Tail Spectral Density Estimation and Its Uncertainty Quantification: Another Look at Tail Dependent Time Series Analysis

By Ting Zhang and Beibei Xu

University of Georgia

October 23, 2022

Abstract

We consider the estimation and uncertainty quantification of the tail spectral density, which provide a foundation for tail spectral analysis of tail dependent time series. The tail spectral density has a particular focus on serial dependence in the tail, and can reveal dependence information that is otherwise not discoverable by the traditional spectral analysis. Understanding the convergence rate of tail spectral density estimators and finding rigorous ways to quantify their statistical uncertainty, however, still stand as a somewhat open problem. The current article aims to fill this gap by providing a novel asymptotic theory on quadratic forms of tail statistics in the double asymptotic setting, based on which we develop the consistency and the long desired central limit theorem for tail spectral density estimators. The results are then used to devise a clean and effective method for constructing confidence intervals to gauge the statistical uncertainty of tail spectral density estimators, and it can be turned into a visualization tool to aid practitioners in examining the tail dependence for their data of interest. Numerical examples including data applications are presented to illustrate the developed results.

Keywords: central limit theorem, quadratic forms, tail adversarial stability, tail dependent time series, tail spectral density.

1 Introduction

A fundamental problem in spectral analysis of time series is to estimate the spectral density function that relates to the discrete-time Fourier transform of autocovariances at different lags. Spectral density estimators and their asymptotic theory have been extensively studied in the literature; see for example Bentkus and Rudzkis (1982), Rosenblatt (1984), Dahlhaus (1985), Velasco and Robinson (2001), Phillips et al. (2006, 2007), Wu and Shao (2007), Liu and Wu (2010), Xiao and Wu (2011), and references therein. We shall here also refer to the books by Grenander and Rosenblatt (1957), Hannan (1970), Anderson (1971), Brillinger (1975), Priestley (1981), Rosenblatt (1985), and Brockwell and Davis (1991) for additional discussions and references. When evaluated at frequency zero, the spectral density becomes the so called long-run variance, which is an important quantity for mean inference of time series data and its estimation has been widely studied in the literature; see for example Newey and West (1987), Andrews (1991), Song and Schmeiser (1995), Lahiri (2003), Wu (2009), Flegal and Jones (2010), Politis (2011), Zhang (2018), and references therein. In the aforementioned works, the batch means method in Song and Schmeiser (1995) is related to the technique of subsampling of Politis and Romano (1994) and Politis et al. (1999), and an alternative to Flegal and Jones (2010) was given by Giakoumatos et al. (1999). Since the conventional spectral density is constructed using the traditional autocovariances, it mainly concerns the dependence in terms of comovements with respect to the mean and therefore may not be suitable for studying tail dependence.

Tail dependence, or extremal dependence, refers to the dependence in the joint extremes of the underlying distribution. The phenomenon has been extensively studied in bivariate or finite-dimensional multivariate distributions; see for example Sibuya (1960), de Haan and Resnick (1977), Joe (1993), Ledford and Tawn (1996), Coles et al. (1999), Embrechts et al. (2002), Draisma et al. (2004), Poon et al. (2004), McNeil et al. (2005), Zhang (2008), Balla et al. (2014), Hoga (2018), and references therein. Zhang (2005) considered a lag-k tail dependence index that summarizes the degree of tail dependence at different lags for a time series; see also the quantilogram of Linton and Whang (2007), the extremogram of Davis and Mikosch (2009) and Hill (2009), and the tail autocorrelation of Zhang (2022) for additional results along this direction. Unlike the traditional autocorrelations, the asymptotic behavior of their tail counterparts can be much more complicated due to the double asymptotic scheme necessary for capturing the tail phenomena. To be more specific, sample tail autocorrelations no longer share a universal convergence rate as traditional autocorrelations otherwise do, and their asymptotic behavior can transit from one phase to the other once the lag index passes the point beyond which tail dependence vanishes; see for example Zhang (2022). This can pose a challenge on studying their infinite sum, a quantity that appears in the central limit theorem of high quantile regression estimators (Zhang, 2021a), or more generally their Fourier transforms that lead to the tail spectral density.

Davis and Mikosch (2009) considered estimating the tail spectral density using the truncated periodogram, and proved its estimation consistency. The associated convergence rate and asymptotic distribution, however, were still unknown at the time and left as an open problem; see the discussion in their Section 5. Mikosch and Zhao (2014) obtained an asymptotic theory on periodogram ordinates at fixed frequencies, which was then used to show that the smoothed periodogram provides a consistent estimator of the tail spectral density. However, their Theorem 5.1 only concerned the estimation consistency and did not provide any information about the convergence rate or the asymptotic distribution of the smoothed periodogram. Mikosch and Zhao (2015) studied the integrated periodogram, but did not cover triangular array weight functions with shrinking support that are needed for consistent estimation of the spectral tail density. To the best of our knowledge, the problem of developing a more comprehensive asymptotic theory on tail spectral density estimators to understand their convergence rate and asymptotic distribution for a general class of tail dependent time series has not been well addressed. The current article aims at filling this gap, for which there are three major difficulties. First, unlike the periodogram ordinates in Mikosch and Zhao (2014) or the leading diagonal term in Mikosch and Zhao (2015) that can be handled by linear forms, the study of tail spectral density estimators beyond their consistency typically requires an asymptotic theory on non-degenerate quadratic forms of tail statistics which has not been well explored and can be more challenging to deal with. Second, compared with traditional spectral density estimators in which sample autocovariances share a universal convergence rate, sample tail autocorrelations can exhibit a two-phase asymptotic behavior with dichotomous convergence rates at different lags (Zhang, 2022), which makes it difficult to analyze their infinite sums or Fourier transforms in the tail setting that relate to the tail spectral density. Third, existing results in this direction were mostly developed under the strong mixing framework of Rosenblatt (1956), which typically had to be used together with additional anti-clustering conditions and regularly varying conditions to handle tail events. As a result, it can often lead to non-trivial conditions that involve the interplay between how fast the strong mixing coefficient decays and how extremal the tail can be, and can result in strong conditions for common extreme value time series models described in a recent review by Zhang (2021b).

We in the current article provide a novel asymptotic theory on quadratic forms of tail statistics in the double asymptotic setting, based on which we develop the consistency and the long desired central limit theorem for tail spectral density estimators. The results are then used to devise a clean and effective method for constructing confidence intervals to gauge the statistical uncertainty of tail spectral density estimators, and it can be turned into a visualization tool to aid practitioners in examining the tail dependence for their data of interest. It can be seen from our data applications that the tail spectral analysis as considered in the current article can be useful in revealing dependence information that is otherwise not discoverable by the traditional spectral analysis. Results developed in the current article provide a foundation for tail spectral analysis and its uncertainty quantification for a general class of tail dependent time series, and shed new light on the difficult problem of handling quadratic forms of tail statistics in the double asymptotic scheme.

The remaining of the article is organized as follows. Section 2 introduces our tail adversarial stability framework, and provides examples to illustrate its difference from the strong mixing framework. Section 3 contains our main results, with proofs provided in the Supplementary Material. Section 4 contains a simulation study and real data applications to illustrate the developed results. Section 5 concludes the article.

2 Framework

2.1 Tail Adversarial Stability

Suppose we observe a stationary time series $X_1, \ldots, X_n$ with marginal distribution function $F(x) = \mathrm{pr}(X_1 \leq x)$, $x \in \mathbb{R}$, according to

$$X_i = G(\mathcal{F}_i), \quad \mathcal{F}_i = (\ldots, \epsilon_{i-1}, \epsilon_i), \tag{1}$$

where ϵ_j, $j \in \mathbb{Z}$, are independent and identically distributed (i.i.d.) innovations, and G is a measurable function such that X_i is properly defined. The causal representation (1) is very general and covers a huge class of stationary processes; see for example the discussions in Wiener (1958), Tong (1990), Wu (2005), and Zhang (2021a). An example of a certain stationary process with trivial backward tail field that is not covered by (1) was given by Rosenblatt (2009). Let $F^{-1}(u) = \inf\{x : \ F(x) \geq u\}$, then $\mathcal{U}_F = \lim_{u\uparrow 1} F^{-1}(u)$ represents the upper end point of the distribution and can take the value of infinity if the distribution is not bounded. Then as x approaches $\mathcal{U}_F$, data points exceeding x can be viewed as tail events, and Zhang (2005) considered the lag-k tail dependence index

$$\nu_k = \lim_{x\uparrow\mathcal{U}_F} \nu_{k,x}, \quad \nu_{k,x} = \mathrm{pr}(X_{1+k} > x \mid X_1 > x), \tag{2}$$

which naturally extends the foundational metric of Sibuya (1960) to the time series setting. Although the quantities in (2) provide a straightforward summary of the underlying serial tail dependence, similar to the traditional autocorrelations, they are typically not directly useful for developing an asymptotic theory for tail dependent time series. We shall here adopt the tail adversarial stability framework of Zhang (2021a), which has been proven to be useful in obtaining limit theorems of tail dependent processes under interpretable conditions. Existing results on tail adversarial stability mainly concern limit theorems of linear forms, and we shall here focus on the more challenging topic of obtaining limit theorems for quadratic forms of tail statistics.

Let $\epsilon_0^\star$ be an innovation that has the same distribution as ϵ_0 but independent of $(\epsilon_k)_{k\in\mathbb{Z}}$, then $\mathcal{F}_i^\star = (\mathcal{F}_{-1}, \epsilon_0^\star, \epsilon_1, \ldots, \epsilon_i)$ is the coupled shift process and $X_i^\star = G(\mathcal{F}_i^\star)$ represents the coupled output at time i when the innovation at time zero is replaced by its i.i.d. copy. We consider

$$\theta_x(i) = \sup_{z\geq x} \mathrm{pr}(X_i^\star \leq z \mid X_i > z),$$

which quantifies the degree to which the innovation at time zero affects whether the output data at time i is a tail observation. Note that if X_i does not depend on ϵ_0, then $X_i^\star = X_i$

and $\theta_x(i) = 0$, meaning that ϵ_0 will not have any tail adversarial effect on X_i. Let

$$\Theta_{x,q}(m) = \sum_{i=m}^{\infty} \{\theta_x(i)\}^{1/q}, \quad m \geq 0,$$

which measures the cumulative tail adversarial effect of ϵ_0 on future observations with gap $i \geq m$, then we say that the process (X_i) is tail adversarial q-stable or $(X_i) \in \mathrm{TAS}_q$ if

$$\lim_{x \uparrow \mathcal{U}_F} \Theta_{x,q}(0) < \infty;$$

and geometrically tail adversarial stable or $(X_i) \in \mathrm{GTAS}$ if there exist constants $c^\star \in (0, \infty)$ and $\phi \in (0, 1)$ such that

$$\theta_x(i) \leq c^\star \phi^i, \quad i \geq 0,$$

holds for some x that is close enough to $\mathcal{U}_F$. Additional discussions on the tail adversarial stability coefficient and its calculation can be found in Zhang (2021a). A geometric moment contraction condition was adopted by Shao and Wu (2007) for studying asymptotic theory of conventional spectral density estimators, and we shall here focus on the tail setting.

2.2 Examples: Illustration and Comparison

We shall here illustrate our tail adversarial stability framework and make a comparison with conditions under the strong mixing framework. We first consider the autoregressive moving average (ARMA) model with regularly varying innovations as considered in Mikosch and Zhao (2014).

Example 2.1. *Let (ϵ_i) be a sequence of i.i.d. regularly varying random variables with index $\gamma > 0$ and*

$$X_i = \sum_{l=0}^{\infty} a_l \epsilon_{i-l},$$

where the coefficients (a_l) are chosen from the ARMA equation $\sum_{l=0}^{\infty} a_l z^l = \beta(z)/\psi(z)$, $z \in \mathbb{C}$, with $\beta(z) = 1 + \beta_1 z + \cdots + \beta_r z^r$ and $\psi(z) = 1 - \psi_1 z - \cdots - \psi_s z^s$ for some $r, s \geq 0$.

Following Mikosch and Zhao (2014), we assume that $\beta(z)$ and $\psi(z)$ do not have common zeros and $\psi(z) \neq 0$ for $|z| \leq 1$, then by the argument in the aforementioned paper the process (X_i) is also regularly varying with the same index $\gamma > 0$ and we can write $\mathrm{pr}(X_i > x) = x^{-\gamma}\ell(x)$ *for some slowly varying function $\ell(\cdot)$. Under certain regularity conditions on the innovation density, Mikosch and Zhao (2014) showed that the conditions needed for their Theorem 5.1 to guarantee the estimation consistency of the smoothed periodogram are satisfied for the ARMA model. Their derivation largely benefited from the fact that the strong mixing coefficients decay geometrically fast under the ARMA structure, which simplified the verification of some of their key conditions.*

For the tail adversarial stability framework, by Theorem 3.9 of Bai and Zhang (2022) we can show that under certain regularity conditions the tail adversarial stability measure satisfies

$$\theta_x(i) \leq c|a_i|^{\gamma'}$$

for sufficiently large x, where $\gamma' < \gamma$ can be chosen arbitrarily close to γ when $\gamma \leq 1$ and $\gamma' = 1$ if $\gamma > 1$. Under the ARMA structure, the coefficients a_i, $i \geq 0$, decay geometrically fast, and thus $(X_i) \in \mathrm{TAS}_q$ for any $q > 0$ and in addition $(X_i) \in \mathrm{GTAS}$. Note that the consistency result in our Theorem 3.1 only requires that $(X_i) \in \mathrm{TAS}_q$ for some $q \geq 4$ and $n\mathrm{pr}(X_i > x_n) \to \infty$. *In comparison, the consistency result in Theorem 5.1 of Mikosch and Zhao (2014) requires that* $n^{1/3}\mathrm{pr}(X_i > x_n) \to \infty$ *which can be much stronger; see their condition (M1). We also remark that the conditions employed by Mikosch and Zhao (2014) in general involve nontrivial interplays between how fast the strong mixing coefficient decays to zero and how extremal the tail can be and thus can be difficult to be made explicit for processes without a geometric decay.*

We shall then consider the moving-maximum process of Hall et al. (2002) that has been widely adopted for extreme value time series modeling. As commented in a recent review paper by Zhang (2021b), the additive structure in traditional time series models cannot describe the extremal clusters and tail dependence satisfactorily in many applications, and the moving-maximum model and its variants and generalizations have been recognized as powerful alternatives for extreme value time series modeling; see also the references therein for additional discussions.

Example 2.2. *Let* (ϵ_i) *be a sequence of independent Fréchet random variables with distribution function* $\mathrm{pr}(\epsilon_i \le z) = \exp(-z^{-\gamma})$ *for some* $\gamma > 0$*, we consider the moving-maximum process*

$$X_i = \max_{0 \le l < \infty} a_l \epsilon_{i-l},$$

which is well defined if the nonnegative coefficients satisfy $\sum_{l=0}^{\infty} a_l^{\gamma} < \infty$*; see the discussion in Hall et al. (2002). The aforementioned paper also showed that the moving-maximum process is dense in the class of stationary processes whose finite-dimensional distributions are extreme-value of a given type; see also Zhang and Smith (2004), Zhang (2005) and Zhang et al. (2017) for additional discussions. Mikosch and Zhao (2014) verified conditions for their Theorem 5.1 regarding the estimation consistency of the smoothed periodogram when the coefficients* (a_i) *follow a geometric decay.*

As a comparison, our consistency result in Theorem 3.1 only requires that $(X_i) \in \mathrm{TAS}_q$ *for some* $q \ge 4$*, for which by the elementary calculation as in Zhang (2021a) a sufficient condition is given by*

$$\sum_{l=0}^{\infty} a_l^{\gamma/q} < \infty.$$

Compared with the existence condition $\sum_{l=0}^{\infty} a_l^{\gamma} < \infty$ *under which the moving-maximum process is well defined, our tail adversarial stability condition seems to be reasonably mild. Although Mikosch and Zhao (2014) only considered the case when* (a_i) *decay geometrically* *fast, we shall here provide additional insights into the conditions of their Theorem 5.1 by considering the case when the coefficients* (a_i) *only follow an algebraic decay. For this, we consider the simple setting when* $\gamma = 1$ *and* $a_i = ci^{-\tau}$ *for some* $0 < c < \infty$ *and* $\tau > 1$*, then by Corollary 2.2 of Dombry and Eyi-Minko (2012) and Karamata's theorem, the conditions in Theorem 5.1 of Mikosch and Zhao (2014) requires at least* $\tau > 10$*. Note that this is only a necessary condition, as it is very complicated to work out exactly what a sufficient condition on* τ *is due to the nontrivial interplay with how extremal the tail can be under their framework. In contrast, for the current tail adversarial stability framework, it suffices to require that* $\tau > 4$ *to guarantee the estimation consistency; see Theorem 3.1. In addition, our results allow more extremal tails in the sense that our Theorem 3.1 only requires* $n\mathrm{pr}(X_i > x_n) \to \infty$ *while Theorem 5.1 of Mikosch and Zhao (2014) requires that* $n^{1/3}\mathrm{pr}(X_i > x_n) \to \infty$*.*

Therefore, the tail adversarial stability framework in Section 2.1 seems to provide an alternative mathematical foundation for asymptotic theory of tail dependent time series, which can possibly lead to more explicit but weaker conditions than the strong mixing framework. In addition, it does not require the use of additional anti-clustering conditions such as the one in (2.4) of Mikosch and Zhao (2014), and can be used to guide the development of asymptotic theory for processes that are not necessarily regularly varying.

3 Main Results

Let $x_n \to \mathcal{U}_F$, we consider the tail autocorrelation

$$\rho_{k,n} = \frac{\mathrm{pr}(X_{1+k} > x_n \mid X_1 > x_n) - \mathrm{pr}(X_{1+k} > x_n)}{1 - \mathrm{pr}(X_1 > x_n)}, \tag{3}$$

which can be viewed as a standardized pre-asymptotic version of the lag-k tail dependence index of Zhang (2005). Zhang (2022) provided a two-phase asymptotic theory on sample tail autocorrelations and used it to guide the construction of a visualization tool with lines of significance. Let $\iota = \sqrt{-1}$ be the imaginary unit, we shall here focus on the frequency domain and consider the tail spectral density

$$f_n(\lambda) = \frac{1}{2\pi} \sum_{|k|<n} \rho_{k,n} e^{\iota k \lambda}, \tag{4}$$

which naturally extends the conventional spectral density to the tail setting using the tail autocorrelations (3). Let $\mathrm{I}(\cdot)$ be the indicator function, $\bar{F}(x) = 1 - F(x)$ and $\hat{\bar{F}}_n(x) = n^{-1} \sum_{i=1}^{n} \mathrm{I}(X_i > x)$, then a sample version of (3) is given by

$$\hat{\rho}_{k,n} = \frac{\hat{\mu}_{k,n}}{\hat{\mu}_{0,n}}, \quad \hat{\mu}_{k,n} = \frac{1}{n} \sum_{i=1}^{n-|k|} \{\mathrm{I}(X_i > x_n) - \hat{\bar{F}}_n(x_n)\}\{\mathrm{I}(X_{i+|k|} > x_n) - \hat{\bar{F}}_n(x_n)\}.$$

To estimate the tail spectral density (4), we consider the lag-window estimator

$$\hat{f}_n(\lambda) = \frac{1}{2\pi} \sum_{|k|<n} \hat{\rho}_{k,n} e^{\iota k \lambda} K\left(\frac{k}{B_n}\right), \tag{5}$$

where $K : \mathbb{R} \to \mathbb{R}$ is a kernel function and $B_n \to \infty$ is a positive bandwidth sequence. In the non-tail setting, lag-window estimators for the conventional spectral density function have been extensively studied in the literature; see for example Anderson (1971), Rosenblatt (1984), Phillips et al. (2006), Shao and Wu (2007), Liu and Wu (2010), and references therein. Although developing a central limit theorem for the conventional spectral density estimator is already a highly nontrivial problem as commented by Liu and Wu (2010), achieving it in the current tail setting with the double asymptotic scheme can be even more challenging. Zhang (2022) provided an asymptotic theory for sample tail autocorrelations, whose unusual two-phase asymptotic behavior distinguishes them from traditional autocorrelations. We shall here provide an asymptotic theory on their infinite sums that lead to the tail spectral density estimator (5). Unlike Zhang (2022) which only focused on individual sample autocorrelations at a fixed lag, the current problem requires the handling of a growing number of sample tail autocorrelations at the same time, which is more challenging and requires a new asymptotic theory on quadratic forms of tail statistics.

Throughout the article, we assume that the kernel function $K \in \mathcal{K}$, the class of symmetric nonnegative functions with $\bar{\kappa} = \sup_{u\in\mathbb{R}} K(u) < \infty$ and $\kappa = \int_{\mathbb{R}} K^2(u)du < \infty$. We first consider a proxy of (5) given by

$$\tilde{f}_n(\lambda) = \frac{1}{2\pi} \sum_{|k|<n} \tilde{\rho}_{k,n} e^{\iota k\lambda} K\left(\frac{k}{B_n}\right),$$

where

$$\tilde{\rho}_{k,n} = \frac{1}{nF(x_n)\bar{F}(x_n)} \sum_{i=1}^{n-|k|} \{\mathrm{I}(X_i > x_n) - \bar{F}(x_n)\}\{\mathrm{I}(X_{i+|k|} > x_n) - \bar{F}(x_n)\}.$$

Theorem 3.1 provides the consistency of $\tilde{f}_n(\lambda)$ and quantifies its distance from $\hat{f}_n(\lambda)$.

Theorem 3.1. *Assume that* $(X_i) \in \mathrm{TAS}_4$ *and* $K \in \mathcal{K}$. *If* $F(x_n) \to 1$, $B_n \to \infty$ *and* $\{n\bar{F}(x_n)\}^{-1}B_n \to 0$, *then for any* $\lambda \in [0, 2\pi)$,

$$\tilde{f}_n(\lambda) - E\{\tilde{f}_n(\lambda)\} \to_p 0,$$

and

$$\hat{f}_n(\lambda) = \{\tilde{f}_n(\lambda) + O_p(n^{-1}B_n)\}(1 + O_p[\{n\bar{F}(x_n)\}^{-1/2}]).$$

By Theorem 3.1, the tail spectral density estimator (5) also satisfies

$$\hat{f}_n(\lambda) - E\{\tilde{f}_n(\lambda)\} \to_p 0,$$

where the asymptotic center

$$E\{\tilde{f}_n(\lambda)\} = f_n(\lambda) + \frac{1}{2\pi} \sum_{|k|<n} \rho_{k,n} e^{\iota k\lambda} \left\{\left(1 - \frac{|k|}{n}\right) K\left(\frac{k}{B_n}\right) - 1\right\}.$$

Therefore, if we choose a kernel function that satisfies $K(u) = 1$ when $|u| \le c_K$ for some $0 < c_K < \infty$, which is referred to as the class of flat-top kernels (Politis, 2011), then

$$|E\{\tilde{f}_n(\lambda)\} - f_n(\lambda)| \le \frac{1}{\pi} \sum_{k=1}^{\lfloor c_K B_n \rfloor} \frac{|k\rho_{k,n}|}{n} + \frac{1}{\pi} \sum_{k=\lfloor c_K B_n \rfloor + 1}^{n-1} (\bar{\kappa} + 1)|\rho_{k,n}| \to 0$$

by the dominated convergence theorem and the proof of Lemma 3 in Zhang (2021a). This leads to the consistency of the tail spectral density estimator (5). As discussed in Section 2.2, our Theorem 3.1 complements the consistency results in Davis and Mikosch (2009) and Mikosch and Zhao (2014) in the sense that our tail adversarial stability framework can possibly lead to weaker conditions on the degree of tail dependence and allow more extremal tails.

We shall now provide a central limit theorem for the tail spectral density estimator (5), which has not been well addressed in the literature. In the following, we use $\mathcal{K}_{FT}$ to denote the class of flat-top kernel functions in $\mathcal{K}$ that have bounded support and are Lipschitz continuous except for a finite number of points. Shao and Wu (2007) provided a central limit theorem for the conventional spectral density estimator under a geometric moment contraction condition, and their proof was based on a big-block-small-block argument, which splits the summation into alternating big and small blocks to take advantage of the asymptotic independence between well separated blocks. We shall here focus on the tail setting under the double asymptotic scheme where the tail level $x_n \to \mathcal{U}_F = \lim_{u \uparrow 1} F^{-1}(u)$ is allowed to approach the end point of the distribution as $n \to \infty$, and the proof is based on using quadratic forms of m-dependent tail martingale differences to approximate the tail spectral density estimator.

Theorem 3.2. *Assume that* $(X_i) \in \mathrm{GTAS}$ *and* $K \in \mathcal{K}$ *has bounded support and is Lipschitz continuous except for a finite number of points. If* $F(x_n) \to 1$, $B_n \bar{F}(x_n)(\log n)^{-7} \to \infty$ *and* $\{n\bar{F}(x_n)\}^{-1} B_n (\log n)^8 \to 0$, *then (i) for* $\lambda \in [0, 2\pi) \setminus \{0, \pi\}$ *where* $f_n(\lambda)$ *is bounded away* *from zero for all large n,*

$$\{\kappa^{1/2} f_n(\lambda)\}^{-1}(B_n^{-1}n)^{1/2}[\tilde{f}_n(\lambda) - E\{\tilde{f}_n(\lambda)\}] \to_d N(0,1);$$

and (ii) for $\lambda \in \{0, \pi\}$ where $f_n(\lambda)$ is bounded away from zero for all large n,

$$\{\kappa^{1/2} f_n(\lambda)\}^{-1}(B_n^{-1}n)^{1/2}[\tilde{f}_n(\lambda) - E\{\tilde{f}_n(\lambda)\}] \to_d N(0,2).$$

Theorem 3.2 discovers a somewhat surprising and unexpected limiting behavior of tail spectral density estimators, and we shall here provide a brief discussion. In particular, unlike traditional autocorrelations that share the universal $n^{1/2}$-convergence rate, the convergence rate of sample tail autocorrelations can be affected by the tail setting and slowed to $\{n\bar{F}(x_n)\}^{1/2}$; see for example Davis and Mikosch (2009), Mikosch and Zhao (2014) and Zhang (2022). As a result, the rate of convergence for tail spectral density estimators in the form of their infinite sums was conjectured to be $\{n\bar{F}(x_n)/B_n\}^{1/2}$, where the additional B_n term is due to the use of kernel tapering. However, Theorem 3.2 discovers that the convergence rate of tail spectral density estimators can actually attain $(n/B_n)^{1/2}$, which is faster than the conjectured rate. This seems a bit surprising, but is intuitively related to the dichotomous asymptotic behavior of sample tail autocorrelations whose convergence rate can transit from $\{n\bar{F}(x_n)\}^{1/2}$ to $n^{1/2}$ for different lags; see for example Zhang (2022). The aforementioned paper considered individual sample tail autocorrelations that can be handled by linear forms, and the results are not directly useful for the current problem but offer a possible intuitive explanation from a different perspective. Corollary 3.1 states that the central limit theorems in Theorem 3.2 will continue to hold for $\hat{f}_n(\lambda) - f_n(\lambda)$ if a flat-top kernel function is used.

Corollary 3.1. *Assume conditions of Theorem 3.2. If in addition $K \in \mathcal{K}_{FT}$, then (i) for*

$\lambda \in [0, 2\pi) \setminus \{0, \pi\}$ *where* $f_n(\lambda)$ *is bounded away from zero for all large* n*,*

$$\{\kappa^{1/2} f_n(\lambda)\}^{-1}(B_n^{-1}n)^{1/2}\{\hat{f}_n(\lambda) - f_n(\lambda)\} \to_d N(0, 1);$$

and (ii) for $\lambda \in \{0, \pi\}$ *where* $f_n(\lambda)$ *is bounded away from zero for all large* n*,*

$$\{\kappa^{1/2} f_n(\lambda)\}^{-1}(B_n^{-1}n)^{1/2}\{\hat{f}_n(\lambda) - f_n(\lambda)\} \to_d N(0, 2).$$

Corollary 3.1 can be used to guide the construction of confidence intervals for the tail spectral density. The special case when $\lambda = 0$ relates to estimating the asymptotic variance in high quantile regression problems for tail dependent time series; see for example Zhang (2021a). The aforementioned paper considered lag-window estimation of $f_n(0)$, but did not explore the difficult problem about how to quantify the uncertainty of such an estimator. Theorem 3.3 states that tail spectral density estimators at different frequencies are asymptotically independent of each other.

Theorem 3.3. *Assume conditions of Theorem 3.2. Let* $\lambda_1, \ldots, \lambda_L$ *be different frequencies in* $[0, 2\pi)$ *with* $(\lambda_l + \lambda_{l'})/\pi \notin \mathbb{Z}$ *and* $(\lambda_l - \lambda_{l'})/\pi \notin \mathbb{Z}$ *for any* $1 \le l < l' \le L$*. If* $K \in \mathcal{K}_{FT}$ *and* $f_n(\lambda_l)$ *is bounded away from zero for all large* n *for* $l = 1, \ldots, L$*, then* $\{f_n(\lambda_l)\}^{-1}(B_n^{-1}n)^{1/2}\{\hat{f}_n(\lambda_l) - f_n(\lambda_l)\}$*,* $l = 1, \ldots, L$*, converge jointly to independent normal random variables.*

4 Statistical Practice

4.1 Tail Spectral Analysis with Uncertainty Quantification

The results developed in Section 3 can be used to devise a clean and effective method for uncertainty quantification in tail spectral analysis, which adds a valuable piece to the toolbox for practitioners to examine the tail dependence of a given data. In particular, by Corollary 3.1, a $(1-\alpha)$-th confidence interval for $f_n(\lambda)$ can then be constructed as

$$\hat{f}_n(\lambda) \pm (n^{-1}B_n)^{1/2}\kappa^{1/2}\hat{f}_n(\lambda)\Phi^{-1}(1-\alpha/2)$$

if $\lambda/\pi \notin \mathbb{Z}$; and

$$\hat{f}_n(\lambda) \pm (n^{-1}B_n)^{1/2}(2\kappa)^{1/2}\hat{f}_n(\lambda)\Phi^{-1}(1-\alpha/2)$$

if $\lambda/\pi \in \mathbb{Z}$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal. Therefore, similar to the traditional spectral density plot, one can construct its tail counterpart using the tail spectral density estimator and the constructed confidence intervals can assist the practitioner in understanding the associated estimation uncertainty. We would like to emphasize that, without the confidence interval to gauge the statistical uncertainty, it can be very difficult to conduct any meaningful assessment of the tail spectral density, as it remains unknown or at least ambiguous if an observed pattern is indeed systematic or simply due to estimation errors. It can be seen from our data applications in Sections 4.3 and 4.4 that the tail spectral density plot with uncertainty quantification can be useful in revealing dependence information that otherwise cannot be discovered by the traditional spectral analysis, and we expect it to become a routine statistical tool for practitioners to examine dependence in the tail for time series data.

4.2 A Simulation Study

We shall here conduct a simulation study to illustrate the developed results and examine the finite-sample performance of confidence intervals constructed using the developed central limit theorem. For this, we follow Zhang (2022) and consider the moving-maximum process

$$X_i = \max(\epsilon_i, \epsilon_{i-1}/2), \quad i = 1, \ldots, n, \tag{6}$$

where (ϵ_i) is a sequence of independent Fréchet random variables with distribution function $\mathrm{pr}(\epsilon_i \leq z) = \exp(-z^{-\gamma})$ for some $\gamma > 0$. Let $n \in \{1000, 2000\}$ and $\gamma \in \{1, 2, 3\}$, we use the developed central limit theorem to construct confidence intervals for the tail spectral density of (6) at frequencies $\lambda \in \{0, 0.25\pi, 0.5\pi\}$. By Theorem 3.2 and Corollary 3.1, the asymptotic variance is estimated as $2n^{-1}B_n\kappa\{\hat{f}_n(\lambda)\}^2$ for $\lambda = 0$ and $n^{-1}B_n\kappa\{\hat{f}_n(\lambda)\}^2$ for $\lambda \in \{0.25\pi, 0.5\pi\}$. We consider $B_n \in \{15, 30\}$ and chose the tail threshold x_n as the 90% and 95% quantiles. Throughout our numerical experiments, the trapezoidal kernel $K(u) = \max[\min\{2(1 - |u|), 1\}, 0]$ is used, which belongs to the family of flat-top kernels and possesses certain favorable properties; see for example the discussion in Politis and Romano (1995), Politis (2011), and references therein. The results are summarized in Tables 1 and 2, from which we can see that the empirical coverage probabilities are mostly reasonably close to their nominal levels as long as the sample size is reasonably large. Note that a sample size of $n = 200$ is in general a challenging case for tail inference, for which we can also observe a certain degree of size distortion in our numerical results. However, as the sample increases, the size distortion generally gets smaller and the results also seem to become more robust to the choice of B_n. Additional simulation results concerning other settings can be found in the supplementary material. We shall in the following consider two data applications to further illustrate the developed results and make a comparison with the traditional spectral density analysis.

4.3 Application to A Temperature Data

We in this section consider an application to a temperature data that contains the anomaly series of monthly averages of daily high temperatures in the United States from 03/1840 to 05/2016. The data and its detailed description are available through Berkeley Earth at

n	λ/π	$\gamma = 1$		$\gamma = 2$		$\gamma = 3$	
		90%	95%	90%	95%	90%	95%
				x_n chosen as the 90% quantile			
200	0	$0.861_{(0.346)}$	$0.892_{(0.412)}$	$0.838_{(0.290)}$	$0.867_{(0.345)}$	$0.825_{(0.252)}$	$0.864_{(0.301)}$
	0.25	$0.900_{(0.238)}$	$0.935_{(0.284)}$	$0.883_{(0.213)}$	$0.907_{(0.254)}$	$0.873_{(0.191)}$	$0.905_{(0.228)}$
	0.50	$0.882_{(0.167)}$	$0.916_{(0.198)}$	$0.896_{(0.165)}$	$0.920_{(0.196)}$	$0.895_{(0.167)}$	$0.924_{(0.199)}$
500	0	$0.888_{(0.231)}$	$0.924_{(0.275)}$	$0.884_{(0.194)}$	$0.912_{(0.231)}$	$0.880_{(0.174)}$	$0.920_{(0.208)}$
	0.25	$0.915_{(0.154)}$	$0.950_{(0.184)}$	$0.910_{(0.133)}$	$0.939_{(0.159)}$	$0.908_{(0.121)}$	$0.942_{(0.144)}$
	0.50	$0.907_{(0.103)}$	$0.940_{(0.123)}$	$0.892_{(0.106)}$	$0.932_{(0.126)}$	$0.900_{(0.105)}$	$0.932_{(0.125)}$
1000	0	$0.893_{(0.169)}$	$0.937_{(0.202)}$	$0.901_{(0.140)}$	$0.933_{(0.167)}$	$0.888_{(0.124)}$	$0.921_{(0.148)}$
	0.25	$0.909_{(0.108)}$	$0.949_{(0.129)}$	$0.924_{(0.094)}$	$0.951_{(0.112)}$	$0.914_{(0.085)}$	$0.948_{(0.101)}$
	0.50	$0.908_{(0.075)}$	$0.950_{(0.089)}$	$0.906_{(0.074)}$	$0.957_{(0.088)}$	$0.920_{(0.075)}$	$0.957_{(0.089)}$
2000	0	$0.913_{(0.120)}$	$0.949_{(0.143)}$	$0.919_{(0.101)}$	$0.957_{(0.120)}$	$0.887_{(0.089)}$	$0.937_{(0.106)}$
	0.25	$0.910_{(0.076)}$	$0.947_{(0.091)}$	$0.909_{(0.067)}$	$0.952_{(0.079)}$	$0.920_{(0.060)}$	$0.959_{(0.072)}$
	0.50	$0.900_{(0.052)}$	$0.945_{(0.062)}$	$0.917_{(0.052)}$	$0.952_{(0.063)}$	$0.909_{(0.053)}$	$0.952_{(0.063)}$
				x_n chosen as the 95% quantile			
200	0	$0.881_{(0.342)}$	$0.910_{(0.408)}$	$0.862_{(0.296)}$	$0.894_{(0.353)}$	$0.877_{(0.249)}$	$0.906_{(0.297)}$
	0.25	$0.905_{(0.243)}$	$0.932_{(0.290)}$	$0.877_{(0.207)}$	$0.913_{(0.247)}$	$0.883_{(0.189)}$	$0.910_{(0.225)}$
	0.50	$0.908_{(0.166)}$	$0.921_{(0.198)}$	$0.900_{(0.164)}$	$0.927_{(0.196)}$	$0.901_{(0.168)}$	$0.927_{(0.200)}$
500	0	$0.914_{(0.232)}$	$0.947_{(0.276)}$	$0.890_{(0.195)}$	$0.936_{(0.233)}$	$0.881_{(0.172)}$	$0.920_{(0.205)}$
	0.25	$0.921_{(0.154)}$	$0.943_{(0.183)}$	$0.884_{(0.134)}$	$0.921_{(0.159)}$	$0.894_{(0.121)}$	$0.932_{(0.144)}$
	0.50	$0.911_{(0.104)}$	$0.941_{(0.124)}$	$0.896_{(0.105)}$	$0.930_{(0.125)}$	$0.922_{(0.105)}$	$0.950_{(0.125)}$
1000	0	$0.888_{(0.171)}$	$0.942_{(0.204)}$	$0.893_{(0.142)}$	$0.940_{(0.170)}$	$0.889_{(0.126)}$	$0.928_{(0.150)}$
	0.25	$0.892_{(0.108)}$	$0.938_{(0.128)}$	$0.896_{(0.094)}$	$0.944_{(0.112)}$	$0.894_{(0.085)}$	$0.932_{(0.101)}$
	0.50	$0.902_{(0.074)}$	$0.945_{(0.088)}$	$0.897_{(0.074)}$	$0.947_{(0.088)}$	$0.891_{(0.074)}$	$0.946_{(0.088)}$
2000	0	$0.898_{(0.122)}$	$0.938_{(0.145)}$	$0.903_{(0.102)}$	$0.955_{(0.121)}$	$0.882_{(0.889)}$	$0.937_{(0.106)}$
	0.25	$0.902_{(0.076)}$	$0.949_{(0.091)}$	$0.896_{(0.067)}$	$0.953_{(0.079)}$	$0.915_{(0.060)}$	$0.955_{(0.072)}$
	0.50	$0.900_{(0.052)}$	$0.946_{(0.062)}$	$0.910_{(0.052)}$	$0.959_{(0.062)}$	$0.928_{(0.052)}$	$0.960_{(0.062)}$

Table 1: Empirical coverage probabilities (with average lengths in parentheses) of confidence intervals at 90% and 95% nominal levels constructed using the developed central limit theorem for $n \in \{200, 500, 1000, 2000\}$, $\gamma \in \{1, 2, 3\}$, and $B_n = 15$ when the trapezoidal kernel is used.

		$\gamma = 1$		$\gamma = 2$		$\gamma = 3$	
n	λ/π	90%	95%	90%	95%	90%	95%
				x_n chosen as the 90% quantile			
200	0	$0.802_{(0.421)}$	$0.836_{(0.502)}$	$0.798_{(0.365)}$	$0.824_{(0.435)}$	$0.791_{(0.312)}$	$0.823_{(0.371)}$
	0.25	$0.861_{(0.340)}$	$0.901_{(0.405)}$	$0.847_{(0.289)}$	$0.881_{(0.345)}$	$0.866_{(0.272)}$	$0.895_{(0.324)}$
	0.50	$0.869_{(0.235)}$	$0.889_{(0.280)}$	$0.860_{(0.232)}$	$0.886_{(0.277)}$	$0.856_{(0.234)}$	$0.892_{(0.279)}$
500	0	$0.879_{(0.318)}$	$0.904_{(0.379)}$	$0.858_{(0.266)}$	$0.888_{(0.317)}$	$0.849_{(0.237)}$	$0.879_{(0.282)}$
	0.25	$0.878_{(0.215)}$	$0.916_{(0.257)}$	$0.880_{(0.186)}$	$0.912_{(0.222)}$	$0.886_{(0.169)}$	$0.919_{(0.202)}$
	0.50	$0.890_{(0.149)}$	$0.925_{(0.178)}$	$0.877_{(0.149)}$	$0.910_{(0.178)}$	$0.884_{(0.148)}$	$0.908_{(0.176)}$
1000	0	$0.878_{(0.233)}$	$0.917_{(0.277)}$	$0.895_{(0.195)}$	$0.923_{(0.233)}$	$0.884_{(0.173)}$	$0.923_{(0.207)}$
	0.25	$0.905_{(0.152)}$	$0.937_{(0.181)}$	$0.913_{(0.136)}$	$0.943_{(0.162)}$	$0.914_{(0.121)}$	$0.952_{(0.144)}$
	0.50	$0.896_{(0.105)}$	$0.939_{(0.125)}$	$0.905_{(0.105)}$	$0.936_{(0.125)}$	$0.894_{(0.105)}$	$0.935_{(0.125)}$
2000	0	$0.909_{(0.168)}$	$0.944_{(0.201)}$	$0.881_{(0.142)}$	$0.931_{(0.169)}$	$0.887_{(0.124)}$	$0.925_{(0.147)}$
	0.25	$0.903_{(0.107)}$	$0.949_{(0.128)}$	$0.907_{(0.095)}$	$0.943_{(0.113)}$	$0.907_{(0.085)}$	$0.951_{(0.102)}$
	0.50	$0.902_{(0.073)}$	$0.956_{(0.087)}$	$0.912_{(0.074)}$	$0.954_{(0.089)}$	$0.906_{(0.074)}$	$0.940_{(0.088)}$
				x_n chosen as the 95% quantile			
200	0	$0.808_{(0.427)}$	$0.839_{(0.508)}$	$0.816_{(0.363)}$	$0.846_{(0.432)}$	$0.803_{(0.313)}$	$0.827_{(0.373)}$
	0.25	$0.881_{(0.342)}$	$0.902_{(0.408)}$	$0.870_{(0.295)}$	$0.889_{(0.351)}$	$0.874_{(0.267)}$	$0.896_{(0.319)}$
	0.50	$0.871_{(0.238)}$	$0.895_{(0.284)}$	$0.885_{(0.236)}$	$0.913_{(0.281)}$	$0.867_{(0.234)}$	$0.889_{(0.279)}$
500	0	$0.855_{(0.311)}$	$0.894_{(0.370)}$	$0.860_{(0.271)}$	$0.891_{(0.323)}$	$0.854_{(0.227)}$	$0.882_{(0.271)}$
	0.25	$0.893_{(0.216)}$	$0.917_{(0.258)}$	$0.886_{(0.190)}$	$0.916_{(0.227)}$	$0.891_{(0.173)}$	$0.917_{(0.206)}$
	0.50	$0.896_{(0.150)}$	$0.926_{(0.179)}$	$0.886_{(0.148)}$	$0.918_{(0.176)}$	$0.883_{(0.148)}$	$0.914_{(0.177)}$
1000	0	$0.882_{(0.230)}$	$0.924_{(0.274)}$	$0.888_{(0.193)}$	$0.925_{(0.230)}$	$0.863_{(0.172)}$	$0.900_{(0.204)}$
	0.25	$0.913_{(0.154)}$	$0.935_{(0.183)}$	$0.899_{(0.133)}$	$0.937_{(0.158)}$	$0.881_{(0.121)}$	$0.930_{(0.144)}$
	0.50	$0.899_{(0.105)}$	$0.945_{(0.125)}$	$0.908_{(0.105)}$	$0.949_{(0.126)}$	$0.891_{(0.106)}$	$0.932_{(0.127)}$
2000	0	$0.896_{(0.169)}$	$0.936_{(0.201)}$	$0.875_{(0.142)}$	$0.919_{(0.170)}$	$0.892_{(0.125)}$	$0.931_{(0.149)}$
	0.25	$0.905_{(0.109)}$	$0.948_{(0.129)}$	$0.895_{(0.095)}$	$0.945_{(0.113)}$	$0.886_{(0.086)}$	$0.933_{(0.102)}$
	0.50	$0.893_{(0.074)}$	$0.943_{(0.088)}$	$0.902_{(0.074)}$	$0.939_{(0.089)}$	$0.907_{(0.075)}$	$0.952_{(0.089)}$

Table 2: Empirical coverage probabilities (with average lengths in parentheses) of confidence intervals at 90% and 95% nominal levels constructed using the developed central limit theorem for $n \in \{200, 500, 1000, 2000\}$, $\gamma \in \{1, 2, 3\}$, and $B_n = 30$ when the trapezoidal kernel is used.

`http://berkeleyearth.lbl.gov/regions/united-states`. A time series plot is provided in Figure 1, and we shall here provide a tail spectral analysis to study the tail dependence among temperatures from the hottest months. For this, we set the tail threshold as the 95% quantile, and we first apply the high quantile trend analysis procedure of Zhang (2021a) which identified a linear trend. We then apply the developed results to perform a tail spectral analysis on the residuals to study the underlying tail dependence, and make a comparison with the traditional spectral analysis. We use the trapezoidal kernel $K(u) = \max[\min\{2(1-|u|), 1\}, 0]$ from Politis (2011) and the rule of thumb bandwidth choice $B_n = \lfloor n^{1/3} \rfloor = 12$ of Zhang and Wu (2011) for both the tail and traditional spectral analyses so that a comparison can be made. Figure 2 provides the tail spectral density plot (left) and the traditional spectral density plot (right), from which we can see that the tail spectral density function seems to be quite different from the traditional spectral density function indicating that the dependence structure in the tail can be very different from that in the non-tail regions. In addition, there seems to be a peak around 0.5 for the tail spectral density function in the left panel of Figure 2, which relates to a period of 12 months or a yearly cycle. Such a peak does not exist in the traditional spectral density function as plotted in the right panel of Figure 2. This is mainly because the data is an anomaly series, where seasonal patterns have already been removed according to certain climate science calibrations. However, our analysis indicates that, although existing climate science calibrations are able to remove seasonal patterns in the mean, they may not be able to remove patterns in high quantiles at the same time as the seasonal pattern may not be homogeneous across different quantiles. This is in line with the finding of Zhang (2021a), which discovered that the temperature of the hottest days may require a more complicated model than that for modeling the average temperature; see also Eastoe and Tawn (2009)

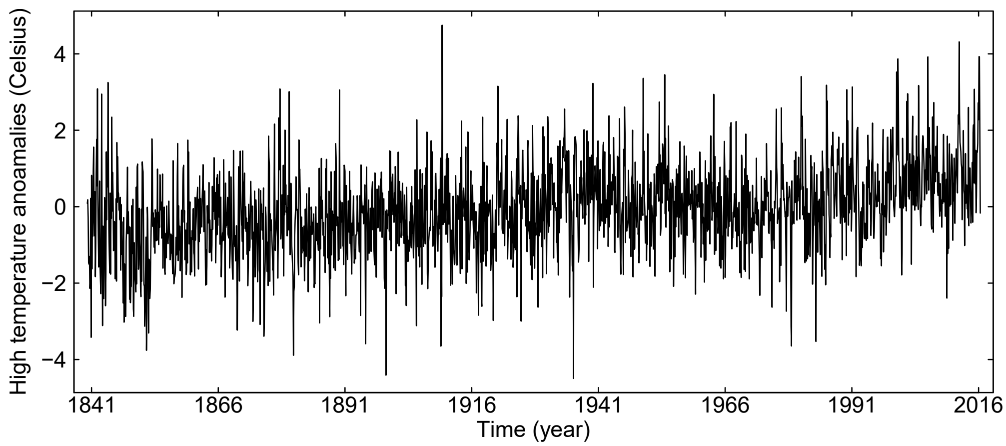

Figure 1: Monthly average anomalies of daily high temperatures in the United States from 03/1840 to 05/2016.

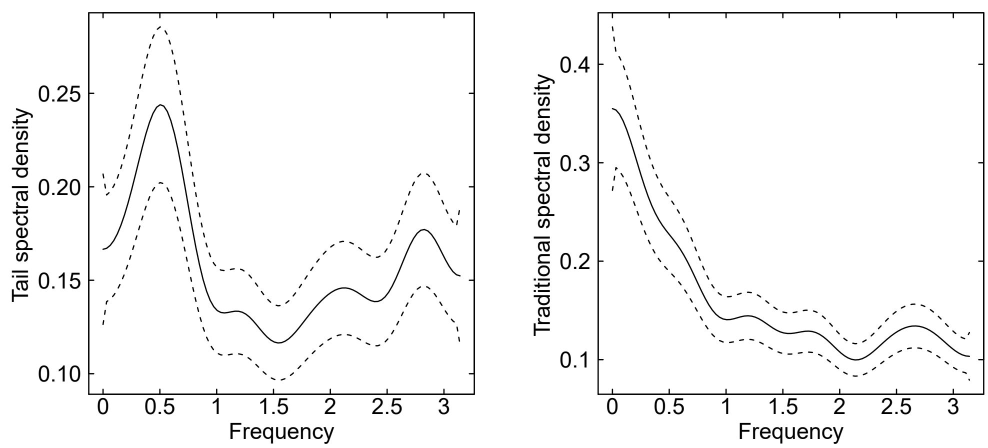

Figure 2: Tail spectral density plot (left) and traditional spectral density plot (right) for the temperature data. In both plots, the two dashed curves correspond to the 95% confidence intervals based on the central limit theorem of (tail) spectral density estimators.

for a discussion on an ozone data.

4.4 Application to A Financial Data

We in this section further illustrate the developed results by considering an application to a financial data that contains the daily adjusted closing price of JPMorgan Chase & Co.

(JPM) from 03/17/1980 to 10/15/2021. The data is available through Yahoo! Finance, and a time series plot is provided in Figure 3. We shall here provide a tail spectral analysis on the lower tail part of the log return series to study the tail dependence among big price drops. The analysis on the upper tail part can be found in the supplementary material. To focus on the lower tail, we set the tail threshold as the 99% quantile of the negative log return series, and use the same kernel and bandwidth choice as in Section 4.3 so that a comparison can be made with the traditional spectral analysis. Figure 4 provides the tail spectral density plot (left) and the traditional spectral density plot (right), from which we can see that the tail spectral density function seems to deviate from a constant more significantly than the traditional spectral density function indicating a higher degree of dependence in the tail. The peak near frequency zero can be related to a positive tail dependence, meaning that big price drops are more often associated with further big price drops. This does not seem to be shared by the traditional spectral density function, indicating that the dependence structure in the tail can be different from that in the mean. In addition, there seems to exist another peak around 0.45 for the tail spectral density function, which can be related to a period of 14 days in tail dependence. In contrast, for the traditional spectral density function, it can be difficult to identify any significant peaks around 0.5 given the confidence interval widths. Even if one chooses to ignore the confidence interval or the associated uncertainty, the first peak of the traditional spectral density estimate seems to be around 0.6 indicating at least a shorter period (if it ever exists). Therefore, the tail spectral density estimation and its uncertainty quantification as considered in the current article seems to provide the practitioners with a useful tool for analyzing tail dependence in the spectral domain, and can lead to discoveries that otherwise cannot be made by the traditional spectral density estimate.

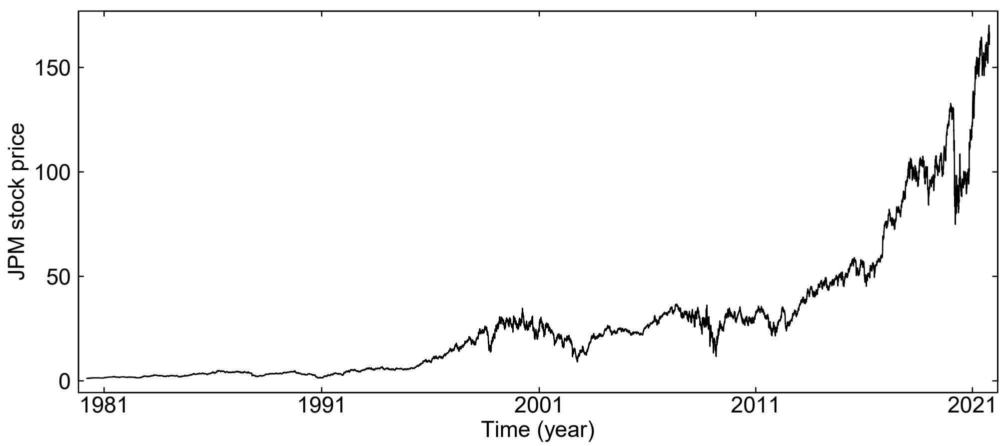

Figure 3: Daily adjusted closing price of JPMorgan Chase & Co. (JPM) from 03/17/1980 to 10/15/2021.

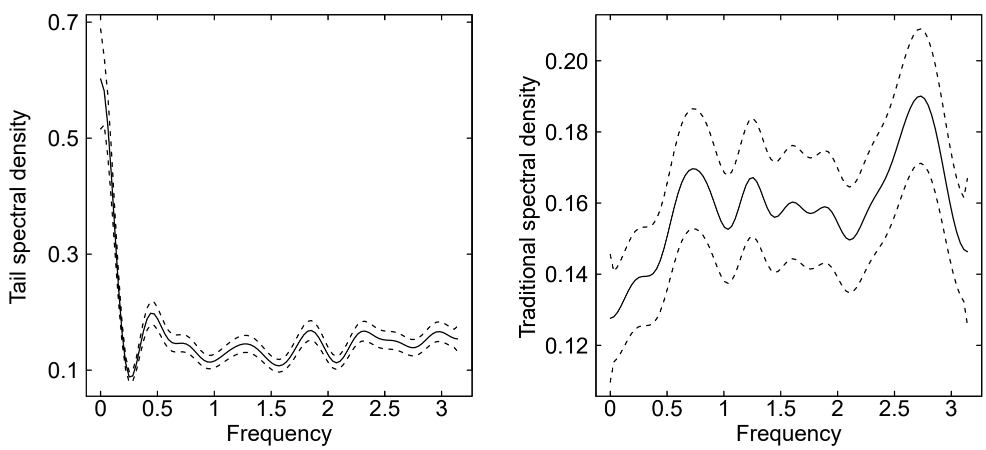

Figure 4: Tail spectral density plot (left) and traditional spectral density plot (right) for the negative log return series of the JPM stock. In both plots, the two dashed curves correspond to the 95% confidence intervals based on the central limit theorem of (tail) spectral density estimators.

5 Conclusion

In the current article, we provide a foundational step toward spectral analysis of tail dependent time series by developing an asymptotic theory on tail spectral density estimators. It is remarkable that developing a central limit theorem for the conventional spectral density estimator is already a highly nontrivial problem as commented by Liu and Wu (2010), and achieving it in the current tail setting can be even more challenging due to the following major difficulties. First, the study of tail phenomena requires a double asymptotic scheme, which, similar to the role of a growing dimension in high-dimensional problems, is typically quite demanding to deal with. Second, unlike the traditional autocorrelations that share a universal convergence rate, sample tail autocorrelations at different lags can exhibit a two-phase asymptotic behavior with different convergence rates (Zhang, 2022), which can pose a challenge to understand the asymptotic behavior of their infinite sums. Third, the study of tail spectral density estimators requires a limit theorem on quadratic forms of tail statistics, which, to the best of our knowledge, still remains largely unknown in the literature and stands as an open problem. We in the current article provide a novel asymptotic theory on quadratic forms of tail statistics in the double asymptotic setting, and use it to develop the desired consistency and central limit theorem for tail spectral density estimators. Apart from the new central limit theorem result that has not been well addressed in the literature, our consistency result itself has already improved over existing ones by allowing more extremal tails and possibly a weaker notion of tail dependence; see the discussions in Section 2.2. Our results provide a foundation for tail spectral analysis of tail dependent time series, and shed new lights on the difficult problem of handling quadratic forms of tail statistics in the double asymptotic scheme. It can be seen from the data applications in Sections 4.3 and 4.4 that the tail spectral density and its uncertainty quantification as considered in the current article is expected to become a useful and powerful tool for practitioners to study tail dependence in the time series setting, and can potentially lead to new scientific discoveries in any discipline that may involve the analysis of tail dependent time series data.

We shall here conclude the article by posing an open question regarding the tail spectral density estimator. In particular, it was shown in the current article that the tail spectral density estimator enjoys a central limit theorem under the geometric tail adversarial stability condition. When the tail adversarial stability measure follows only an algebraic decay as in Theorem 3.1, the tail spectral density estimator is still shown to be consistent but it remains unknown if the central limit theorem will continue to hold in this case. This is intuitively due to the nontrivial asymptotic behavior of sample tail autocorrelations that appear in the construction of tail spectral density estimators. Unlike traditional sample autocorrelations that share a universal convergence rate, sample tail autocorrelations can exhibit a two-phase asymptotic behavior (Zhang, 2022) with dichotomous convergence rates in the current double asymptotic setting. In the algebraic tail adversarial stability case, it becomes nontrivial to quantify the tradeoff between the two phases, and the tradeoff can become ambiguous when a growing number of lags are involved at the same time as in the current problem. We conjecture that the asymptotic distribution of tail spectral density estimators in this case may possibly belong to the more general normal mixture family, and shall here pose it as an open question for future investigation.

Acknowledgement

We thank the Editor, the Associate Editor, two anonymous referees, and a reproducibility reviewer for their helpful comments and suggestions. The research is supported by the NSF CAREER Award DMS-2131821.

Supplementary Material

Supplementary material contains proofs for our main results in Section 3, where some of the results can be useful on their own. It also contains additional simulation results and a supplementary analysis on the financial data studied in Section 4.4.

References

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

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59**, 817–858.

Bai, S. and Zhang, T. (2022). Tail adversarial stability for regularly varying linear processes and their extensions. *arXiv*, https://arxiv.org/abs/2205.00043.

Balla, E., Ergen, I. and Migueis, M. (2014). Tail dependence and indicators of systemic risk for large us depositories. *Journal of Financial Stability*, **15**, 195–209.

Bentkus, R. Yu. and Rudzkis, R. A. (1982). On the distribution of some statistical estimates of spectral density. *Theory of Probability & Its Applications*, **27**, 795–814.

Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory.* Holden-Day, San Francisco.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods.* Springer-Verlag, New York, 2nd ed.

Coles, S., Heffernan, J. and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, **2**, 339–365.

Dahlhaus, R. (1985). On a spectral density estimate obtained by averaging periodograms. *Journal of Applied Probability*, **22**, 598–610.

Davis, R. A. and Mikosch, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli*, **15**, 977–1009.

de Haan, L. and Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **40**, 317–337.

Dombry, C. and Eyi-Minko, F. (2012). Strong mixing properties of max-infinitely divisible random fields. *Stochastic Processes and their Applications*, **122**, 3790–3811.

Draisma, G., Drees, H., Ferreira, A. and De Haan, L. (2004). Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli*, **10**, 251–280.

Eastoe, E. F. and Tawn, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **58**, 25–45.

Embrechts, P., Mcneil, A. and Straumann, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. In M. A. H. Dempster, editor, *Risk Management: Value at Risk and Beyond*. Cambridge University Press, pp. 176–223.

Flegal, J. M. and Jones, G. L. (2010). Batch means and spectral variance estimators in markov chain monte carlo. *The Annals of Statistics*, **38**, 1034–1070.

Giakoumatos, S. G., Vrontos, I. D., Dellaportas, P. and Politis, D. N. (1999). A Markov chain Monte Carlo convergence diagnostic using subsampling. *Journal of Computational and Graphical Statistics*, **8**, 431–451.

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

Hall, P., Peng, L. and Yao, Q. (2002). Moving-maximum models for extrema of time series. *Journal of Statistical Planning and Inference*, **103**, 51–63.

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

Hill, J. B. (2009). On functional central limit theorems for dependent, heterogeneous arrays with applications to tail index and tail dependence estimation. *Journal of Statistical Planning and Inference*, **139**, 2091–2110.

Hoga, Y. (2018). A structural break test for extremal dependence in β-mixing random vectors. *Biometrika*, **105**, 627–643.

Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis*, **46**, 262–282.

Lahiri, S. (2003). *Resampling Methods for Dependent Data*. Springer, New York.

Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, **83**, 169–187.

Linton, O. B. and Whang, Y.-J. (2007). The quantilogram: With an application to evaluating directional predictability. *Journal of Econometrics*, **141**, 250–282.

Liu, W. and Wu, W. B. (2010). Asymptotics of spectral density estimates. *Econometric Theory*, **26**, 1218–1245.

McNeil, A. J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management–Concept, Techniques and Tools*. Princeton University Press, Princeton, NJ.

Mikosch, T. and Zhao, Y. (2014). A fourier analysis of extreme events. *Bernoulli*, **20**, 803–845.

Mikosch, T. and Zhao, Y. (2015). The integrated periodogram of a dependent extremal event sequence. *Stochastic Processes and their Applications*, **125**, 3126–3169.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**, 703–708.

Phillips, P. C. B., Sun, Y. and Jin, S. (2006). Spectral density estimation and robust hypothesis testing using steep origin kernels without truncation. *International Economic*

Review, **47**, 837–894.

Phillips, P. C. B., Sun, Y. and Jin, S. (2007). Long run variance estimation and robust regression testing using sharp origin kernels with no truncation. *Journal of Statistical Planning and Inference*, **137**, 985–1023.

Politis, D. N. (2011). Higher-order accurate, positive semi-definite estimation of large-sample covariance and spectral density matrices. *Econometric Theory*, **27**, 703–744.

Politis, D. N. and Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics*, **22**, 2031–2050.

Politis, D. N. and Romano, J. P. (1995). Bias-corrected nonparametric spectral estimation. *Journal of Time Series Analysis*, **16**, 67–103.

Politis, D. N., Romano, J. P. and Wolf, M. (1999). *Subsampling*. Springer, New York.

Poon, S.-H., Rockinger, M. and Tawn, J. (2004). Extreme value dependence in financial markets: diagnostics, models, and financial implications. *The Review of Financial Studies*, **17**, 581–610.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*, vol. 1. Academic Press.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*, **42**, 43–47.

Rosenblatt, M. (1984). Asymptotic normality, strong mixing and spectral density estimates. *The Annals of Probability*, **12**, 1167–1180.

Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Birkhäuser, Boston.

Rosenblatt, M. (2009). A comment on a conjecture of N. Wiener. *Statistics & Probability Letters*, **79**, 347–348.

Shao, X. and Wu, W. B. (2007). Asymptotic spectral theory for nonlinear time series.

The Annals of Statistics, **35**, 1773–1801.

Sibuya, M. (1960). Bivariate extreme statistics, I. *Annals of the Institute of Statistical Mathematics*, **11**, 195–210.

Song, W. T. and Schmeiser, B. W. (1995). Optimal mean-squared-error batch sizes. *Management Science*, **41**, 110–123.

Tong, H. (1990). *Non-Linear Time Series: A Dynamical System Approach (Oxford Statistical Science Series, 6).* Oxford University Press, U.K.

Velasco, C. and Robinson, P. M. (2001). Edgeworth expansions for spectral density estimates and Studentized sample mean. *Econometric Theory*, **17**, 497–539.

Wiener, N. (1958). *Nonlinear Problems In Random Theory (Technology Press Research Monographs).* The MIT Press, Cambridge.

Wu, W. B. (2005). Nonlinear 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. (2009). Recursive estimation of time-average variance constants. *The Annals of Applied Probability*, **19**, 1529–1552.

Wu, W. B. and Shao, X. (2007). A limit theorem for quadratic forms and its applications. *Econometric Theory*, **23**, 930–951.

Xiao, H. and Wu, W. B. (2011). A single-pass algorithm for spectrum estimation with fast convergence. *IEEE Transactions on Information Theory*, **57**, 4720–4731.

Zhang, T. (2018). A thresholding-based prewhitened long-run variance estimator and its dependence oracle property. *Statistica Sinica*, **28**, 319–338.

Zhang, T. (2021a). High-quantile regression for tail-dependent time series. *Biometrika*, **108**, 113–126.

Zhang, T. (2022). Asymptotics of sample tail autocorrelations for tail dependent time series: phase transition and visualization. *Biometrika*, forthcoming.

Zhang, T. and Wu, W. B. (2011). Testing parametric assumptions of trends of a nonstationary time series. *Biometrika*, **98**, 599–614.

Zhang, Z. (2005). A new class of tail-dependent time series models and its applications in financial time series. *Advances in Econometrics*, **20**, 323–358.

Zhang, Z. (2008). Quotient correlation: A sample based alternative to Pearson's correlation. *The Annals of Statistics*, **36**, 1007–1030.

Zhang, Z. (2021b). On studying extreme values and systematic risks with nonlinear time series models and tail dependence measures. *Statistical Theory and Related Fields*, **5**, 1–25.

Zhang, Z. and Smith, R. L. (2004). The behavior of multivariate maxima of moving maxima processes. *Journal of Applied Probability*, **41**, 1113–1123.

Zhang, Z., Zhang, C. and Cui, Q. (2017). Random threshold driven tail dependence measures with application to precipitation data analysis. *Statistica Sinica*, **27**, 685–709.