Biometrika (2021), **xx**, x, *pp*. 1–15
Printed in Great Britain

Asymptotics of Sample Tail Autocorrelations for Tail Dependent Time Series: Phase Transition and Visualization

BY TING ZHANG
Department of Statistics, University of Georgia
310 Herty Drive, Athens, GA 30602, U.S.A.
tingzhang@uga.edu

SUMMARY

This article develops an asymptotic theory on sample tail autocorrelations of time series data that can exhibit serial dependence in both tail and non-tail regions. Unlike the traditional autocorrelation function, the study of tail autocorrelations requires a double asymptotic scheme to capture the tail phenomena, and our results do not impose any restriction on the dependence structure in non-tail regions and allow processes that are not necessarily strong mixing. The newly developed asymptotic theory reveals a previously undiscovered phase transition phenomenon for sample tail autocorrelations, whose asymptotic behavior including the convergence rate can transit from one phase to the other when the lag index moves past the point beyond which serial tail dependence vanishes. The phase transition discovery fills the gap of existing research on tail autocorrelations, and can be used to construct the lines of significance, in analogy to the traditional autocorrelation plot, when visualizing sample tail autocorrelations to assess the existence of serial tail dependence or to identify the maximal lag of tail dependence.

Some key words: Double asymptotics; phase transition; tail adversarial stability; tail autocorrelation function; tail dependent time series

1. INTRODUCTION

Tail dependence, also known as asymptotic dependence or extremal dependence, appears in data applications from a wide range of disciplines, including actuarial science, climate science, economics, finance, hydrology, and internet traffic engineering. The phenomenon in the bivariate or finite-dimensional multivariate setting has been extensively studied by Sibuya (1960), de Haan & Resnick (1977), Ledford & Tawn (1996), Embrechts et al. (2002), Draisma et al. (2004), Poon et al. (2004), Zhang (2008) and Balla et al. (2014) among others; see also Joe (1993), Coles et al. (1999) and McNeil et al. (2005) for copula-based approaches. Compared with the vast literature on tail dependence when independent samples are available for the finite-dimensional joint distribution of interest, the problem has been much less explored in the time series setting and was regarded as a somewhat open topic by Davis & Mikosch (2009). There have been two main approaches of summarizing the strength of serial tail dependence in a time series. The first is through the extremal index (Leadbetter et al., 1983; Smith & Weissman, 1994; Ferro & Segers, 2003), which serves as an adjustment factor for the distribution of the maximum and is similar in spirit to the adjustment of replacing the marginal variance with the long-run variance in mean inference. The second main approach is to find a tail counterpart for the traditional autocorrelation function (ACF), and results along this line include the lag-k tail dependence index of Zhang (2005), the quantilogram of Linton & Whang (2007), and the extremogram of

Davis & Mikosch (2009). Unlike the extremal index, such an approach makes it possible to visualize tail dependence at different lags in analogy to the widely used traditional ACF plot.

The tail autocorrelation is a remarkable type of extremogram that relates to the quantilogram of Linton & Whang (2007) by allowing the quantile level to grow with the sample size, which can also be viewed as a standardized pre-asymptotic counterpart of the lag-k tail dependence index of Zhang (2005). Davis & Mikosch (2009) obtained a central limit theorem for sample tail autocorrelations using the mixing framework of Rosenblatt (1956) along with an additional anti-clustering condition that further controls the strength of tail dependence; see also Hill (2009) for processes that can be well approximated by functions of a mixing sequence. The task of obtaining a tail counterpart of the traditional ACF plot, however, can still be nontrivial due to the lack of theoretical understanding of sample tail autocorrelations when there exists a lag beyond which tail dependence vanishes; see for example the discussion in Section 3.2 of Davis et al. (2012). In particular, the central limit theorem developed in Davis & Mikosch (2009) is not directly useful to fill this gap as it will lead to a degenerate and noninformative limit in this case. Such a degeneracy issue does not exist for the traditional autocorrelation, and is mainly caused by the double asymptotic scheme that is necessary for studying tail autocorrelations. In addition, the anti-clustering condition of Davis & Mikosch (2009) involves an interplay between how fast the mixing coefficients decay to zero and how fast the quantile level approaches the extreme, which can lead to additional inexplicit restrictions on how extremal the tail can be; see also the discussion in Davis et al. (2013).

The current article aims at providing a more comprehensive theoretical understanding of sample tail autocorrelations to discover and study their unusual but characteristic phase transition phenomenon. In particular, our newly developed asymptotic theory reveals that the asymptotic behavior of sample tail autocorrelations can transit from one phase to the other when the lag index passes the point beyond which serial tail dependence vanishes. To be more specific, before the lag index reaches such a point, the convergence rate of sample tail autocorrelations is the same as that in Davis & Mikosch (2009), which is the square root of the expected number of tail observations. Compared with Davis & Mikosch (2009), our results do not impose any restriction on the dependence structure in non-tail regions, involve less restrictive conditions on how extremal the tail can be, and allow processes that are not necessarily strong mixing or regularly varying. In addition, unlike the big blocks small blocks argument that has been commonly used for mixing processes as in Davis & Mikosch (2009), our proof uses an m-dependent martingale approximation scheme to obtain the desired limit theorem. Once the lag index moves past the point beyond which tail dependence vanishes, then the central limit theorem of Davis & Mikosch (2009) becomes degenerate and noninformative, while our results provide a deeper insight and discover that the asymptotic behavior of sample tail autocorrelations in this case actually transits into a second phase, for which the convergence rate becomes the square root of the sample size, which is faster in magnitude than that in the first phase. Our phase transition discovery fills the gap of existing research on tail autocorrelations, and can be used to construct the lines of significance when visualizing tail autocorrelations to assess the existence of serial tail dependence. Such a theoretical approach of obtaining the lines of significance is in analogy to the widely used traditional ACF plot, which provides a clean and convenient alternative to the bootstrap and permutation approaches of Davis et al. (2012) and addresses the open problem posed in their Section 3.2. In addition, our results can also be useful toward the problem of identifying the maximal lag of serial tail dependence as considered by Zhang (2005).

2. Tail Adversarial Stability: Framework

Suppose we observe $Y_1, \ldots, Y_n$ from a stationary process

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

where ϵ_j, $j \in \mathbb{Z}$, are independent and identically distributed (i.i.d.) innovations, and G is a measurable function such that Y_i is properly defined. The causal representation (1) covers a wide range of processes, and the function G can be interpreted as a physical system with $\mathcal{F}_i$ being the input and Y_i being the output; see for example Wiener (1958), Tong (1990), Wu (2011), Zhang (2018) and references therein. Let $F(y) = \mathrm{pr}(Y_1 \leq y)$, $y \in \mathbb{R}$, denote the distribution of (1) with inverse $F^{-1}(u) = \inf\{y : F(y) \geq u\}$, we write $\mathcal{U}_F = \lim_{u \uparrow 1} F^{-1}(u)$ which represents the upper end point of the distribution and can take the value of infinity if the distribution is not bounded. Then as y approaches $\mathcal{U}_F$, data points exceeding y can be viewed as tail events, and a popular approach of summarizing tail dependence in a time series is to use the conditional probability

$$\nu(k) = \lim_{y \uparrow \mathcal{U}_F} \nu_y(k), \quad \nu_y(k) = \mathrm{pr}(Y_{1+k} > y \mid Y_1 > y), \tag{2}$$

which naturally generalizes the bivariate tail dependence coefficient of Sibuya (1960); see for example Ledford & Tawn (2003), Zhang (2005), Zhang & Huang (2006), Linton & Whang (2007), Davis & Mikosch (2009) and references therein. Similar to the role of traditional autocorrelations, although quantities in (2) provide a straightforward summary of the underlying tail dependence, they are generally not directly useful for developing limit theorems by themselves. For this, a common approach in the literature is to use the strong mixing framework of Rosenblatt (1956) and its variants; see for example Drees (2003), Chernozhukov (2005), Davis & Mikosch (2009), Drees & Rootzén (2010), Davis et al. (2018), Hoga (2018) and references therein. To handle tail events, however, the mixing condition often has to be used together with additional conditions that control the strength of tail dependence; see for example condition (9.67) of Chernozhukov (2005), conditions (3.3) and (3.10) of Davis & Mikosch (2009), and conditions (2.2) and (2.4) of Davis et al. (2013), which can possibly lead to additional inexplicit technical restrictions. Additional contributions can be found in Basrak & Segers (2009), Hill (2009), Kulik et al. (2019) and references therein.

We shall here follow Zhang (2021) and consider an alternative framework that relies on understanding the tail effect of adversarial innovations. 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 $Y_i^\star = G(\mathcal{F}_i^\star)$ represents the output of the physical system G when the innovation at time zero is replaced by its i.i.d. copy. We consider

$$\theta_y(i) = \sup_{z \geq y} \mathrm{pr}(Y_i^\star \leq z \mid Y_i > z), \tag{3}$$

which quantifies the degree to which the input innovation at time zero affects whether the output data at time i is a tail observation. In particular, if Y_i does not depend on ϵ_0, then $Y_i^\star = Y_i$ and thus $\theta_y(i) = 0$. Since replacing ϵ_0 by its i.i.d. copy changes whether Y_i is a tail observation, (3) measures the tail adversarial effect of ϵ_0 on Y_i. Let

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

which measures the cumulative tail adversarial effect of the current innovation ϵ_0 on future observations from time m. We say that the process $Y_1, Y_2, \ldots$ is tail adversarial q-stable or

$(Y_i) \in \mathrm{TAS}_q$ if

$$\lim_{y \uparrow \mathcal{U}_F} \Theta_{y,q}(0) < \infty, \tag{4}$$

namely the current innovation has a finite cumulative tail adversarial effect on future observations. Since y can be chosen arbitrarily close to the upper end point $\mathcal{U}_F$, the tail adversarial stability condition (4) only concerns the dependence in the upper tail region and does not impose any restriction on the middle range or lower tail region. Compared with the bivariate tail dependence coefficient (2) of Ledford & Tawn (2003) and Zhang (2005), it incorporates additional time series structures through tail adversarial coupling in the causal representation of Wiener (1958) and is directly useful in coordinating with an m-dependent martingale approximation scheme to obtain desired limit theorems for tail dependent time series. In the context of high quantile regression models, Zhang (2021) observed that the tail adversarial stability framework can lead to weaker conditions than existing results using the strong mixing framework of Rosenblatt (1956). We shall in the following section consider the problem of tail autocorrelation inference and study its unique phase transition phenomenon. The practical meaning of (4) is illustrated using the popular moving-maximum process of Hall et al. (2002) in Section 4.3.

3. ASYMPTOTICS OF SAMPLE TAIL AUTOCORRELATIONS

3.1. *Sample Tail Autocorrelation*

Given a stationary time series $Y_1, \ldots, Y_n$, the tail autocorrelation is defined as

$$\begin{aligned} \tau_y(k) &= \frac{\mathrm{pr}(Y_{1+k} > y \mid Y_1 > y) - \mathrm{pr}(Y_{1+k} > y)}{1 - \mathrm{pr}(Y_1 > y)} \\ &= \frac{\mathrm{pr}(Y_{1+k} > y, Y_1 > y) - \mathrm{pr}(Y_{1+k} > y)\mathrm{pr}(Y_1 > y)}{\mathrm{pr}(Y_1 > y)\{1 - \mathrm{pr}(Y_1 > y)\}}, \end{aligned} \tag{5}$$

which centers and standardizes the serial tail dependence coefficient $\nu_y(k)$ in (2) in the spirit of a correlation coefficient. By choosing the hit threshold y in the form of a quantile as suggested in Davis & Mikosch (2009), (5) gives the quantilogram of Linton & Whang (2007). It also relates to the lag-k tail dependence index of Zhang (2005) for constructing the M3 process to model tail dependent time series. Let $\mathrm{I}(\cdot)$ denote the indicator function, $\bar{F}(y) = 1 - F(y)$ and $\hat{\bar{F}}_n(y) = n^{-1} \sum_{i=1}^n \mathrm{I}(Y_i > y)$, then a sample version of (5) is given by

$$\hat{\tau}_{n,y}(k) = \frac{\hat{\mu}_{n,y}(k)}{\hat{\mu}_{n,y}(0)}, \quad \hat{\mu}_{n,y}(k) = \frac{1}{n} \sum_{i=1}^{n-|k|} \{\mathrm{I}(Y_i > y) - \hat{\bar{F}}_n(y)\}\{\mathrm{I}(Y_{i+|k|} > y) - \hat{\bar{F}}_n(y)\}.$$

Unlike the setting of Linton & Whang (2007) and Han et al. (2016), we shall here consider the situation when $y = y_n$ is allowed to approach the upper end point $\mathcal{U}_F$ as $n \to \infty$ to focus on the tail. Under this double asymptotic scheme, we study the Phase I and Phase II asymptotic behavior of sample autocorrelations in Sections 3.2 and 3.3 respectively.

3.2. *Asymptotic Theory: Phase I*

We first provide an asymptotic theory on the triangular array $\hat{\mu}_{n,y_n}(k)$, which serves as the sample tail autocovariance that estimates $\mathrm{pr}(Y_{1+k} > y_n, Y_1 > y_n) - \mathrm{pr}(Y_{1+k} > y_n)\mathrm{pr}(Y_1 > y_n) = \tau_{y_n}(k)\bar{F}(y_n)F(y_n)$. For this, we need the following additional notations. Let $\mathcal{P}_l(\cdot) = E(\cdot \mid \mathcal{F}_l) - E(\cdot \mid \mathcal{F}_{l-1})$, $l \in \mathbb{Z}$, be the projection operator, and we write $\|Y\|_p = \{E(|Y|^p)\}^{1/p}$

with the convention that $\|\cdot\| = \|\cdot\|_2$. Let $\zeta_{i,n} = \mathrm{I}\{G(\mathcal{F}_i) > y_n\} - \bar{F}(y_n)$, and define

$$\sigma_{\mu,n}(k)^2 = \{\bar{F}(y_n)\}^{-1} \left\| \mathcal{P}_{|k|} \left(\sum_{l=0}^{\infty} \zeta_{l,n}\zeta_{l+|k|,n} \right) \right\|^2 . \tag{6}$$

Theorem 1 provides the Phase I convergence rate and central limit theorem for $\hat{\mu}_{n,y_n}(k)$ when y_n approaches the upper end point $\mathcal{U}_F$ as $n \to \infty$.

THEOREM 1. *Assume that $(Y_i) \in \mathrm{TAS}_q$ for some $q > 4$. If $y_n \uparrow \mathcal{U}_F$ satisfies $\bar{F}(y_n) \to 0$ and $n\bar{F}(y_n) \to \infty$, then*

$$\hat{\mu}_{n,y_n}(k) - \tau_{y_n}(k)\bar{F}(y_n)F(y_n) = O_p\left[\left\{ \frac{\bar{F}(y_n)}{n} \right\}^{1/2} \right],$$

and $\limsup_{n\to\infty} \sigma_{\mu,n}(k)^2 < \infty$. If in addition $\liminf_{n\to\infty} \sigma_{\mu,n}(k)^2 > 0$, then

$$\left\{ \frac{n}{\bar{F}(y_n)\sigma_{\mu,n}(k)^2} \right\}^{1/2} \{\hat{\mu}_{n,y_n}(k) - \tau_{y_n}(k)\bar{F}(y_n)F(y_n)\} \to_d N(0,1),$$

where $\to_d$ denotes convergence in distribution.

We shall here provide a discussion on the quantity (6) that appears in the central limit theorem of $\hat{\mu}_{n,y_n}(k)$. In particular, if the time series $(Y_i) \in \mathrm{TAS}_q$ for some $q \geq 4$, then one can show that $\limsup_{n\to\infty} \sigma_{\mu,n}(k)^2 < \infty$, which motivated the scale adjustment $\{\bar{F}(y_n)\}^{-1}$ in (6). The condition that it is bounded away from zero for all large n is needed to ensure that the central limit theorem of $\hat{\mu}_{n,y_n}(k)$ as in Theorem 1 is not degenerate, and we shall here provide an explicit calculation of the quantity (6) for regularly varying time series models. In particular, assuming that the time series (Y_i) is regularly varying as in the framework of Davis & Mikosch (2009), then there exist constants $c_{l_1,\ldots,l_d}$ such that

$$\{\bar{F}(y_n)\}^{-1} \cdot \mathrm{pr}(Y_{l_1} > y_n, \ldots, Y_{l_d} > y_n) \to c_{l_1,\ldots,l_d} \tag{7}$$

holds for any tuple $l_1, \ldots, l_d$; see assumptions (1.1) and (1.2) of Davis & Mikosch (2009). In this case, one can show that $\sigma_{\mu,n}(k)^2$ converges to the limit $\sum_{l\in\mathbb{Z}} c_{0,|k|,l,l+|k|}$, whose summands are all nonnegative in view of (7). Besides the regularly varying joint distribution, the result of Davis & Mikosch (2009) also requires the condition that $n\{\bar{F}(y_n)\}^3 \to \infty$, which is more restrictive than the current condition $n\bar{F}(y_n) \to \infty$ about how extremal the tail can be. Although Davis & Mikosch (2009) made an effort in providing an alternative to $n\{\bar{F}(y_n)\}^3 \to \infty$, their alternative condition depends on a nontrivial interplay between how fast the mixing coefficient decays to zero and how extremal the tail can be, which in general is still stronger than the current $n\bar{F}(y_n) \to \infty$ especially when the dependence follows an algebraic decay; see for example the application to the moving-maximum process of Hall et al. (2002) in Section 4.3. Let

$$\sigma_{\nu,n}(k)^2 = \{\bar{F}(y_n)\}^{-1} \left\| \mathcal{P}_{|k|} \left[\sum_{l=0}^{\infty} \{\zeta_{l,n}\zeta_{l+|k|,n} - \tau_{y_n}(k)\zeta_{l,n}^2\} \right] \right\|^2 .$$

Corollary 1 provides the Phase I convergence rate and central limit theorem for the sample autocorrelation $\hat{\tau}_{n,y_n}(k)$.

COROLLARY 1. *Assume that $(Y_i) \in \mathrm{TAS}_q$ for some $q > 4$. If $y_n \uparrow \mathcal{U}_F$ satisfies $\bar{F}(y_n) \to 0$ and $n\bar{F}(y_n) \to \infty$, then*

$$\hat{\tau}_{n,y_n}(k) - \tau_{y_n}(k) = O_p[\{n\bar{F}(y_n)\}^{-1/2}],$$

and $\limsup_{n\to\infty} \sigma_{\nu,n}(k)^2 < \infty$. *If in addition* $\liminf_{n\to\infty} \sigma_{\nu,n}(k)^2 > 0$, *then*

$$\left\{\frac{n\bar{F}(y_n)}{\sigma_{\nu,n}(k)^2}\right\}^{1/2} \{\hat{\tau}_{n,y_n}(k) - \tau_{y_n}(k)\} \to_d N(0,1).$$

It can be seen from the proof of Corollary 1 that the randomness of $\hat{\tau}_{n,y_n}(k) - \tau_{y_n}(k)$ is dominated by the leading term $\{\bar{F}(y_n)F(y_n)\}^{-1}\{\hat{\mu}_{n,y_n}(k) - \tau_{y_n}(k)\hat{\mu}_{n,y_n}(0)\}$, which involves a linear combination of $\hat{\mu}_{n,y_n}(k)$ and $\hat{\mu}_{n,y_n}(0)$ that has zero expectation. Unlike the traditional autocovariance and autocorrelation, the convergence rate of the sample tail autocovariance $\hat{\mu}_{n,y_n}(k)$ is faster than the traditional $n^{1/2}$, while the convergence rate of the sample tail autocorrelation $\hat{\tau}_{n,y_n}(k)$ is slower. Such a discrepancy is due to the fact that, in the current tail setting, $\hat{\mu}_{n,y_n}(k)$ estimates $\tau_{y_n}(k)\bar{F}(y_n)F(y_n)$ which itself goes to zero as $n \to \infty$. Therefore, if we scale it and focus on the tail autocorrelation $\tau_{y_n}(k)$ that takes value in $[-1, 1]$, then the convergence rate in this case becomes slower.

3.3. *Asymptotic Theory: Phase II*

The traditional ACF plot has been widely used by practitioners from various disciplines to determine the existence of a lag beyond which the dependence vanishes, so that a model with finite-order dependence can be used for the data. This motivates us to consider the situation when

$$Y_i = G(\mathcal{F}_{i-K,i}), \quad \mathcal{F}_{i-K,i} = (\epsilon_{i-K}, \ldots, \epsilon_i), \tag{8}$$

for some $K \geq 0$, which represents the lag beyond which there is no serial dependence. The successfulness of the traditional ACF plot that leads to its wide recognition stems from the well-known asymptotic theory of sample autocorrelations under (8) that provides a clean and effective way of adding the dashed lines of significance in a traditional ACF plot; see for example Tsay (2010). The problem of developing a counterpart for the tail autocorrelation, however, is still an open one due to the lack of theoretical understanding of sample tail autocorrelations under (8); see for example the discussion in Section 3.2 of Davis et al. (2012). The aforementioned paper considered a bootstrap approach based on the central limit theorem of Davis & Mikosch (2009). In the current setting, however, once the lag index passes K, then the central limit theorem of Davis & Mikosch (2009) becomes degenerate and noninformative for inference. This is because under the regularly varying time series model (7) of Davis & Mikosch (2009) and Davis et al. (2012), the constants $c_{0,|k|,l,l+|k|} = 0$ for $|l| \geq |k| > K$, making the asymptotic variance in their central limit theorem zero. It can also be seen from our Phase I asymptotic theory in Section 3.2, where one can show that the quantity in (6) satisfies $\sigma_{\mu,n}(k)^2 \to 0$ if $|k| > K$. In this case, although $\hat{\mu}_{n,y_n}(0)$ may continue to have a non-degenerate Phase I central limit theorem, it will no longer appear in the leading term of the sample autocorrelation due to the fact that $\tau_{y_n}(k) = 0$ for $|k| > K$. Therefore, it seems desirable if we can develop a more detailed distribution theory on sample tail autocorrelations when the Phase I central limit theorem becomes degenerate and noninformative for inference. Let

$$\varsigma_{\mu,n}(k)^2 = \{F(y_n)\bar{F}(y_n)\}^{-2} \sum_{|l|\leq K} E(\zeta_{0,n}\zeta_{l,n}\zeta_{|k|,n}\zeta_{l+|k|,n}), \tag{9}$$

we shall here fill this gap and provide the Phase II convergence rates and central limit theorems for $\hat{\mu}_{n,y_n}(k)$ and $\hat{\tau}_{n,y_n}(k)$ in Theorem 2 and Corollary 2 respectively.

THEOREM 2. *Assume that* (8) *holds for some* $K \geq 0$. *If* $y_n \uparrow \mathcal{U}_F$ *satisfies* $\bar{F}(y_n) \to 0$ *and* $n\bar{F}(y_n) \to \infty$, *then for any* $|k| > K$,

$$\hat{\mu}_{n,y_n}(k) = O_p\{n^{-1/2}\bar{F}(y_n)\},$$

and $\limsup_{n\to\infty} \varsigma_{\mu,n}(k)^2 < \infty$. *If in addition* $n\{\bar{F}(y_n)\}^2 \to \infty$ *and* $\liminf_{n\to\infty} \varsigma_{\mu,n}(k)^2 > 0$, *then*

$$\left[\frac{n}{\{\bar{F}(y_n)\}^2 \varsigma_{\mu,n}(k)^2}\right]^{1/2} \hat{\mu}_{n,y_n}(k) \to_d N(0,1).$$

Compared with the Phase I asymptotic theory in Theorem 1, the Phase II convergence rate in Theorem 2 is faster by a factor of $\{\bar{F}(y_n)\}^{-1/2}$. Therefore, when the Phase I central limit theorem in Theorem 1 becomes degenerate and noninformative for inference, the Phase II asymptotic theory in Theorem 2 not only discovers the rate at which it degenerates to zero but also establishes the updated limiting distribution when the new scale adjustment is used. Such a phase transition phenomenon typically does not exist in traditional autocovariances, for which the convergence rate is the universal $n^{1/2}$. Our results in Theorems 1 and 2 reveal that, when the lag index passes the point beyond which tail dependence vanishes, the asymptotic behavior of sample tail autocovariances can exhibit a phase transition, where the convergence rate transits from the Phase I rate of $n^{1/2}\{\bar{F}(y_n)\}^{-1/2}$ to the Phase II rate of $n^{1/2}\{\bar{F}(y_n)\}^{-1}$.

COROLLARY 2. *Assume that* (8) *holds for some* $K \geq 0$. *If* $y_n \uparrow \mathcal{U}_F$ *satisfies* $\bar{F}(y_n) \to 0$ *and* $n\bar{F}(y_n) \to \infty$, *then for any* $|k| > K$,

$$\hat{\tau}_{n,y_n}(k) = O_p(n^{-1/2}).$$

If in addition $n\{\bar{F}(y_n)\}^2 \to \infty$ *and* $\liminf_{n\to\infty} \varsigma_{\mu,n}(k)^2 > 0$, *then*

$$\left\{\frac{n}{\varsigma_{\mu,n}(k)^2}\right\}^{1/2} \hat{\tau}_{n,y_n}(k) \to_d N(0,1).$$

By Corollary 2, sample tail autocorrelations continue to exhibit the phase transition as in sample tail autocovariances. In particular, the Phase I convergence rate in Corollary 1 is $\{n\bar{F}(y_n)\}^{1/2}$, where $n\bar{F}(y_n)$ serves as the expected sample size in the tail region that can be used to interpret the intuition of such a convergence rate. The Phase II convergence rate in Corollary 2, however, is no longer affected by the factor $\{\bar{F}(y_n)\}^{1/2}$ and improves to $n^{1/2}$, which seems a little surprising. An investigation of the proof, however, reveals that it is intuitively caused by the degeneracy of the Phase I asymptotic theory, in which case a faster scale has to be used to appropriately normalize the sample autocorrelation that leads to the improved convergence rate in Phase II. We shall in the following section consider implications of the developed results on obtaining a theoretically justifiable tail counterpart of the widely used ACF plot and on identifying the maximal lag of serial dependence as considered by Zhang (2005). An application to the moving-maximum process of Hall et al. (2002) is also presented to illustrate the developed results.

4. IMPLICATIONS ON STATISTICAL PRACTICE

4.1. *Tail Autocorrelation Plot: A Visualization Tool*

A crucial component of the traditional ACF plot is the two lines of significance that enable practitioners to make their decision statistically about whether the autocorrelation is zero and whether a model with finite-order dependence can be used for the data. The line of significance is obtained from a well-known asymptotic theory of sample autocorrelations when the data are independent, and the fact that the cut-off has a mathematically simple and statistically justifiable form contributes largely to the popularity and wide recognition of the ACF plot among practitioners and applied scientists from different disciplines. The task of developing a counterpart for tail autocorrelations, however, can be nontrivial due to the lack of an appropriate asymptotic theory. Davis et al. (2012) commented that, even for independent data, the cut-off for sample tail autocorrelations does not seem to be easily computable from any existing theory. We shall here use the results developed in Section 3 to fill this gap. In particular, for independent data, $E(\zeta_{0,n}\zeta_{l,n}\zeta_{|k|,n}\zeta_{l+|k|,n}) = 0$ holds for any $0 < |l| < |k|$, and as a result the quantity in (9) satisfies

$$\varsigma_{\mu,n}(k)^2 = \{F(y_n)\bar{F}(y_n)\}^{-2}E(\zeta_{0,n}^2\zeta_{|k|,n}^2) \to 1 \tag{10}$$

for $|k| > 0$. Therefore, by the Phase II central limit theorem in Corollary 2, we have

$$n^{1/2}\hat{\tau}_{n,y_n}(k) \to_d N(0,1)$$

for $|k| > 0$ when the underlying data are independent. This suggests that, although the convergence rate of sample tail autocorrelations can be slowed down to $\{n\bar{F}(y_n)\}^{1/2}$ with a nontrivial asymptotic variance in Phase I, for the purpose of constructing a counterpart of the widely recognized ACF plot in the tail setting, the lines of significance can actually be set neatly at the vertical position $\pm n^{-1/2}\Phi^{-1}(1-\alpha/2)$ according to our Phase II asymptotic theory, where $\Phi(\cdot)$ is the distribution function of a standard normal. We call the resulting visualization tool the tail autocorrelation function (TACF) plot, which we believe can have a great potential in assisting practitioners and applied scientists for analyzing serial dependence in the tail region.

4.2. *Identifying the Maximal Lag of Tail Dependence*

In addition to the above approach that finds the lines of significance by applying the Phase II asymptotic theory to independent data analogously to the traditional ACF plot, we consider here an alternative that computes the cut-off by treating the previous lag as the maximal lag of serial tail dependence. This relates to the problem of identifying the maximal lag of serial tail dependence as considered by Zhang (2005), where one seeks a test for $H_0 : \tau_{y_n}(k) = 0$ under (8) with $K = k-1$, namely when $k-1$ is the maximal lag of dependence; see also Zhang (2008). The gamma test of Zhang (2005) and Zhang (2008) assumes that $y_n = y$ is fixed, and we shall here consider the double asymptotic scheme when $y_n \uparrow \mathcal{U}_F$ to better accommodate the tail setting. For this, by the Phase II asymptotic theory in Corollary 2, we propose to reject the null hypothesis at level α if

$$|\hat{\tau}_{n,y_n}(k)| > n^{-1/2}\varsigma_{\mu,n}(k)\Phi^{-1}(1-\alpha/2), \tag{11}$$

where $\varsigma_{\mu,n}(k)^2$ is given in (9) with $K = |k|-1$. Compared with the independence setting where $\varsigma_{\mu,n}(k)^2 \to 1$ is known from (10), in the current setup $\varsigma_{\mu,n}(k)^2$ is typically unknown and needs to be estimated. For this, we propose to plug in its empirical version

$$\hat{\varsigma}_{\mu,n}(k)^2 = 1 + \frac{2}{n-|k|}\sum_{l=1}^{|k|-1}\sum_{i=1}^{n-l-|k|}\frac{\hat{\zeta}_{i,n}\hat{\zeta}_{i+l,n}\hat{\zeta}_{i+|k|,n}\hat{\zeta}_{i+l+|k|,n}}{[\hat{\bar{F}}(y_n)\{1-\hat{\bar{F}}(y_n)\}]^2}, \quad \hat{\zeta}_{i,n} = \mathrm{I}(Y_i > y_n) - \hat{\bar{F}}_n(y_n).$$

Therefore, if one seeks a statistical tool that provides a counterpart of the traditional ACF plot in the tail setting, then the clean but effective $\pm n^{-1/2}\Phi^{-1}(1-\alpha/2)$ lines of significance proposed in Section 4.1 can be more desirable. However, if one is interested in identifying the maximal lag of tail dependence as in the setting of Zhang (2005), then the decision rule in (11) can be more suitable, and thus we also include it here as a valuable alternative for practitioners to choose.

4.3. *Application to the Moving-Maximum Process*

We shall in this section use the moving-maximum process of Hall et al. (2002) to illustrate our results, including the tail adversarial stability condition and the discovered phase transition phenomenon of sample tail autocorrelations. As commented by Hall et al. (2002), the moving-maximum model encompasses a range of stochastic processes that are of interest in the context of extreme-value data, and in the same paper it was shown to be dense in the class of stationary processes whose finite-dimensional distributions are extreme-value of a given type; see also Zhang & Smith (2004), Zhang (2005) and Zhang et al. (2017) for additional discussions. Let $\epsilon_j,\ j \in \mathbb{Z}$, be independent Fréchet random variables with distribution function $F_\epsilon(z) = \mathrm{pr}(\epsilon_j \le z) = \exp(-z^{-\gamma})$ for some $\gamma > 0$, we consider the moving-maximum process

$$Y_i = \max_{0 \le l < \infty} a_l \epsilon_{i-l}, \quad i = 1, \ldots, n, \tag{12}$$

which is well defined if the nonnegative coefficients satisfy $\sum_{l=0}^{\infty} a_l^{\gamma} < \infty$; see the discussion in Section 2.2 of Hall et al. (2002). Zhang (2005) required a similar summability condition to define the M3 process with unit Fréchet innovations. We shall first illustrate the meaning of the tail adversarial q-stability condition for the moving-maximum process (12). For this, by the results in Hall et al. (2002), one can show that the joint probability

$$\mathrm{pr}(Y_i^\star \le y,\ Y_i > y) = \{1 - \exp(-a_i^{\gamma} y^{-\gamma})\} \cdot \exp\left(-\sum_{l=0}^{\infty} a_l^{\gamma} y^{-\gamma}\right) \le 1 - \exp(-a_i^{\gamma} y^{-\gamma}),$$

and as a result

$$\mathrm{pr}(Y_i^\star \le y \mid Y_i > y) \le \frac{1 - \exp(-a_i^{\gamma} y^{-\gamma})}{1 - \exp(-\sum_{l=0}^{\infty} a_l^{\gamma} y^{-\gamma})} \le \frac{2 a_i^{\gamma}}{\sum_{l=0}^{\infty} a_l^{\gamma}}$$

holds if $\mathrm{pr}(Y_i > y) \le 1/2$. Therefore, when $y \uparrow \mathcal{U}_F$, we have

$$\lim_{y \uparrow \mathcal{U}_F} \Theta_{y,q}(0) \le \frac{2 \sum_{i=0}^{\infty} a_i^{\gamma/q}}{\sum_{l=0}^{\infty} a_l^{\gamma}},$$

and the tail adversarial q-stability condition is satisfied if $\sum_{l=0}^{\infty} a_l^{\gamma} > 0$ and $\sum_{i=0}^{\infty} a_i^{\gamma/q} < \infty$. The first condition $\sum_{l=0}^{\infty} a_l^{\gamma} > 0$ essentially prevents the process from being degenerate; see for example Zhang (2005). The second condition $\sum_{i=0}^{\infty} a_i^{\gamma/q} < \infty$ controls the degree of tail dependence, and compared to the existence condition $\sum_{i=0}^{\infty} a_i^{\gamma} < \infty$ under which the moving-maximum process is well defined (Hall et al., 2002), it seems to be reasonable and mild. Calculating the strong mixing coefficient or the β-mixing coefficient for the moving-maximum process (12) can be a nontrivial task (Dombry & Eyi-Minko, 2012) and may possibly lead to stronger conditions on the coefficients, let alone the additional technical conditions often needed under the strong mixing framework that involve the interplay between how fast the mixing coefficient decays to zero and how far the tail can reach.

To illustrate the phase transition phenomenon, for simplicity, we consider the special case when $a_0 = 1$, $a_1 = 1/2$ and $a_l = 0$ for $l \ge 2$, namely when

$$Y_i = \max(\epsilon_i, \epsilon_{i-1}/2), \quad i = 1, \ldots, n. \tag{13}$$

In this case, the phase transition will occur at lag $k = 2$. To be more specific, when $|k| < 2$, one can show that the quantity $\sigma_{\mu,n}(k)^2$ defined in (6) satisfies

$$\lim_{n \to \infty} \sigma_{\mu,n}(0)^2 = \frac{2^{\gamma} + 3}{2^{\gamma} + 1}, \quad \lim_{n \to \infty} \sigma_{\mu,n}(1)^2 = \frac{1}{2^{\gamma} + 1}.$$

Since both of them are bounded away from zero, the Phase I asymptotic theory in Theorem 1 applies, and we have

$$\left\{\frac{n}{\bar{F}(y_n)}\right\}^{1/2}\{\hat{\mu}_{n,y_n}(0)-\bar{F}(y_n)F(y_n)\}\to_d N\left(0,\frac{2^\gamma+3}{2^\gamma+1}\right), \tag{14}$$

and

$$\left\{\frac{n}{\bar{F}(y_n)}\right\}^{1/2}\{\hat{\mu}_{n,y_n}(1)-\tau_{y_n}(1)\bar{F}(y_n)F(y_n)\}\to_d N\left(0,\frac{1}{2^\gamma+1}\right).$$

When the lag moves to $k=2$, however, the quantity $\sigma_{\mu,n}(k)^2$ defined in (6) now degenerates to zero, making the Phase I asymptotic theory no longer applicable. In this case, we shall use the Phase II asymptotic theory in Theorem 2, by which we can obtain that

$$\left[\frac{n}{\{\bar{F}(y_n)\}^2}\right]^{1/2}\hat{\mu}_{n,y_n}(2)\to_d N\left\{0,1+\frac{2}{(2^\gamma+1)^2}\right\}.$$

The convergence rate in this case is faster by a factor of $\{\bar{F}(y_n)\}^{-1/2}$ than the Phase I rate of $n^{1/2}\{\bar{F}(y_n)\}^{-1/2}$. Since $\hat{\mu}_{n,y_n}(0)=\bar{F}(y_n)F(y_n)\{1+o_p(1)\}$ from (14) and $F(y_n)\to 1$ as $n\to\infty$, by an application of Slutsky's theorem we have the Phase II central limit theorem for sample autocorrelations

$$n^{1/2}\hat{\tau}_{n,y_n}(2)\to_d N\left\{0,1+\frac{2}{(2^\gamma+1)^2}\right\},$$

which can be useful in testing if $k=2$ is the first lag of tail independence; see the discussion in Section 4.2 and Zhang (2005).

We also include here a small simulation study to illustrate the TACF plot in Section 4.1 that visualizes tail autocorrelations and the test described in Section 4.2 for determining the maximal lag of tail dependence. For this, we generate data according to the moving-maximum process (13) and provide sample TACF plots in Figure 1 for different choices of sample size $n\in\{500,1500\}$ and shape parameter $\gamma\in\{1,2\}$ where y_n is chosen as the 90% quantile. It can be seen from Figure 1 that a first-order tail dependence model indeed seems to be plausible for the generated data, and therefore the TACF plots can provide a useful visualization tool to practitioners for investigating tail dependence as a starting point. We also consider applying the test in Section 4.2 to the moving-maximum process (13), and the results are summarized in Table 1 based on 5000 realizations for each configuration. We can observe the followings from Table 1. First, the case when $k\geq 2$ relates to the null, for which the empirical sizes are reasonably close to their nominal levels of 90% and 95%. On the other hand, the case when $k=1$ relates to the alternative, for which the empirical sizes are generally close to zero indicating a reasonably good power performance of the test. Third, when we increase y_n from the 90% quantile to the 95% quantile, we are focusing on events that are more extremal, which can typically lead to larger size distortions and less power of the test. However, if we increase the sample size, then the size distortion gets smaller and the power increases, which corroborates our asymptotic theory. Additional simulation results on seasonal moving maximum processes are provided in the supplementary material, where similar observations can be made.

4.4. *An Application to Tropical Cyclone Data*

We shall here apply the developed results to assess tail dependence of a tropical cyclone data. The data contains satellite-derived lifetime-maximum wind speeds of 2098 tropical cyclones

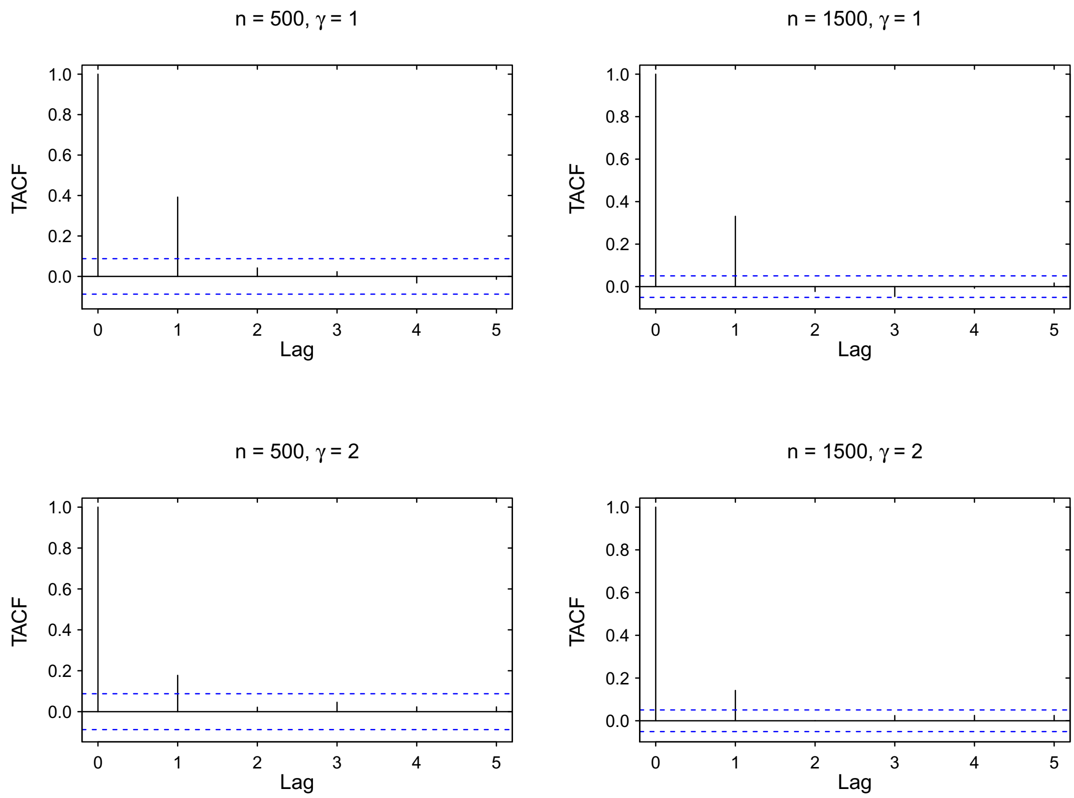

Fig. 1. Sample TACF plots for the moving-maximum process (13) with sample size $n \in \{500, 1500\}$ and shape parameter $\gamma \in \{1, 2\}$ where y_n is chosen as the 90% quantile. In all the plots, the blue dashed lines represent the 95% lines of significance based on the developed asymptotic theory as discussed in Section 4.1.

over the globe during 1981–2006, and we refer to Elsner et al. (2008) for a more detailed description. The aforementioned paper fitted linear trends to quantiles by assuming that the errors in the quantile regression model are independent and identically distributed; see also Zhou (2010), Zhang & Wu (2011) and Zhang & Lavitas (2018) for additional trend analyses. Figure 2 provides the time series plot of the data, and we shall here focus on examining the tail dependence. For this, we first follow Elsner et al. (2008) and fit a linear trend to the 95% quantile, and then apply the TACF visualization tool introduced in Section 4.1 to the detrended time series. The resulting TACF plot is provided in Figure 3, from which we can see that there seems to be a nonnegligible lag-12 tail dependence, which also seems to serve as the maximal lag of tail dependence. By the test introduced in Section 4.2, we can obtain a p-value of 0.000 for the lag-12 tail autocorrelation, indicating strong statistical evidence for a lag-12 tail dependence. If we rely on the traditional autocorrelation to examine the dependence, then the lag-12 traditional autocorrelation will be 0.025 which falls within the 95% lines of significance $\pm 1.96 n^{-1/2} = \pm 0.043$ in the usual ACF plot. Therefore, it seems desirable to develop convenient and rigorous statistical tools for visualizing and assessing tail dependence as in the current paper, which can be different from the usual

		$k=1$		$k=2$		$k=3$	
n	γ	90%	95%	90%	95%	90%	95%
				y_n chosen as the 90% quantile			
500	1	0.000	0.000	0.913	0.965	0.904	0.960
	2	0.019	0.033	0.911	0.961	0.917	0.964
1500	1	0.000	0.000	0.900	0.951	0.905	0.952
	2	0.000	0.000	0.904	0.958	0.897	0.947
				y_n chosen as the 95% quantile			
500	1	0.001	0.001	0.950	0.973	0.947	0.971
	2	0.060	0.087	0.936	0.965	0.936	0.960
1500	1	0.000	0.000	0.915	0.968	0.913	0.966
	2	0.000	0.000	0.914	0.967	0.909	0.963

Table 1. *Empirical size of the test described in Section* 4.2 *for determining the maximal lag of tail dependence.*

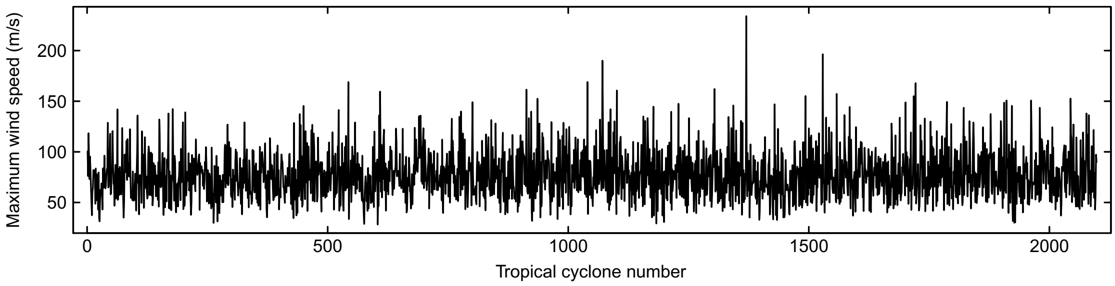

Fig. 2. Satellite-derived lifetime-maximum wind speeds of tropical cyclones during 1981–2006.

concept of dependence that concerns comovements around the mean. Our finding also indicates that the analysis of Elsner et al. (2008) conducted by assuming independence can be revisited and improved as a future topic in climate science by taking into account the discovered significant lag-12 tail dependence.

5. DISCUSSION

We develop an asymptotic theory on sample tail autocorrelations for time series data that can exhibit serial dependence in both tail and non-tail regions. Compared with the strong mixing framework of Rosenblatt (1956) which often has to be used along with additional technical conditions that can largely limit the degree of tail dependence (Chernozhukov, 2005; Chernozhukov & Fernández-Val, 2011) or cause additional inexplicit restrictions on how extremal the tail can be through an interplay with how fast the mixing coefficient decays to zero (Davis & Mikosch, 2009), our results are developed under a tail adversarial stability framework which does not impose any restriction on the dependence structure in non-tail regions and allows processes that are not necessarily strong mixing or regularly varying. It also leads to more explicit and less restrictive conditions on the tail and covers more extremal tails than existing ones; see the discussion in Section 3.2. In addition, our newly developed asymptotic theory reveals a previously undiscovered phase transition phenomenon for sample tail autocorrelations, where we find that the

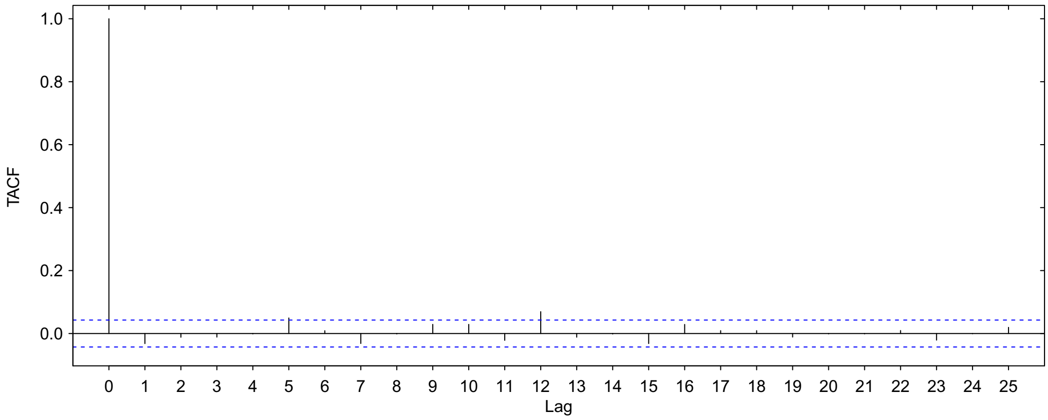

Fig. 3. Sample TACF plot of the tropical cyclone data after removing a linear trend to the 95% quantile. The blue dashed lines represent the 95% lines of significance based on the developed asymptotic theory as discussed in Section 4.1.

asymptotic behavior of sample tail autocorrelations including their convergence rate can transit from one phase to the other when the lag index moves past the point beyond which serial tail dependence vanishes. Such a phase transition phenomenon does not exist in traditional autocorrelations where the convergence rate is the universal square root of the sample size, and is a characteristic product of the double asymptotic scheme in the current setting that is necessary for studying tail phenomena. To illustrate the developed results, we in Section 4 consider the problem of developing a theoretically justified analogue of the traditional ACF plot in the tail setting and the problem of identifying the maximal lag of tail dependence as posed in Zhang (2005). For both problems, existing central limit theorems on sample tail autocorrelations can degenerate and become noninformative for inference, while the two-phase asymptotic theory developed in the current paper fills this gap and provides an essential foundation for finding statistical solutions to the aforementioned problems.

Acknowledgement

I am grateful to the editor, the associate editor and two anonymous referees for their helpful comments and suggestions. The research is supported by the NSF CAREER Award DMS-1848035/DMS-2131821.

Supplementary Material

Supplementary material includes additional simulation results and technical proofs of the results in Section 3.

References

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

Basrak, B. & Segers, J. (2009). Regularly varying multivariate time series. *Stochastic Processes and their Applications* **119**, 1055–1080.

CHERNOZHUKOV, V. (2005). Extremal quantile regression. *The Annals of Statistics* **33**, 806–839.

CHERNOZHUKOV, V. & FERNÁNDEZ-VAL, I. (2011). Inference for extremal conditional quantile models, with an application to market and birthweight risks. *The Review of Economic Studies* **78**, 559–589.

COLES, S., HEFFERNAN, J. & TAWN, J. (1999). Dependence measures for extreme value analyses. *Extremes* **2**, 339–365.

DAVIS, R. A., DREES, H., SEGERS, J. & WARCHOŁ, M. (2018). Inference on the tail process with application to financial time series modeling. *Journal of Econometrics* **205**, 508–525.

DAVIS, R. A. & MIKOSCH, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli* **15**, 977–1009.

DAVIS, R. A., MIKOSCH, T. & CRIBBEN, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. *Journal of Econometrics* **170**, 142–152.

DAVIS, R. A., MIKOSCH, T. & ZHAO, Y. (2013). Measures of serial extremal dependence and their estimation. *Stochastic Processes and their Applications* **123**, 2575–2602.

DE HAAN, L. & RESNICK, S. I. (1977). Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **40**, 317–337.

DOMBRY, C. & 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. & DE HAAN, L. (2004). Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli* **10**, 251–280.

DREES, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **9**, 617–657.

DREES, H. & ROOTZÉN, H. (2010). Limit theorems for empirical processes of cluster functionals. *The Annals of Statistics* **38**, 2145–2186.

ELSNER, J. B., KOSSIN, J. P. & JAGGER, T. H. (2008). The increasing intensity of the strongest tropical cyclones. *Nature* **455**, 92–95.

EMBRECHTS, P., MCNEIL, A. & STRAUMANN, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. In *Risk Management: Value at Risk and Beyond*, M. A. H. Dempster, ed. Cambridge University Press, pp. 176–223.

FERRO, C. A. T. & SEGERS, J. (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65**, 545–556.

HALL, P., PENG, L. & YAO, Q. (2002). Moving-maximum models for extrema of time series. *Journal of Statistical Planning and Inference* **103**, 51–63.

HAN, H., LINTON, O., OKA, T. & WHANG, Y.-J. (2016). The cross-quantilogram: Measuring quantile dependence and testing directional predictability between time series. *Journal of Econometrics* **193**, 251–270.

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.

KULIK, R., SOULIER, P. & WINTENBERGER, O. (2019). The tail empirical process of regularly varying functions of geometrically ergodic markov chains. *Stochastic Processes and their Applications* **129**, 4209–4238.

LEADBETTER, M. R., LINDGREN, G. & ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Berlin: Springer.

LEDFORD, A. W. & TAWN, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169–187.

LEDFORD, A. W. & TAWN, J. A. (2003). Diagnostics for dependence within time series extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65**, 521–543.

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

MCNEIL, A. J., FREY, R. & EMBRECHTS, P. (2005). *Quantitative Risk Management–Concept, Techniques and Tools*. Princeton, NJ: Princeton University Press.

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

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.

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

SMITH, R. L. & WEISSMAN, I. (1994). Estimating the extremal index. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **56**, 515–528.

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

TSAY, R. S. (2010). *Analysis of financial time series*. Hoboken, New Jersey: John Wiley & Sons, 3rd ed.

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

WU, W. B. (2011). Asymptotic theory for stationary processes. *Statistics and Its Interface* **4**, 207–226.

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

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

ZHANG, T. & LAVITAS, L. (2018). Unsupervised self-normalized change-point testing for time series. *Journal of the American Statistical Association* **113**, 637–648.

ZHANG, T. & 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. & HUANG, J. (2006). Extremal financial risk models and portfolio evaluation. *Computational Statistics & Data Analysis* **51**, 2313–2338.

ZHANG, Z. & SMITH, R. L. (2004). The behavior of multivariate maxima of moving maxima processes. *Journal of Applied Probability* **41**, 1113–1123.

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

ZHOU, Z. (2010). Nonparametric inference of quantile curves for nonstationary time series. *The Annals of Statistics* **38**, 2187–2217.