max,T}\left(\omega\right)$. Define $f^*_{h,T}\left(j/T,\,\omega\right)=f_{h,T}\left(j/T,\,\omega\right)-\mathbb{E}\left(f_{h,T}\left(j/T,\,\omega\right)\right)$. For $\omega\in[-\pi,\,\pi]$ let $S_{r+1}\left(\omega\right)=\sum_{j\in\{\mathbf{S}_s,\,s=1,\dots,\,r+1\}}f^*_{h,T}\left(j/T,\,\omega\right)$ and

$$R_{r,T}\left(\omega\right)=\frac{1}{M_{S,T}}\left(S_{r+1}\left(\omega\right)-\sum_{j\in\{\mathbf{S}_s,\,s=1,\dots,\,r+1\}}\mathscr{W}_j\left(\omega\right)-\left(S_r\left(\omega\right)-\sum_{j\in\{\mathbf{S}_s,\,s=1,\dots,\,r\}}\mathscr{W}_j\left(\omega\right)\right)\right),$$

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

$$\begin{aligned}\widetilde{f}_{r,T}\left(\omega\right)&=M_{S,T}^{-1}\sum_{j\in\mathbf{S}_r}f_{h,T}\left(j/T,\,\omega\right)\\&=M_{S,T}^{-1}\sum_{j\in\mathbf{S}_r}\left(f^*_{h,T}\left(j/T,\,\omega\right)+\mathbb{E}\left(f_{h,T}\left(j/T,\,\omega\right)\right)\right)\\&=\frac{1}{M_{S,T}}\sum_{j\in\mathbf{S}_r}\mathscr{W}_j\left(\omega\right)+R_{r,T}+\frac{1}{M_{S,T}}\sum_{j\in\mathbf{S}_r}\mathbb{E}\left(f_{h,T}\left(j/T,\,\omega\right)\right).\end{aligned}\tag{S.27}$$

Under Assumption 3.2, Theorem 1 in Wu and Zhou (2011) yields $\max_{0\leq r\leq M_{S,T}-1}|R_{r,T}|=O_{\mathbb{P}}\left(\tau_T/M_{S,T}\right)$. By Theorem S.A.5,

$$\mathbb{E}\left(f_{h,T}\left(j/T,\,\omega\right)\right)=f\left(j/T,\,\omega\right)+O\left(\left(n_T/T\right)^2\right)+O\left(b_{W,T}^2\right)+O\left(\log\left(n_T\right)/n_T\right).$$

Using (S.23) we obtain

$$\begin{aligned}&\sqrt{M_{S,T}}\left(\widetilde{f}_{r+1,T}\left(\omega\right)-\widetilde{f}_{r,T}\left(\omega\right)\right)\\&=\frac{1}{\sqrt{M_{S,T}}}\left(\sum_{j\in\mathbf{S}_{r+1}}\mathscr{W}_j\left(\omega\right)-\sum_{j\in\mathbf{S}_r}\mathscr{W}_j\left(\omega\right)\right)\\&\quad+O\left(M_{S,T}^{1/2}m_T^{\theta}/T^{\theta}\right)+O_{\mathbb{P}}\left(\tau_T/M_{S,T}^{1/2}\right)+O_{\mathbb{P}}\left(M_{S,T}^{1/2}\left(n_T/T\right)^2+M_{S,T}^{1/2}b_{W,T}^2+M_{S,T}^{1/2}\log\left(n_T\right)/n_T\right)\\&=\frac{1}{\sqrt{m_T}}\left(\sum_{j\in\mathbf{S}_{r+1}}\mathscr{W}_j\left(\omega\right)-\sum_{j\in\mathbf{S}_r}\mathscr{W}_j\left(\omega\right)\right)\\&\quad+o_{\mathbb{P}}\left(\left(\log M_T\right)^{-1/2}\right).\end{aligned}\tag{S.28}$$

The result then follows from Lemma 1 in Wu and Zhao (2007). □

S.B.3.4 Proof of Theorem 3.2

Lemma S.B.8. *Let $\mathscr{V}\left(\omega\right)$ denote a random variable defined by $\mathbb{P}\left(\mathscr{V}\left(\omega\right)\leq v\right)=\exp(-\pi^{-1/2}\exp(-v))$ for $\omega\in\Pi$. Assume that for $\omega,\,\omega'\in\Pi$ the variables $\mathscr{V}\left(\omega\right)$ and $\mathscr{V}\left(\omega'\right)$ are independent. Let $\mathscr{V}^*\triangleq\max_{\omega\in\Pi}\mathscr{V}\left(\omega\right)-\log\left(n_\omega\right)$. Then, $\mathbb{P}\left(\mathscr{V}^*\leq v\right)=\exp(-\pi^{-1/2}\exp\left(-v\right))$.*

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

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

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

Proof of Theorem 3.2. From Theorem S.A.6, it follows that $f_{h,T}\left(u,\,\omega_j\right)$ and $f_{h,T}\left(u,\,\omega_k\right)$ are asymptotically independent if $\omega_k\pm\omega_k\not\equiv0\left(\mathrm{mod}\,2\pi\right)$, $1\leq j<k\leq n_\omega$. The result then follows from Lemma S.B.6 and S.B.8, and Theorem 3.1. □

S.B.3.5 Proof of Theorem 3.3

Due to the self-normalization nature of the test statistic, we can use Lemma S.B.7 and steps similar to Proposition A1-A.3 in Bibinger et al. (2017) to show that it is sufficient to consider the behavior of

$$\widetilde{\mathrm{R}}^*\left(\omega\right)=\max_{r=1,\dots,\,M_T-2}\left|M_{S,T}^{-1}\sum_{j\in\mathbf{S}_r}g_T\left(j/T,\,\omega\right)-M_{S,T}^{-1}\sum_{j\in\mathbf{S}_{r+1}}g_T\left(j/T,\,\omega\right)\right|,$$

where $g_T(j/T,\,\omega)$ are random variables with mean $\mathbb{E}(f_{h,T}(j/T,\,\omega))$, unit variance and satisfying Assumption 3.2. For $\omega\in[-\pi,\,\pi]$ let $S_{r+1}(\omega)=\sum_{j\in\{\mathbf{S}_s,\,s=1,\dots,\,r+1\}}(g_T(j/T,\,\omega)-\mathbb{E}(f_{h,T}(j/T,\,\omega)))$ and

$$R_{r,T}(\omega)=\frac{1}{M_{S,T}}\left(S_{r+1}(\omega)-\sum_{j\in\{\mathbf{S}_s,\,s=1,\dots,\,r+1\}}\mathscr{W}_j(\omega)-\left(S_r(\omega)-\sum_{j\in\{\mathbf{S}_s,\,s=1,\dots,\,r\}}\mathscr{W}_j(\omega)\right)\right),$$

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

$$\begin{aligned}M_{S,T}^{-1}\sum_{j\in\mathbf{S}_r}g_T(j/T,\,\omega)&=M_{S,T}^{-1}\sum_{j\in\mathbf{S}_r}\left((g_T(j/T,\,\omega)-\mathbb{E}(g_T(j/T,\,\omega)))+\mathbb{E}(g_T(j/T,\,\omega))\right)\\&=\frac{1}{M_{S,T}}\sum_{j\in\mathbf{S}_r}\mathscr{W}_j(\omega)+R_{r,T}+\frac{1}{M_{S,T}}\sum_{j\in\mathbf{S}_r}\mathbb{E}(g_T(j/T,\,\omega)).\end{aligned}\tag{S.29}$$

As in the proof of Theorem 3.1, we have $\max_{0\leq r\leq M_{S,T}-1}|R_{r,T}|=O_{\mathbb{P}}(\tau_T/M_{S,T})$. By Theorem S.A.5, $\mathbb{E}(g_T(j/T,\,\omega))=f(j/T,\,\omega)+O\left((n_T/T)^2\right)+O\left(b_{W,T}^2\right)+O(\log(n_T)/n_T)$. Using (S.23), we obtain

$$\begin{aligned}&\sqrt{M_{S,T}}\left(M_{S,T}^{-1}\sum_{j\in\mathbf{S}_{r+1}}g_T(j/T,\,\omega)-M_{S,T}^{-1}\sum_{j\in\mathbf{S}_r}g_T(j/T,\,\omega)\right)\\&\quad=\frac{1}{\sqrt{M_{S,T}}}\left(\sum_{j\in\mathbf{S}_{r+1}}\mathscr{W}_j(\omega)-\sum_{j\in\mathbf{S}_r}\mathscr{W}_j(\omega)\right)\\&\qquad+O\left(M_{S,T}^{1/2}m_T^{\theta}/T^{\theta}\right)+O_{\mathbb{P}}\left(\tau_T/M_{S,T}^{1/2}\right)+O_{\mathbb{P}}\left(M_{S,T}^{1/2}(n_T/T)^2+M_{S,T}^{1/2}b_{W,T}^2+M_{S,T}^{1/2}\log(n_T)/n_T\right)\\&\quad=\frac{1}{\sqrt{m_T}}\left(\sum_{j\in\mathbf{S}_{r+1}}\mathscr{W}_j(\omega)-\sum_{j\in\mathbf{S}_r}\mathscr{W}_j(\omega)\right)\\&\qquad+o_{\mathbb{P}}\left((\log M_T)^{-1/2}\right).\end{aligned}\tag{S.30}$$

The result about $\mathrm{R}_{\max,T}(\omega)$ follows from Lemma 1 in Wu and Zhao (2007). The result concerning $\mathrm{R}_{\mathrm{Dmax},T}$ follows using the same argument as in the proof of Theorem 3.2. □

S.B.4 Proofs of the Results in Section 4

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.B.4.1 Proof of Theorem 4.1

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 et al. (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 standard normal variables.

We first consider alternatives as in $\mathcal{H}_1^{\mathrm{B}}$. Throughout the proof we set

$$m_T = C_T \left(\sqrt{\log (M_T)} T^{\theta} / D \right)^{\frac{2}{2\theta+1}}, \tag{S.31}$$

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 et al. (2017). We set $f_-(\omega_0) = 1$ and let

$$f(j/T, \omega_0) = \begin{cases} 1 + (m_T - j \bmod m_T)^{\theta} T^{-\theta}, & T\lambda_b^0 < j \leq T\lambda_b^0 + m_T \\ 1, & \text{else} \end{cases}. \tag{S.32}$$

We discuss the general case $f_-(\omega_0) \neq 1$ at the end of this proof. Eq. (S.32) 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 S.A.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 S.A.4-(ii) since by Assumption 4.1 $f(\cdot, \omega_0)$ 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 $S_J = \{n_T + 1, n_T + 1 + m_{S,T}, \ldots\}$. 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. Define $\widetilde{\zeta}_j^* = \widetilde{f}(j/T, \omega_0)\chi_2^2/2$ where $\widetilde{\zeta}_j^*$ are independent across j.

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

(i) Case $\theta > 1/2$. Let us consider the following distinct experiments:

$\mathcal{E}_1$: Observe $\{\zeta_j\}_{j=n_T+1}^T$ 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 $\{\zeta_j^*\}_{j=n_T+1}^T$ 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 $\{\widetilde{\zeta}_j^*\}_{j=n_T+1}^T$ 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 $\{\zeta_j\}$ and $\{\zeta_j^*\}$ 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 $\{\zeta_j\}$ 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:

$$\mathcal{E}_1\approx\mathcal{E}_2\approx\mathcal{E}_3\sim\mathcal{E}_4\approx\mathcal{E}_5. \tag{S.33}$$

Therefore, given the relation (S.33), the lower bound for $\mathcal{E}_5$ carries over to the less informative experiment $\mathcal{E}_1$. We prove (S.33) 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 3.2, we have $\sum_{j=1}^{\infty}\nu_j<\infty$, which in turn yields,

$$\sum_{j=1}^{\infty}\left|\zeta_j-\zeta_j^*\right|<\infty,\qquad\mathbb{P}-\text{almost surely}. \tag{S.34}$$

Note that

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

Choose J_1 large enough such that $\sum_{j\in S_J,\,j>J_1}\left|\zeta_j-\zeta_j^*\right|=o_{\mathrm{a.s}}(|S_J|)$. Thus, $|S_J|^{-1}\sum_{j\in S_J}\left|\zeta_j-\zeta_j^*\right|\to 0$ $\mathbb{P}$-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

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

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

$$\begin{aligned}D_{\mathrm{KL}}\left(\mathbb{P}_c||\,\mathbb{P}_{\widetilde{c}}\right)&=\log\left(\frac{\widetilde{c}+\delta}{\widetilde{c}}\right)+\frac{\widetilde{c}-(\widetilde{c}+\delta)}{\widetilde{c}+\delta}\\&=-\frac{\delta^2}{2\widetilde{c}^2}+O\left(\delta^2\right)+O\left(\delta^3\right).\end{aligned} \tag{S.35}$$

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.35) 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.31) and $m_T^{-1}a_T\rightarrow 0$.

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

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

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

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

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,

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

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

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

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 $Tm_T^{-2}\rightarrow 0$. Finally, since $m_T^{-1}a_T\rightarrow 0$, we can choose a_T increasing sufficiently fast such that $m_Ta_T^{-2}\rightarrow 0$. Thus, we have $||\mathbb{P}_\chi-\mathbb{P}_\xi||_{\mathrm{TV}}\rightarrow 0$.

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 the observations

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

and the hypothesis

$$\mathcal{H}_0 : \sup_j \left(\widetilde{f}\,(j/T,\, \omega_0) - 1\right) = 0 \qquad \text{versus} \qquad \mathcal{H}_1 : \sup_j \left(\widetilde{f}\,(j/T,\, \omega_0) - 1\right) \geq b_T. \tag{S.36}$$

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 $\lim_{T\to\infty} \inf_\psi \gamma_\psi\,(\theta,\, b_T) = 1$. 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

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

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 et al. (2017).

It remains to consider the case $\theta \leq 1/2$. In a different setting, Bibinger et al. (2017) considered separately the case where their regularity exponent $\mathfrak{a}$ satisfies $\mathfrak{a} \leq 1/2$ 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.36). 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 of the lower bound for the alternative $\mathcal{H}_1^{\text{S}}$ is similar to the proof discussed above. The minor differences in the proof outlined by Bibinger et al. (2017) also apply here. □

S.B.4.2 Proof of Theorem 4.2

We present the proof for the statistic $\text{S}_{\max,T}$. The proof for the other test statistics discussed in Section 3 is similar and omitted. From the same reasoning as in the proofs of the results of Section 3, we can replace $\widehat{\sigma}_{L,r}\,(\omega)$ by $\sigma_{L,r}\,(\omega)$ throughout the proof. Without loss of generality, we assume that $\omega_0 \neq \pm\pi$. Let $M_{S,T}^* = m_T^*/m_{S,T}^*$ and $m_{S,T}/m_T^* \to [0,\, \infty)$. If $\lfloor T\lambda_b^0 \rfloor \notin \{\{\mathbf{S}_r\} \cup \{\mathbf{S}_{r+1}\}\}$ or if $\omega \neq \omega_0$, then

$$\left|\frac{\widetilde{f}_{L,r,T}\,(\omega) - \widetilde{f}_{R,r+1,T}\,(\omega)}{\sigma_{L,r}\,(\omega)}\right|$$

$$
\begin{aligned}
&= \left| \frac{\left(M_{S,T}^*\right)^{-1} \sum_{j\in\mathbf{S}_r} \left(f_{L,h,T}^*\left(j/T,\, \omega\right) + \mathbb{E}\left(f_{L,h,T}\left(j/T,\, \omega\right)\right)\right)}{\sigma_{L,r}\left(\omega\right)} \right. \\
&\quad \left. - \frac{\left(M_{S,T}^*\right)^{-1} \sum_{j\in\mathbf{S}_{r+1}} \left(f_{R,h,T}^*\left(j/T,\, \omega\right) + \mathbb{E}\left(f_{R,h,T}\left(j/T,\, \omega\right)\right)\right)}{\sigma_{L,r}\left(\omega\right)} \right| \\
&= \left| \frac{\left(M_{S,T}^*\right)^{-1} \sum_{j\in\mathbf{S}_r} f_{L,h,T}^*\left(j/T,\, \omega\right) - \left(M_{S,T}^*\right)^{-1} \sum_{j\in\mathbf{S}_{r+1}} f_{R,h,T}^*\left(j/T,\, \omega\right)}{\sigma_{L,r}\left(\omega\right)} \right| \\
&\quad + O\left(\left(m_T^*/T\right)^{\theta}\right) + O_{\mathbb{P}}\left(\left(n_T/T\right)^2 + \log\left(n_T\right)/n_T\right) + o\left(b_{W,T}^2\right) \\
&\triangleq \mathring{f}_{r,T}\left(\omega\right) + O\left(\left(m_T^*/T\right)^{\theta}\right) + O_{\mathbb{P}}\left(\left(n_T/T\right)^2 + \log\left(n_T\right)/n_T\right) + o\left(b_{W,T}^2\right) \\
&= \mathring{f}_{r,T}\left(\omega\right) + o_{\mathbb{P}}\left(\left(\sqrt{m_T^*}\right)^{-1}\right),
\end{aligned}
$$

where the last inequality follows from (4.1). As in the proof of Theorem 3.1, we have $\sqrt{M_{S,T}^*}\mathring{f}_{r,T}\left(\omega\right) = O_{\mathbb{P}}\left(1\right)$ for $1 \leq r \leq M_T^* - 2$. This can be used to obtain the following inequality, if $\lfloor T\lambda_b^0 \rfloor \in \{\{\mathbf{S}_r\} \cup \{\mathbf{S}_{r+1}\}\}$ and $\omega = \omega_0$,

$$
\begin{aligned}
\mathrm{S}_{\max,T}\left(\omega_0\right) &\geq -\mathring{f}_{r,T}\left(\omega_0\right) \\
&\quad + \frac{T}{m_T^*} \left| \int_{\left(rm_T^* - m_T^*/2 + n_T/2 + 1\right)/T}^{\lambda_b^0} f\left(u,\, \omega_0\right) du - \int_{\lambda_b^0}^{\left((r+1)m_T^* + n_T/2 + m_{S,T}M_{S,T}/2\right)/T} f\left(u,\, \omega_0\right) du \right| \\
&\quad \times \frac{\left(1 - o_{\mathbb{P}}\left(1\right)\right)}{\sup_u f(u,\, \omega_0)} \\
&\geq -O_{\mathbb{P}}\left(\left(m_T^*\right)^{-1/2}\right) \qquad \text{(S.37)} \\
&\quad + \frac{T}{m_T^*} \left| \int_{\left(rm_T^* - m_T^*/2 + n_T/2 + 1\right)/T}^{\lambda_b^0} f\left(u,\, \omega_0\right) du - \int_{\lambda_b^0}^{\left((r+1)m_T^* + n_T/2 + m_{S,T}M_{S,T}/2\right)/T} f\left(u,\, \omega_0\right) du \right| \\
&\quad \times \frac{\left(1 - o_{\mathbb{P}}\left(1\right)\right)}{\sup_u f(u,\, \omega_0)}.
\end{aligned}
$$

Note that $\gamma_{\psi^*}\left(\theta,\, b_T^*\right) \to 0$ follows from

$$\mathbb{P}\left(\mathrm{S}_{\max,T}\left(\omega\right) < 2D^*\sqrt{\log\left(M_T^*\right)/m_T^*}\right) \to 1, \quad \text{for all}\, \omega \in \left[-\pi,\, \pi\right], \qquad \text{under}\, \mathcal{H}_0 \qquad \text{(S.38)}$$

$$\mathbb{P}\left(\mathrm{S}_{\max,T}\left(\omega\right) \geq 2D^*\sqrt{\log\left(M_T^*\right)/m_T^*}\right) \to 1, \quad \text{for some}\, \omega \in \left[-\pi,\, \pi\right], \quad \text{under}\, \mathcal{H}_1^{\mathrm{B}}\, \text{or}\, \mathcal{H}_1^{\mathrm{S}}. \qquad \text{(S.39)}$$

We first show (S.38). Note that

$$2D^*\sqrt{\log\left(M_T^*\right)/m_T^*} \geq 2\sqrt{\log\left(M_T^*\right)/m_T^*} + D\left(m_T^*/T\right)^{\theta}.$$

Under $\mathcal{H}_0$, since $\theta' < \theta$ we have for all $\omega \in [-\pi,\, \pi]$,

$$\mathrm{S}_{\max,T}\left(\omega\right) \leq \max_{1\leq r\leq M_T^*-2} \mathring{f}_{r,T}\left(\omega\right) + D\left(m_T^*/T\right)^{\theta'} + O_{\mathbb{P}}\left(\left(n_T/T\right)^2 + \log\left(n_T\right)/n_T + o\left(b_{W,T}^2\right)\right).$$

Given (4.1), to conclude the proof, we have to show

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

The latter result follows from $\sqrt{\log\left(M_T^*\right)/m_T^*}\leq\sqrt{\log\left(M_T^*\right)/M_{S,T}^*}$ which is implied by Theorem 3.1.

We now prove (S.39) under $\mathcal{H}_1^{\mathrm{B}}$. We have to show that the second term on the right hand side of (S.39) is greater than or equal to $2D^*\sqrt{\log\left(M_T^*\right)/m_T^*}$. The term in question is larger than $b_T^*-2D\left(m_T^*/T\right)^{\theta}$. In view of (4.2) with $\theta'=0$ the result follows.

We now prove (S.39) under $\mathcal{H}_1^{\mathrm{S}}$. For $h\leq 2m_T^*/T$, we have $f\left(\lambda_b^0+h,\,\omega_0\right)\geq f\left(\lambda_b^0,\,\omega_0\right)+b_T^*h^{\theta'}$ or $f\left(\lambda_b^0+h,\,\omega_0\right)\leq f\left(\lambda_b^0,\,\omega_0\right)-b_T^*h^{\theta'}$. Thus,

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

where the second equality follows from (4.2). □

S.B.5 Proofs of the Results of Section 5

From the same reasoning as in the proofs of the results of Section 3, we can replace $\widehat{\sigma}_{L,r}\left(\omega\right)$ by $\sigma_{L,r}\left(\omega\right)$ throughout the proofs of this section.

S.B.5.1 Proof of Proposition 5.1

The following lemma is simple to verify.

Lemma S.B.9. *Let $C\left(u\right)$ and $d\left(u\right)$ be functions on $\left[0,\,\lambda_b^0\right]$ such that $d\left(u\right)$ is increasing. As long as $d\left(\lambda_b^0\right)-d\left(\lambda_b^0-\kappa\right)\geq\sup_{0\leq u\leq\lambda_b^0}\left|C\left(u\right)\right|$ for some $\kappa\in\left[0,\,\lambda_b^0\right]$ we have that,*

$$\underset{0\leq u\leq\lambda_b^0}{\operatorname{argmax}}\left(d\left(u\right)+C\left(u\right)\right)\geq\lambda_b^0-\kappa.\tag{S.40}$$

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

Proof of Proposition 5.1. For $\lambda_b^0\in\left(0,\,1\right)$ 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 than or equal to λ_b^0+1/T. Denote by $\{\widetilde{f}\left(u,\,\omega_0\right)\}_{u\in[0,\,1]}$ 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

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

and $\{d\left(u,\,\omega_0\right)\}_{u\in[0,\,1]}$ is the associated piecewise constant increasing step function. By Lemma S.B.6 it is sufficient to consider

$$\mathrm{D}'_{r,T}(\omega)=M_{S,T}^{-1/2}\left|\sum_{j\in\mathbf{S}_{L,r}}f_{h,T}(j/T,\,\omega)-\sum_{j\in\mathbf{S}_{R,r}}f_{h,T}(j/T,\,\omega)\right|,\qquad\omega\in[-\pi,\,\pi]\,. \tag{S.41}$$

For $r=2m_T,\,2m_T+\dot{m}_T,\,2m_T+2\dot{m}_T\dots$ write

$$\begin{aligned}\sum_{j\in\mathbf{S}_{L,r}}&f_{h,T}(j/T,\,\omega_0)-\sum_{j\in\mathbf{S}_{R,r}}f_{h,T}(j/T,\,\omega_0)\\&=\sum_{j\in\mathbf{S}_{L,r}}\left(f_{h,T}(j/T,\,\omega_0)-\mathbb{E}\left(f_{h,T}(j/T,\,\omega_0)\right)\right)-\sum_{j\in\mathbf{S}_{R,r}}\left(f_{h,T}(j/T,\,\omega_0)-\mathbb{E}\left(f_{h,T}(j/T,\,\omega_0)\right)\right)\\&\quad+\sum_{j\in\mathbf{S}_{L,r}}\left(\mathbb{E}\left(f_{h,T}(j/T,\,\omega_0)\right)-\widetilde{f}(j/T,\,\omega_0)\right)-\sum_{j\in\mathbf{S}_{R,r}}\left(\mathbb{E}\left(f_{h,T}(j/T,\,\omega_0)\right)-f(j/T,\,\omega_0)\right)\\&\quad+\sum_{j\in\mathbf{S}_{L,r}}\widetilde{f}(j/T,\,\omega_0)-\sum_{j\in\mathbf{S}_{R,r}}\widetilde{f}(j/T,\,\omega_0)-\sum_{j\in\mathbf{S}_{R,r}}\left(f(j/T,\,\omega_0)-\widetilde{f}(j/T,\,\omega_0)\right).\end{aligned}$$

For $r=2m_T,\,2m_T+\dot{m}_T,\,2m_T+2\dot{m}_T,\dots,\,\overline{r}_b$ let $C(r/T,\,\omega)=\mathrm{D}'_{r,T}(\omega)$ for $\omega\neq\omega_0$ and

$$\begin{aligned}C(r/T,\,\omega_0)=M_{S,T}^{-1/2}&\left(\sum_{j\in\mathbf{S}_{L,r}}f_{h,T}(j/T,\,\omega_0)-\sum_{j\in\mathbf{S}_{R,r}}f_{h,T}(j/T,\,\omega_0)\right.\\&\left.\quad+\sum_{j\in\mathbf{S}_{R,r},\,j>\overline{r}_b}\left(f(j/T,\,\omega_0)-\widetilde{f}(j/T,\,\omega_0)\right)\right),\end{aligned}$$

for $\omega=\omega_0$. 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 3. That is, we exploit the smoothness of $f(\cdot,\,\cdot)$ under $\mathcal{H}_0$ to yield $\sup_{u\in[0,\,\lambda_b^0]}\sup_{\omega\in[-\pi,\,\pi]}|C(u,\,\omega)|=O_{\mathbb{P}}(\sqrt{\log(T)})$. This combined with the definition of $d(r/T,\,\omega_0)$ implies that for each $r=\overline{r}_b-\lfloor m_T/B\rfloor,\dots,\,\overline{r}_b$, where B is any finite integer with $B>1$,

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

with probability approaching one and

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

By the definition of $d(\cdot,\,\omega_0)$, for $\kappa_T\in[0,\,\dot{m}_T/(BT)]$,

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

In order to apply Lemma S.B.9, we need to choose κ_T such that $\lfloor\kappa_T T\rfloor\,m_{S,T}^{-1}\delta_T M_{S,T}^{-1/2}/\sqrt{\log(T)}\geq 1$ or $\sqrt{M_{S,T}\log(T)}m_{S,T}/(\delta_T T)=o(\kappa_T)$. Lemma S.B.9 then yields

$$\frac{\overline{r}_b}{T}\geq\underset{r=2m_T,\,2m_T+\dot{m}_T,\,2m_T+2\dot{m}_T,\dots;\,r<\overline{r}_b}{\mathrm{argmax}}\max_{\omega\in[-\pi,\,\pi]}T^{-1}\mathrm{D}_{r,T}(\omega)=\underset{r=2m_T,\dots;\,r<\overline{r}_b}{\mathrm{argmax}}T^{-1}\mathrm{D}_{r,T}(\omega_0)\geq\frac{\overline{r}_b}{T}-\kappa_T.$$

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

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

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

S.B.5.2 Proof of Proposition 5.2

Set $\widehat{\mathcal{I}} = \left\{2m_T,\, 2m_T + \dot{m}_T, \ldots,\, 2m_T + \left(\dot{M}_T - 1\right)\dot{m}_T - n_T\right\} \setminus \{2m_T\}$ and $\widehat{\mathcal{T}} = \emptyset$. Under $\mathcal{H}_{1,\mathrm{M}}$, the arguments in the proof of Theorem 3.2 yields,

$$\max_{r\in\widehat{\mathcal{I}}} \max_{\omega\neq\omega_1,\ldots,\,\omega_{m_0}} \mathrm{D}_{r,T}\left(\omega\right) = O_{\mathbb{P}}\left(\sqrt{\log\left(T\right)}\right).$$

Let $r_{L,l},\, r_{R,l} \in \widehat{\mathcal{I}}$ $(l = 1, \ldots, m_0)$ such that $r_{R,l} = r_{L,l} + \dot{m}_T$ and $r_{L,l} \leq T_l^0 < r_{R,l}$. For any ω, we have

$$\max_{r\in\widehat{\mathcal{I}}\setminus\left\{r_{L,1},r_{R,1},\ldots,r_{L,m_0},r_{R,m_0}\right\}} \mathrm{D}_{r,T}\left(\omega\right) = O_{\mathbb{P}}\left(\sqrt{\log\left(T\right)}\right).$$

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

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

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 5.2 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 $T_l^0 \in [r_l,\, r_{l+1})$. As in the proof of Proposition 5.1,

$$\mathrm{D}_{r_l,T}\left(\omega_l\right) = |O_{\mathbb{P}}\left(M_{S,T}^{-1/2}\delta_{l,T}\left(M_{S,T} - \left(r_l - T_l^0\right)\mathbf{1}\left\{r_{l-1} < T_l^0 \leq r_l\right\} + \left(r_{l+1} - T_l^0\right)\mathbf{1}\left\{r_l < T_l^0 < r_{l+1}\right\}\right)\right)|.$$

Note that if $\mathrm{D}_{r_l,T}\left(\omega_l\right) / \left(\delta_{l,T}\sqrt{M_{S,T}}\right) \overset{\mathbb{P}}{\to} 0$ then we must have $\mathrm{D}_{r_{l+1},T}\left(\omega_l\right) = O_{\mathbb{P}}\left(\delta_{l,T}\sqrt{M_{S,T}}\right)$. Using a similar argument as in Lemma S.B.6 one can show that $\mathrm{S}_{\mathrm{Dmax},T}\left(\widehat{\mathcal{I}}\right)$ is asymptotically equivalent to $\max_{r\in\widehat{\mathcal{I}}} \max_{k\in K} \max_{\omega\in[-\pi,\,\pi]} \mathrm{D}_{r_{k,r}^{\diamond},T}\left(\omega\right)$. Thus, in step (2) $\psi(\{X_{t,T}\}, \widehat{\mathcal{I}}) = 1$ because for large enough T,

$$\begin{aligned}
\max_{r\in\widehat{\mathcal{I}}} \max_{k\in K} \max_{\omega\in[-\pi,\,\pi]} \mathrm{D}_{r_{k,r}^{\diamond},T}\left(\omega_0\right) &\geq \max_{r\in\widehat{\mathcal{I}}} \max_{\omega\in[-\pi,\,\pi]} \mathrm{D}_{r,T}\left(\omega\right) \\
&= |\delta_{l,T} O_{\mathbb{P}}\left(\sqrt{M_{S,T}}\right)| \\
&\geq \inf_{1\leq l\leq m_0} |\delta_{l,T} O_{\mathbb{P}}\left(\sqrt{M_{S,T}}\right)| \\
&= 2D^*\left(\log\left(T\right)\right)^{2/3} \\
&> 2D^*\sqrt{\log\left(M_T^*\right)},
\end{aligned}$$

where the last equality follows from Assumption 5.2. We now move to step (3). By the arguments in the proof of Proposition 5.1, there exists $1 \leq l \leq m_0$ such that $|\lambda_l^0 - \widehat{\lambda}_T(\widehat{\mathcal{I}})| \leq \dot{m}_T/T$. Since $\inf_{1 \leq l \leq m_0 - 1} |\lambda_{l+1}^0 - \lambda_l^0| \geq \nu_T^{-1}$ and $m_T/v_T \to 0$ there can exist exactly one l that satisfies $|\lambda_l^0 - \widehat{\lambda}_T(\widehat{\mathcal{I}})| \leq \dot{m}_T/T$. For such a λ_l^0 define $\overline{r}_{l,b} = \lceil T\lambda_l^0 + 1 \rceil$, the smallest integer such that $\overline{r}_{l,b}/T$ is larger than or equal to $\lambda_l^0 + 1/T$. Denote by $\{\widetilde{f}(u, \omega)\}_{u \in [0,1]}$ the path of the spectrum $f(\cdot, \omega)$ without the break $\delta_{l,T}$:

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

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

$$d_l(r/T, \omega_l) = \begin{cases} 0 & \text{if } r + m_T < \overline{r}_{l,b}, \\ (r + m_T - \overline{r}_{l,b}) m_{S,T}^{-1} M_{S,T}^{-1/2} \delta_{l,T} & \text{if } r = \overline{r}_{l,b} - m_T, \overline{r}_{l,b} - m_T + m_{S,T}, \ldots, \overline{r}_{l,b}, \\ M_{S,T}^{1/2} \delta_{l,T} & \text{if } r > \overline{r}_{l,b}, \end{cases}$$

for $\omega = \omega_l$. Let $\{d(u)\}_{u \in [0,1]}$ be the associated piecewise constant increasing step function. For any $r \in \widehat{\mathcal{I}}$, write

$$\begin{aligned} &\sum_{j \in \mathbf{S}_{L,r}} f_{h,T}(j/T, \omega) - \sum_{j \in \mathbf{S}_{R,r}} f_{h,T}(j/T, \omega) \\ &\quad = \sum_{j \in \mathbf{S}_{L,r}} (f_{h,T}(j/T, \omega) - \mathbb{E}(f_{h,T}(j/T, \omega))) - \sum_{j \in \mathbf{S}_{R,r}} (f_{h,T}(j/T, \omega) - \mathbb{E}(f_{h,T}(j/T, \omega))) \\ &\quad\quad + \sum_{j \in \mathbf{S}_{L,r}} \left(\mathbb{E}(f_{h,T}(j/T, \omega)) - \widetilde{f}(j/T, \omega)\right) - \sum_{j \in \mathbf{S}_{R,r}} (\mathbb{E}(f_{h,T}(j/T, \omega)) - f(j/T, \omega)) \\ &\quad\quad + \sum_{j \in \mathbf{S}_{L,r}} \widetilde{f}(j/T, \omega) - \sum_{j \in \mathbf{S}_{R,r}} \widetilde{f}(j/T, \omega) - \sum_{j \in \mathbf{S}_{R,r}} \left(f(j/T, \omega) - \widetilde{f}(j/T, \omega)\right). \end{aligned}$$

For $r = 2m_T, 2m_T + \dot{m}_T, \ldots, \overline{r}_b$, let $C_l(r/T, \omega) = \mathrm{D}'_{r,T}(\omega)$ for $\omega \neq \omega_l$, where $\mathrm{D}'_{r,T}(\omega)$ is given in (S.41) and

$$\begin{aligned} C_l(r/T, \omega_l) = M_{S,T}^{-1/2} \Bigg(&\sum_{j \in \mathbf{S}_{L,r}} f_{h,T}(j/T, \omega_l) - \sum_{j \in \mathbf{S}_{R,r}} f_{h,T}(j/T, \omega_l) \\ &+ \sum_{j \in \mathbf{S}_{R,r}, j > \overline{r}_{l,b}} \left(f(j/T, \omega_l) - \widetilde{f}(j/T, \omega_l)\right) \Bigg), \end{aligned}$$

for $\omega = \omega_l$. We proceed as in the proof of Proposition 5.1. We have

$$d(r/T, \omega_l) \geq \max_{\omega \in \{[-\pi, \pi]/\{\omega_1, \ldots, \omega_{m_0}\}\}} |d(r/T, \omega)| > 0,$$

with probability approaching one. Exploiting the smoothness on $(\lambda_{l-1}^0, \lambda_l^0]$, we have

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

This implies

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

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

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

In order to apply Lemma S.B.9, we need to choose κ_T such that $\lfloor\kappa_T T\rfloor\,\delta_{l,T}M_{S,T}^{-1/2}/m_{S,T}\sqrt{\log(T)}\geq 1$ or $m_{S,T}\sqrt{M_{S,T}\log(T)}/\delta_{l,T}T=o(\kappa_T)$. Lemma S.B.9 then yields

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

The case $r>\overline{r}_{l,b}$ can be treated similarly by symmetry. It results in

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

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

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 $|\widehat{\lambda}_T^{(1)}-\lambda_1^0|\overset{\mathbb{P}}{\to}0$ we need the following event to occur with sufficiently high probability, $\mathbf{W}=\{\text{for } l=1\,\exists r\in\widehat{\mathcal{I}}\text{ and } k=1,\ldots,K\text{ s.t. } r_{r,k}^{\diamond}=T_1^0\}$. Note that

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

Thus,

$$\begin{aligned}1-\mathbb{P}(\mathbf{W}^c)&=1-\frac{\dot{m}_T-1}{\dot{m}_T}\times\frac{\dot{m}_T-2}{\dot{m}_T-1}\times\cdots\times\frac{\dot{m}_T-K}{\dot{m}_T-K+1}\\&=1-\frac{\dot{m}_T-K}{\dot{m}_T}\\&\to 1,\end{aligned}$$

only if $K=O(a_T\dot{m}_T)$ with $a_T\in(0,\,1]$ such that $a_T\to 1$. Note that $K\leq\dot{m}_T$ by construction. The same argument can be repeated for $l=2,\ldots,\,m_0$. □

S.C Sensitivity Analyses

In this section, we conduct Monte Carlo simulations to assess how the finite-sample performance of the test statistics and change-point estimators change when we implement them with different choices for the tuning parameters. Recall that our recommended choices are $m_T = T^{0.66}$, $n_T = T^{5/8}$, $b_{W,T} = n_T^{-1/6}$ and $n_\omega = 7$. Since our choice for m_T and n_T corresponds to the upper bound allowed by Assumption 3.6, here we consider smaller values of m_T and n_T. We consider Model M1 and $T = 250$, 500 and 1000. Table S.1 shows that the null rejection rates are not much affected by the change in the choice of the tuning parameters. The null rejection rates become less accurate only for substantially smaller values of m_T, n_T and $b_{W,T}$. For example, for $m_T = T^{0.58}$ and $n_T = T^{0.56}$ or for $b_{W,T} = n_T^{-0.25}$, some of the tests show some over-rejection. The choice of the number of frequencies n_ω and of their locations do not matter much for the finite-sample performance of S_{Dmax} and R_{Dmax}. Table S.2 shows the results about the power. Any tuning parameter choice results in good monotonic power for all tests. Overall, the results suggest that reasonable changes in the choice of the tuning parameters yield little changes in the finite-sample performance of the tests. The change in the results become larger as the smoothing bandwidths m_T, n_T and $b_{W,T}$ are set too small.

We move to the results about the change-point estimator. We consider Model M6 with $T = 1000$. Table S.3 shows that small changes in the tuning parameters result in little changes in the precision of the change-point estimator and of the estimator of the number of change-points.

Table S.1: Empirical small-sample size for model M1

	$m_T = T^{0.63}$, $n_T = T^{0.6}$			$m_T = T^{0.6}$, $n_T = T^{0.58}$			$m_T = T^{0.58}$, $n_T = T^{0.56}$		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$
$\text{S}_{\max,T}(0)$	0.079	0.061	0.056	0.059	0.054	0.059	0.109	0.088	0.076
$\text{S}_{\text{Dmax},T}$	0.045	0.067	0.063	0.058	0.070	0.058	0.132	0.069	0.062
$\text{R}_{\max,T}(0)$	0.088	0.085	0.077	0.109	0.096	0.082	0.169	0114	0.102
$\text{R}_{\text{Dmax},T}$	0.039	0.050	0.034	0.093	0.086	0.072	0.101	0.092	0.074
	$b_{W,T} = n_T^{-0.15}$			$b_{W,T} = n_T^{-0.2}$			$b_{W,T} = n_T^{-0.25}$		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$
$\text{S}_{\max,T}(0)$	0.025	0.038	0.037	0.072	0.068	0.063	0.084	0.103	0.092
$\text{S}_{\text{Dmax},T}$	0.032	0.041	0.044	0.047	0.061	0.057	0.021	0.067	0.096
$\text{R}_{\max,T}(0)$	0.032	0.031	0.045	0.063	0.061	0.058	0.094	0.089	0.084
$\text{R}_{\text{Dmax},T}$	0.029	0.031	0.032	0.027	0.038	0.042	0.091	0.085	0.082
	$n_\omega = 15$			$n_\omega = 11$			$n_\omega = 5$		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$
$\text{S}_{\text{Dmax},T}$	0.011	0.037	0.043	0.016	0.037	0.039	0.023	0.070	0.064
$\text{R}_{\text{Dmax},T}$	0.022	0.024	0.030	0.024	0.027	0.033	0.026	0.032	0.035

Table S.2: Empirical small-sample power for model M1

	$m_T = T^{0.63}$, $n_T = T^{0.6}$			$m_T = T^{0.6}$, $n_T = T^{0.58}$			$m_T = T^{0.58}$, $n_T = T^{0.56}$		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}\,(0)$	0.686	0.791	0.913	0.605	0.748	0.886	0.682	0.783	0.892
$\mathrm{S}_{\mathrm{Dmax},T}$	0.771	0.846	0.936	0.747	0.828	0.936	0.822	0.914	0.942
$\mathrm{R}_{\max,T}\,(0)$	0.813	0.899	0.956	0.821	0.915	0.942	0.891	0.877	0.932
$\mathrm{R}_{\mathrm{Dmax},T}$	0.582	0.745	0.871	0.649	0.827	0.864	0.784	0.791	0.863
	$b_{W,T} = n_T^{-0.15}$			$b_{W,T} = n_T^{-0.2}$			$b_{W,T} = n_T^{-0.25}$		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}\,(0)$	0.607	0.815	0.886	0.786	0.921	0.996	0.779	0.926	0.969
$\mathrm{S}_{\mathrm{Dmax},T}$	0.628	0.899	0.907	0.754	0.921	0.995	0.759	0.915	0.953
$\mathrm{R}_{\max,T}\,(0)$	0.771	0.902	0.940	0.821	0.952	0.996	0.926	0.982	0.996
$\mathrm{R}_{\mathrm{Dmax},T}$	0.458	0.641	0.704	0.604	0.821	0.952	0.852	0.928	0.952
	$n_\omega = 15$			$n_\omega = 11$			$n_\omega = 5$		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\mathrm{Dmax},T}$	0.769	0.977	0.907	0.731	0.966	0.962	0.645	0.756	0.864
$\mathrm{R}_{\mathrm{Dmax},T}$	0.401	0.635	0.708	0.455	0.634	0.701	0.742	0.756	0.834

Table S.3: Empirical distribution of $\widehat{m} - m_0$ for model M6

$m_T = T^{0.63}$, $n_T = T^{0.6}$					$m_T = T^{0.60}$, $n_T = T^{0.58}$				
% time $\widehat{m} = m_0$		$Q_{0.25}$	Median	$Q_{0.75}$	% time $\widehat{m} = m_0$		$Q_{0.25}$	Median	$Q_{0.75}$
80.42	$\widehat{T}_1$	298	329	348	78.60	$\widehat{T}_1$	302	328	348
	$\widehat{T}_2$	605	651	682		$\widehat{T}_2$	613	646	675
$n_\omega = 15$					$n_\omega = 11$				
% time $\widehat{m} = m_0$		$Q_{0.25}$	Median	$Q_{0.75}$	% time $\widehat{m} = m_0$		$Q_{0.25}$	Median	$Q_{0.75}$
84.56	$\widehat{T}_1$	300	333	355	84.78		300	333	354
	$\widehat{T}_2$	629	663	691			627	662	690
$n_\omega = 5$									
% time $\widehat{m} = m_0$		$Q_{0.25}$	Median	$Q_{0.75}$					
81.85	$\widehat{T}_1$	299	333	353					
	$\widehat{T}_2$	618	657	696					

S.D Additional Monte Carlo Results

In this section, we report simulations results for Model M6 where the errors are drawn from the t_ν distribution. In particular, in Model M6 $e_t \sim$ i.i.d. t_ν with $\nu = 5, 10$. Table S.4-S.5 show that the proposed test statistics have accurate null rejection rates and good monotonic power similar to the case of Gaussian errors. Note that the statistic $\widehat{D}$ continues to suffer from large size distortions.

Table S.4: Empirical small-sample size for model M1 with t-distributed errors

	$t_\nu,\ \nu = 5$			$t_\nu,\ \nu = 10$		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}(0)$	0.065	0.062	0.058	0.039	0.054	0.035
$\mathrm{S}_{\mathrm{Dmax},T}$	0.054	0.066	0.053	0.028	0.063	0.051
$\mathrm{R}_{\max,T}(0)$	0.071	0.056	0.049	0.038	0.065	0.039
$\mathrm{R}_{\mathrm{Dmax},T}$	0.043	0.014	0.007	0.008	0.026	0.011
$\widehat{D}$ statistic	0.661	0.515	0.083	0.598	0.454	0.051

Table S.5: Empirical small-sample power for model M6 with t-distributed errors

	$t_\nu,\ \nu = 5$			$t_\nu,\ \nu = 10$		
$\alpha = 0.05$	$T = 250$	$T = 500$	$T = 1000$	$T = 250$	$T = 500$	$T = 1000$
$\mathrm{S}_{\max,T}(0)$	0.777	0.896	0.900	0.738	0.875	0.915
$\mathrm{S}_{\mathrm{Dmax},T}$	0.788	0.894	0.908	0.717	0.859	0.890
$\mathrm{R}_{\max,T}(0)$	0.782	0.928	0.934	0.734	0.920	0.967
$\mathrm{R}_{\mathrm{Dmax},T}$	0.582	0.803	0.800	0.522	0.796	0.895
$\widehat{D}$ statistic	0.989	0.996	0.992	0.982	0.979	0.856

References

Berkes, I., Philipp, W., 1979. Approximation theorems for independent and weakly dependent random vectors. Annals of Probability 7, 29–54.

Bibinger, M., Jirak, M., Vetter, M., 2017. Nonparametric change-point analysis of volatility. Annals of Statistics 45, 1542–1578.

Bobkov, S.G., Chistyakov, G.P., Götze, F., 2013. Rate of convergence and edgeworth-type expansion in the entropic central limit theorem. Annals of Probability 41, 2479–2512.

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

Brown, L.D., Low, M.G., 1996. Asymptotic equivalence of nonparametric regression and white noise. Annals of Statistics 24, 2384–2398.

Dahlhaus, R., 1997. Fitting time series models to nonstationary processes. Annals of Statistics 25, 1–37.

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

Wu, W.B., 2007. Strong invariance principles for dependent random variables. Annals of Probability 35, 2294–2320.

Wu, W.B., Shao, Z., 2004. Limit theorems for iterated random functions. Journal of Applied Probability 41, 425–436.

Wu, W.B., Zhao, Z., 2007. Inference of trends in time series. Journal of the Royal Statistical Society. Series B 69, 391–410.

Wu, W.B., Zhou, Z., 2011. Gaussian approximation for non-stationary multiple time series. Statistica Sinica 21, 1397–1413.