

## Estimation of the marginal expected shortfall under asymptotic independence

Cai, Juanjuan; Musta, Eni

**DOI**

[10.1111/sjos.12397](https://doi.org/10.1111/sjos.12397)

**Publication date**

2019

**Document Version**

Final published version

**Published in**

Scandinavian Journal of Statistics

**Citation (APA)**

Cai, J., & Musta, E. (2019). Estimation of the marginal expected shortfall under asymptotic independence. *Scandinavian Journal of Statistics*, 47 (2020)(1), 56-83. <https://doi.org/10.1111/sjos.12397>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.


**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Estimation of the marginal expected shortfall under asymptotic independence

Juan-Juan Cai | Eni Musta 

Delft Institute of Applied  
Mathematics, Delft University of  
Technology, Delft, The Netherlands

## Correspondence

Eni Musta, Delft Institute of Applied  
Mathematics, Delft University of  
Technology, Delft 2628 XE,  
The Netherlands.  
Email: e.musta@tudelft.nl

## Abstract

We study the asymptotic behavior of the marginal expected shortfall when the two random variables are asymptotic independent but positively associated, which is modeled by the so-called tail dependent coefficient. We construct an estimator of the marginal expected shortfall, which is shown to be asymptotically normal. The finite sample performance of the estimator is investigated in a small simulation study. The method is also applied to estimate the expected amount of rainfall at a weather station given that there is a once every 100 years rainfall at another weather station nearby.

## KEYWORDS

asymptotic independence, marginal expected shortfall, tail dependence coefficient

## 1 | INTRODUCTION

Let  $X$  and  $Y$  denote two risk factors. The marginal expected shortfall (MES) is defined as  $\mathbb{E}[X | Y > Q_Y(1 - p)]$ , where  $Q_Y$  is the quantile function of  $Y$  and  $p$  is a *small* probability. The name MES originates from its application in finance as an important ingredient for constructing a systemic risk measure; see for instance Acharya, Pedersen, Philippon, and Richardson (2017) and Caporin and Santucci de Magistris (2012). In actuarial science, this quantity is known as the multivariate extensions of tail conditional expectation (or conditional tail expectation); see for instance Cai and Li (2005) and Cousin and Di Bernardino (2014).

Under the assumption that  $X$  is in the Fréchet domain of attraction, Cai, Einmahl, de Haan, and Zhou (2015) have established the following asymptotic limit (see Proposition 1 in that paper). With  $Q_X$ , the quantile function of  $X$ ,

$$\lim_{p \rightarrow 0} \frac{\mathbb{E}[X | Y > Q_Y(1 - p)]}{Q_X(1 - p)} = a \in [0, \infty), \quad (1)$$

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors Scandinavian Journal of Statistics published by John Wiley & Sons Ltd on behalf of The Board of the Foundation of the Scandinavian Journal of Statistics

where  $a > 0$  if  $X$  and  $Y$  are asymptotically dependent and  $a = 0$  if they are asymptotically independent. Based on this result, an estimator for MES is established in Cai et al. (2015), which is not applicable for asymptotically independent data. It is the goal of this paper to study the asymptotic behavior of MES and to develop an estimation of MES for asymptotically independent data.

Under the framework of multivariate extreme value theory, there are various ways to describe asymptotic dependence, for instance, by means of exponent measure, spectral measure, or Pickands dependence functions (cf. de Haan & Ferreira, 2006, Chapter 6; Beirlant, Goegebeur, Segers, & Teugels, 2004, Chapter 8). However, these measures do not distinguish the relative strength of the extremal dependence for an asymptotically independent pair. The so-called coefficient of tail dependence introduced by Ledford and Tawn (1996) is mostly used to measure the strength of the extremal dependence for an asymptotically independent pair. In this paper, we make use of the coefficient of tail dependence, denoted as  $\eta$  to model asymptotic independence. Namely, we assume that there exists an  $\eta \in (0, 1]$  such that the following limit exists and is positive:

$$\lim_{p \rightarrow 0} p^{-\frac{1}{\eta}} \mathbb{P}(X > Q_X(1-p) \text{ and } Y > Q_Y(1-p)).$$

We are interested in the scenario that  $\eta \in (1/2, 1)$ , which corresponds to asymptotic independence but positive association of  $X$  and  $Y$ . For this type of distributions, for  $p$  close to zero, one has

$$\mathbb{P}(X > Q_X(1-p) \text{ and } Y > Q_Y(1-p)) \gg \mathbb{P}(X > Q_X(1-p)) \mathbb{P}(Y > Q_Y(1-p)),$$

that is, the joint extremes of  $(X, Y)$  happen much more often than those of a distribution with independent components of  $X$  and  $Y$ . This gives an intuitive explanation that, even if the pair are asymptotically independent, the extremal dependence can still be strong and thus needs to be accounted for. We also assume that  $X$  is in the Fréchet domain of attraction, so it has a heavy right tail. As our result shall reveal, the risk represented by MES can also be very big under the combination of positive association and  $X$  being heavy tailed (cf. Proposition 1). Thus, from the application point of view, it is very important to consider MES for such a model assumption.

The study on MES for an asymptotically independent pair has attracted increasing attention in recent literature. Various approaches have been used to model asymptotic independence and none of them is superior to the others. Das and Fasen-Hartmann (2018) have constructed an estimator of MES and have shown the consistency of the estimator, assuming that the distribution of the pair  $(X, Y)$  possesses hidden regular variation, which is a similar setting with ours. However, from the modeling point view, Das and Fasen-Hartmann (2018) also assume a heavy right tail for the distribution of the conditioning variable, that is,  $Y$  in our notation, which also results in a different estimator of MES than ours. Apart from the continuity, we do not impose any condition on the marginal distribution of  $Y$ . Kulik and Soulier (2015) have provided the asymptotic property of MES for regularly varying time series  $\{X_h\}$  with extremal independence (meaning that  $X_0$  and  $X_h$  are asymptotically independent for  $h > 0$ ), using conditional extreme value approach, which is an alternative way to model asymptotic independence, introduced in Heffernan and Resnick (2007) and Heffernan and Tawn (2004). No estimators are yet proposed. In view of these existing papers, our contribution is that, first, the asymptotic property of MES is obtained by using only the coefficient of tail dependence to model the asymptotic independence; in this way, only the strength of the extremal dependence between  $X$  and  $Y$  matters and the marginal distribution of  $Y$  does not play a role in the modeling part and, thus, in the estimation part as well, and second, we prove the asymptotic normality for the constructed estimator of MES.

This paper is organized as follows. Section 2 contains the main theoretical results on the limit behavior of MES and the asymptotic normality of the proposed estimator of MES. The

performance of the estimation method is illustrated by a simulation study in Section 3 and by an application to precipitation data in Section 4. The proofs of the main theorems are provided in Section 5.

## 2 | MAIN RESULTS

We first derive the asymptotic limit for MES as  $p \rightarrow 0$ , based on which we shall then construct an estimator for MES. Let  $F_1$  and  $F_2$  denote the marginal distribution functions of  $X$  and  $Y$ , respectively. Assume that the marginal distributions are continuous. As usual in extreme value analysis, it is more convenient to work with, instead of the quantile function, the tail quantile defined as  $U_j = \left(\frac{1}{1-F_j}\right)^{\leftarrow}$ ,  $j = 1, 2$ , where  $\leftarrow$  denotes the left continuous inverse. Then, MES can be written as

$$\mathbb{E}[X | Y > U_2(1/p)] =: \theta_p.$$

We now present our model, namely, assumptions on the tail distribution of  $X$  and the extremal dependence of  $X$  and  $Y$ . First, we assume that  $X$  has a heavy right tail, that is, there exists a  $\gamma_1 > 0$  such that

$$\lim_{t \rightarrow \infty} \frac{U_1(tx)}{U_1(t)} = x^{\gamma_1}, \quad x > 0. \quad (2)$$

$\gamma_1$  is the so-called extreme value index of the distribution of  $X$ . In this paper, we only consider the case that  $\gamma_1 < 1$  to guarantee that  $\mathbb{E}[X]$  is finite.

Second, we assume the positive association of  $X$  and  $Y$ . Precisely, there exists an  $\eta \in (1/2, 1]$  such that, for all  $(x, y) \in (0, \infty)^2$ , the following limit exists:

$$\lim_{t \rightarrow \infty} t^{\frac{1}{\eta}} \mathbb{P}(X > U_1(t/x), Y > U_2(t/y)) =: c(x, y) \in (0, \infty). \quad (3)$$

As a consequence,  $c$  is a homogeneous function of order  $\frac{1}{\eta}$ . The condition of (3) is also assumed in Draisma, Drees, Ferreira, and de Haan (2004) for estimating  $\eta$ , and several examples of models satisfying (3) are given in Ledford and Tawn (1997). Note that if  $\eta = 1$ , it corresponds to  $X$  and  $Y$  being asymptotically dependent. For  $\eta < 1$ , this condition is linked to the so-called hidden regular variation (cf. Resnick, 2002) in the following way, provided that the tails of  $X$  and  $Y$  are equivalent:

$$v^*((x, \infty] \times (y, \infty]) = c(x^{-1/\gamma_1}, y^{-1/\gamma_1}),$$

where  $v^*$  is defined in (3) of Resnick (2002).

In order to obtain the limiting result on  $\theta_p$  for  $p \rightarrow 0$ , we need a second-order strengthening condition of (2).

A(1) There exists  $d \in (0, \infty)$  such that

$$\lim_{t \rightarrow \infty} \frac{U_1(t)}{t^{\gamma_1}} = d.$$

We also need some technical conditions on the extremal dependence of  $X$  and  $Y$ . For  $t > 0$ , define

$$c_t(x, y) = t^{\frac{1}{\eta}} \mathbb{P}(X > U_1(t/x), Y > U_2(t/y)), \quad 0 < x, y < t. \quad (4)$$

A(2) There exists  $\beta_1 > \gamma_1$  such that  $\lim_{t \rightarrow \infty} \sup_{x \leq 1} |c_t(x, 1) - c(x, 1)| x^{-\beta_1} = 0$ .

A(3) There exists  $0 < \beta_2 < \gamma_1$  such that  $\lim_{t \rightarrow \infty} \sup_{1 < x < t} |c_t(x, 1) - c(x, 1)| x^{-\beta_2} = 0$ .

**Proposition 1.** Assume that  $X$  takes values in  $(0, \infty)$ ,  $\gamma_1 \in (0, 1)$  and conditions A(1)–A(3) hold. If  $-\frac{1}{\eta} + 1 + \gamma_1 > 0$ ,  $\int_0^\infty c\left(x^{-\frac{1}{\eta}}, 1\right) dx < \infty$  and  $x \mapsto c(x, 1)$  is a continuous function, then we have

$$\lim_{t \rightarrow \infty} \frac{\theta_{1/t}}{t^{-\frac{1}{\eta}+1} U_1(t)} = \int_0^\infty c\left(x^{-\frac{1}{\eta}}, 1\right) dx. \tag{5}$$

*Remark 1.* If  $\eta = 1$ , then the result of (5) coincides with the result of Proposition 1 in Cai et al. (2015). The condition of  $-\frac{1}{\eta} + 1 + \gamma_1 > 0$  implies that  $t^{-\frac{1}{\eta}+1} U_1(t) \rightarrow \infty$ ; thus,  $\theta_{1/t} \rightarrow \infty$  as  $t \rightarrow \infty$ .

Below, we discuss the main challenge of obtaining the limiting result in Proposition 1 and the technical assumptions A(1)–A(3).

*Remark 2.* As a conditional expectation, the MES can be written as a scaled integral as follows (cf. (13)):

$$\theta_{1/t} = t^{-\frac{1}{\eta}+1} U_1(t) \int_0^\infty t^{\frac{1}{\eta}} \mathbb{P}(X > x U_1(t), Y > U_2(t)) dx.$$

In the current setting ( $\eta < 1$ ), the main challenge is to validate the generalized dominated convergence condition for obtaining the integrability. For this purpose, we impose conditions A(1)–A(3). In Das and Fasen-Hartmann (2018), Assumption B was assumed to guarantee the integrability, and the conditions in Lemma 2.3 in Kulik and Soulier (2015) played a similar role. Note that if  $\eta = 1$ , the integrand is a conditional probability and it is thus easier to show the integrability (cf. Cai et al, 2015).

Given a random sample  $(X_1, Y_1), \dots, (X_n, Y_n)$ , we now construct an estimation of  $\theta_p$ , where  $p = p(n) \rightarrow 0$ , as  $n \rightarrow \infty$ . Proposition 1 suggests the following approximation. With  $t$  sufficiently large,

$$\theta_p \sim \left(\frac{1}{pt}\right)^{-\frac{1}{\eta}+1} \frac{U_1(1/p)}{U_1(t)} \theta_{\frac{1}{t}} \sim \left(\frac{1}{pt}\right)^{-\frac{1}{\eta}+1+\gamma_1} \theta_{\frac{1}{t}}.$$

We choose  $t = \frac{n}{k}$ , where  $k = k(n)$  is a sequence of integers such that  $k \rightarrow \infty$  and  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$ . Then,

$$\theta_p \sim \left(\frac{k}{pn}\right)^{-\frac{1}{\eta}+1+\gamma_1} \theta_{\frac{k}{n}}.$$

From this extrapolation relation, the remaining task is to estimate  $\eta$ ,  $\gamma_1$ , and  $\theta_{\frac{k}{n}}$ . There are well-known existing methods for estimating  $\gamma_1$  and  $\eta$ ; see Chapters 3 and 7 of de Haan and Ferreira (2006). For  $\theta_{\frac{k}{n}}$ , we propose a nonparametric estimator given by

$$\hat{\theta}_{k/n} = \frac{1}{k} \sum_{i=1}^n X_i \mathbb{1}_{\{Y_i > Y_{n-k,n}\}}. \tag{6}$$

Let  $\hat{\gamma}_1$  and  $\hat{\eta}$  denote estimators of  $\gamma_1$  and  $\eta$ , respectively. We construct the following estimator for  $\theta_p$ :

$$\hat{\theta}_p = \hat{\theta}_{k/n} \left(\frac{k}{np}\right)^{-\frac{1}{\hat{\eta}}+1+\hat{\gamma}_1}. \tag{7}$$

Next, we prove the asymptotic normality of  $\hat{\theta}_p$ . The following conditions will be needed.

B(1)  $-\frac{1}{\eta} + 1 + \gamma_1 > 0$  and there exists  $\bar{\delta} > 0$  such that

$$\left| \int_0^1 c(x, 2) dx^{-(2+\bar{\delta})\gamma_1} \right| < \infty, \quad \text{and} \quad \left| \int_1^\infty c(x, 2)^2 dx^{-\gamma_1} \right| < \infty.$$

B(2) There exists  $\beta_1 > (2 + \bar{\delta})\gamma_1$  and  $\tau < 0$  such that

$$\sup_{\substack{x \leq 1 \\ 1/2 \leq y \leq 2}} |c_t(x, y) - c(x, y)| x^{-\beta_1} = O(t^\tau).$$

B(3) There exists  $\beta_2 < \min \{-\tau/(1 - \gamma_1), \gamma_1\}$  such that

$$\sup_{\substack{1 < x < t \\ 1/2 \leq y \leq 2}} |c_t(x, y) - c(x, y)| x^{-\beta_2} = O(t^\tau)$$

with the same  $\tau$  as in B(2).

B(4) There exists  $\rho_1 < \frac{1}{2} - \frac{1}{2\eta}$  and a regularly varying function  $A_1$  with index  $\rho_1$  such that

$$\sup_{x>1} \left| x^{-\gamma_1} \frac{U_1(tx)}{U_1(t)} - 1 \right| = O\{A_1(t)\}.$$

B(5) As  $n \rightarrow \infty$ ,  $k = O(n^\alpha)$  for some  $\alpha$  that satisfies the following condition:

$$1 - \eta < \alpha < \min \left( 1 - \frac{\eta}{1 + \eta\gamma_1\lambda}, 1 + \frac{\eta}{1 - 2\eta - 2\eta\gamma_1\lambda}, \frac{-\frac{1}{\eta} + 1 + 2\tau + 2\beta_2(1 - \gamma_1)}{-\frac{1}{\eta} + 2\tau + 2\beta_2(1 - \gamma_1)}, 1 + \frac{1}{2\rho_1 - 1} \right)$$

with some  $\max \left( \beta_2, \frac{1-\eta}{\gamma_1\eta} \right) < \lambda < 1$ .

B(6)  $\hat{\gamma}_1$  is such that  $\sqrt{k}(\hat{\gamma}_1 - \gamma_1) = O_p(1)$ , and  $\hat{\eta}$  is such that  $\sqrt{k}(\hat{\eta} - \eta) = O_p(1)$ .

**Theorem 1.** Suppose that  $X$  takes values in  $(0, \infty)$ ,  $F_1$  is strictly increasing,  $\gamma_1 \in (0, 1)$ , and Conditions B(1)–B(6) hold. Assume that  $d_n = k/(n\rho) \geq 1$  and  $\lim_{n \rightarrow \infty} (n/k)^{1/2-1/(2\eta)} \log(d_n) = 0$ . Then, as  $n \rightarrow \infty$ ,

$$\left( \frac{n}{k} \right)^{-\frac{1}{2\eta} + \frac{1}{2}} \sqrt{k} \left( \frac{\hat{\theta}_p}{\theta_p} - 1 \right) \xrightarrow{d} N(0, \sigma^2),$$

where

$$\sigma^2 = - \int_0^\infty c(x, 1) dx^{-2\gamma_1} \left| \int_0^\infty c \left( x^{-\frac{1}{\gamma_1}}, 1 \right) dx \right|^{-2}.$$

*Remark 3.* Note that the condition  $\lim_{n \rightarrow \infty} (n/k)^{1/2-1/(2\eta)} \log(d_n) = 0$  implies that  $\eta < 1$ . Moreover, from  $\gamma_1 < 1$  and  $-\frac{1}{\eta} + 1 + \gamma_1 > 0$  (see B(1)), it follows that  $\eta > 1/2$ .

In the remarks below, we discuss the limitation of the model and the choice of  $k$ .

*Remark 4.* From Assumption B(1) and the monotonicity property of  $c(x, y)$ , for  $\rho \in \{1, 2, 2 + \bar{\delta}\}$ , we have

$$\sup_{y \in [1/2, 2]} \left| \int_0^\infty c(x, y) dx^{-\rho\gamma_1} \right| < \infty.$$

We will deal with such integral throughout the proof. This condition rules out models for  $c$ , which contain a power term of  $x$ , such as  $c(x, y) = (xy)^{1/\eta}$ . Assumptions B(2)–B(4) are technical conditions to obtain the rate for the convergence in Proposition 1, that is, to control the bias of the estimation. Assumption B(4), which by Theorem B.2.2 in de Haan and Ferreira (2006) implies Assumption A(1), is an ordinary second-order condition for  $F_1$ .

*Remark 5.* In order to obtain a nondegenerate limiting distribution in Theorem 1, a slower convergence rate is needed, compared to the usual rate of  $\sqrt{k}$ . Note that if  $\eta = 1$ , then the rate becomes  $\sqrt{k}$ . Consider  $k = O(n^\alpha)$  for  $\alpha \in (0, 1)$ . Assumption B(5) imposes both lower and upper bounds for  $\alpha$ . The upper bound of  $\alpha$  is a typical constraint in extreme value theory literature to ensure that the observations used in the estimation are from the tail. The lower bound is used to guarantee a proper convergence rate:  $(n/k)^{1/2-1/(2\eta)}\sqrt{k} \rightarrow \infty$  is implied by  $\alpha > 1 - \eta$ .

### 3 | SIMULATION STUDY

In this section, we study the finite sample performance of our method. We apply our estimator given by (7) to data of  $(X, Y)$  generated from the following two type of distributions. All the distributions of  $(X, Y)$  depend on two parameters:  $\alpha_1, \alpha_2 \in (0, 1)$ .

*Model 1.* Let  $Z_1, Z_2$ , and  $Z_3$  be independent Pareto random variables with parameters  $\alpha_1, \alpha_2$ , and  $\alpha_1$ , respectively. Here, a Pareto distribution with parameter  $\alpha$  means that the probability density function is given by  $f(x) = \frac{1}{\alpha}x^{-1/\alpha-1}$ . Define

$$(X, Y) = B(Z_1, Z_3) + (1 - B)(Z_2, Z_2),$$

where  $B$  is a Bernoulli  $(1/2)$  random variable independent of  $Z_i$ 's. For this model, we have  $\gamma_1 = \alpha_1$ ,  $\rho_1 = 1 - \alpha_1/\alpha_2$ ,  $\eta = \frac{\alpha_2}{\alpha_1}$ , and  $c(x, y) = 2^{\alpha_1/\alpha_2-1}(x \wedge y)^{\alpha_1/\alpha_2}$ . We consider four distributions from this model with parameters specified in Table 1.

*Model 2.* Define  $(X, Y) = ((1 - \Phi(\tilde{X}))^{-\alpha_1}, \tilde{Y})$ , where  $\tilde{X}$  and  $\tilde{Y}$  are two standard normal random variables with covariance  $\alpha_2$  and  $\Phi$  is the distribution function of  $\tilde{X}$ . Thus,  $X$  follows from a Pareto distribution with parameter  $\alpha_1$ , and  $(X, Y)$  has a Gaussian copula. For this model,  $\gamma_1 = \alpha_1$ ,  $\rho_1 = 0$ ,  $\eta = (1 + \alpha_2)/2$ , and  $c(x, y) = d(xy)^{2/(1+\alpha_2)}$ . As noted in Remark 4, distributions from this model do not satisfy our assumptions.

In terms of computing the theoretical value of  $\theta_p$ , basic calculation leads to the following closed form for distributions from Model 1:

$$\theta_p = \frac{1}{2p} \left( u_p^{-1/\alpha_1}(1 - \alpha_1)^{-1} + u_p^{(\alpha_2-1)/\alpha_2}(1 - \alpha_2)^{-1} \right), \tag{8}$$

where  $u_p$  is such that  $(u_p^{-1/\alpha_1} + u_p^{-1/\alpha_2})/2 = p$ , that is, the tail quantile of  $Y$ . For Model 2, there is no explicit form and  $\theta_p$  is obtained via numerical integral. Table 1 shows the parameters for five distributions and the values of  $\theta_p$ . Note that (5) does not hold for the last two distributions because  $-\frac{1}{\eta} + 1 + \gamma_1 < 0$  for Model 1(d) and  $\int_0^\infty c\left(x^{-\frac{1}{\eta}}, 1\right) dx = \infty$  for Model 2(a). For Model 1(d),  $\theta_p \rightarrow 2$  as  $p \rightarrow 0$  by (8) and  $\theta_p \rightarrow \infty$  as  $p \rightarrow 0$  for Model 2(a).

**TABLE 1** Parameters of five distributions and values of  $\theta_p$

Distribution	$(\alpha_1, \alpha_2)$	$\gamma_1$	$\eta$	$-\frac{1}{\eta} + 1 + \gamma_1$	$\theta_{p_1}$	$\theta_{p_2}$	$\theta_{p_3}$
Model 1(a)	(0.4, 0.3)	0.4	0.75	0.067	3.26	3.81	4.33
Model 1(b)	(0.4, 0.35)	0.4	0.875	0.257	6.04	10.66	19.20
Model 1(c)	(0.6, 0.4)	0.6	$\frac{2}{3}$	0.1	5.10	6.03	7.05
Model 1(d)	(0.5, 0.3)	0.5	0.6	-0.166	2.50	2.37	2.26
Model 2(a)	(0.4, 0.5)	0.4	0.75	0.067	4.23	6.54	9.78

As for the three “good” distributions, we have checked that Model 1(b) satisfies all the Conditions A(1)–A(3) and B(1)–B(5), whereas Model 1(a) and Model 1(c) satisfy all the conditions except B(5). Note that B(5) is a technical condition about the convergence speed of the asymptotic normality. From the simulation results presented below, one can see that Model 1(a) and Model 1(b) lead to very similar results, which are slightly better than Model 1(c). This suggests that the combination of all the technical conditions is rather restrictive; however, the estimation works for a wider range of models.

We consider sample size  $n = 5000$ , and  $p_i = 10^{2-i}/n$ , for  $i = 1, 2, 3$ . To complete our estimator given by (7), we use the Hill estimator for  $\gamma_1$  and an estimator for  $\eta$  proposed in Draisma et al. (2004). Let  $k_1$  and  $k_2$  be two intermediate sequences. Define

$$\hat{\gamma}_1 = \frac{1}{k_1} \sum_{i=1}^{k_1} \log(X_{n-i+1,n}) - \log(X_{n-k_1,n}),$$

and

$$\hat{\eta} = \frac{1}{k_2} \sum_{i=1}^{k_2} \log \frac{T_{n,n-i+1}^{(n)}}{T_{n,n-k_2}^{(n)}}, \quad (9)$$

where  $T_i^{(n)} = \frac{n+1}{n+1-R_i^X} \wedge \frac{n+1}{n+1-R_i^Y}$  with  $R_i^X$  and  $R_i^Y$  denoting the ranks of  $X_i$  and  $Y_i$  in their respective samples. These proposed estimators  $\hat{\gamma}_1$  and  $\hat{\eta}$  satisfy Condition B(6).

For  $p = 10/n$ , we compare our estimator with the nonparametric estimator, namely,

$$\hat{\theta}_{\text{emp}} = \frac{1}{10} \sum_{i=1}^n X_i \mathbb{1}_{\{Y_i > Y_{n-10,n}\}},$$

which is obtained by letting  $k/n = p$  in (6).

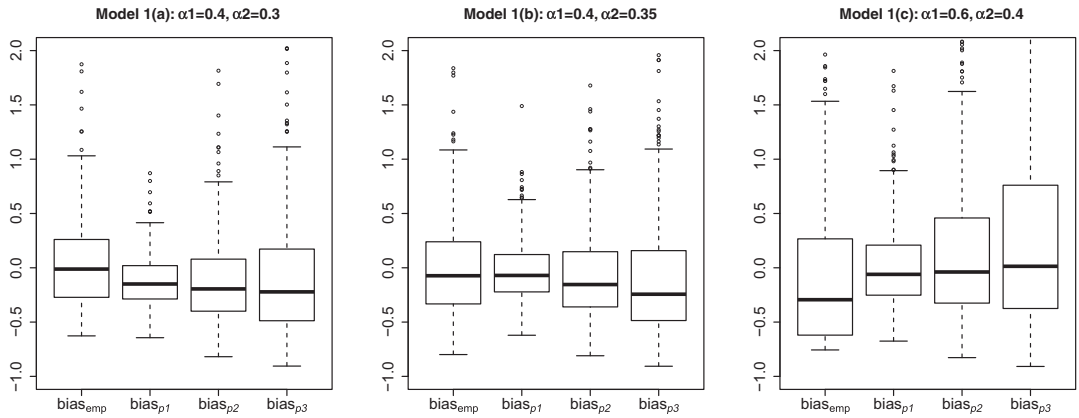
For each estimator, we compute the relative error defined as  $\text{bias}_p = \frac{1}{m} \sum_{i=1}^m \frac{\hat{\theta}_{p,i}}{\theta_p} - 1$ , where  $\hat{\theta}_{p,i}$  is an estimate based on the  $i$ th sample. A relative error for  $\hat{\theta}_{\text{emp}}$  is computed in the same way, denoted as  $\text{bias}_{\text{emp}}$ . For each scenario, the relative error is obtained by generating  $m = 500$  samples and choosing  $k = k_1 = k_2 = 200$ . Figure 1 shows the results for the first three distributions in Table 1. From the box plots, for the situation where the empirical estimator is applicable, that is,  $p = 10/n$ , our estimator has a smaller variance and similar or even smaller bias. As  $p$  becomes smaller, the empirical estimator is not applicable, yet our estimator still has decent performance with growing variance. Figure 2 shows the results for the last two distributions in Table 1, which do not satisfy our model assumptions. Our method underestimates the MES  $\theta_p$  for all the cases, which suggests that two conditions, namely,  $-\frac{1}{\eta} + 1 + \gamma_1 > 0$  and  $\int_0^\infty c\left(x^{-\frac{1}{\eta}}, 1\right) dx < \infty$ , are necessary for the estimation to work.

The proper choice of  $k$ 's, that is, the number of tail observations used in the estimation is always a delicate problem in the extreme value theory. To investigate how sensitive our result is with respect to the choice of the  $k$ 's and to see the range of suitable  $k$ 's, we compute for Model 1(a) and Model 1(b), the scaled mean squared errors:

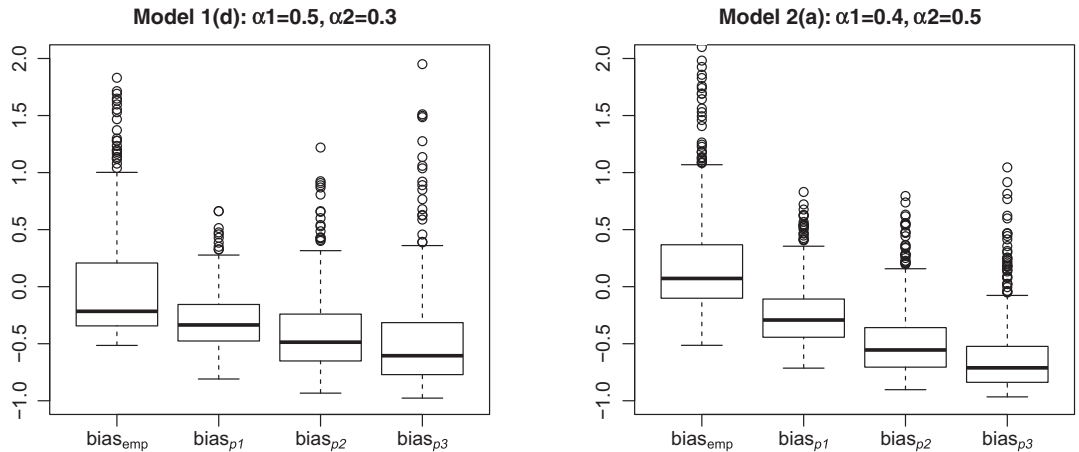
$$sMSE(k, k_1, k_2) = \frac{1}{m} \sum_{i=1}^m \left( \frac{\hat{\theta}_{p,i}(k, k_1, k_2)}{\theta_p} - 1 \right)^2. \quad (10)$$

Figure 3 shows the results with three curves. For each curve, we fix the values of two  $k$ 's to be 200 and let the remaining  $k$  vary. This plot suggests that the scaled mean squared error ( $sMSE$ ) is





**FIGURE 1** The relative errors of the estimators with  $n = 5000$ ,  $p_i = 10^{2-i}/n$ , for  $i = 1, 2, 3$ , and  $k = k_1 = k_2 = 200$ , for Model 1(a), Model 1(b), and Model 1(c) listed in Table 1



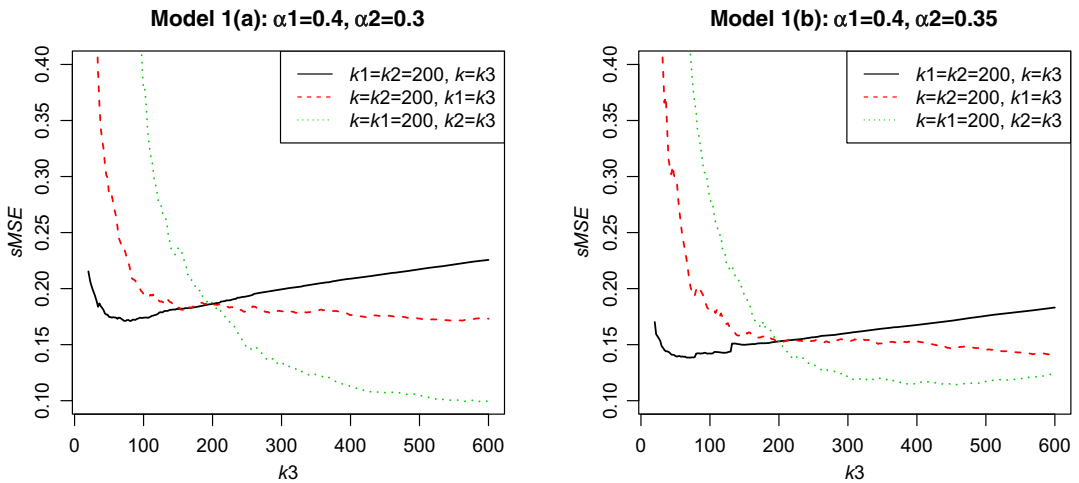
**FIGURE 2** The relative errors of the estimators with  $n = 5000$ ,  $p_i = 10^{2-i}/n$ , for  $i = 1, 2, 3$ , and  $k = k_1 = k_2 = 200$ , for Model 1(d) and Model 2(a) in Table 1

rather stable for a wide range of  $k$  and  $k_1$ , and for this model, an “optimal”  $k_2$  is typically larger than  $k$  and  $k_1$ . For Model 1(c), we obtained a similar result (not presented here) with larger  $sMSE$ .

Theorem 1 states the asymptotic normality result for the estimated error  $\hat{\theta}_p/\theta_p - 1$ , which suggests the following  $100(1 - \alpha)\%$  confidence interval for  $\theta_p$ :

$$\left[ \frac{\hat{\theta}_p}{1 - \sigma\left(\frac{n}{k}\right)^{\frac{1}{2n}-\frac{1}{2}} k^{-\frac{1}{2}} Z_{\alpha/2}}, \frac{\hat{\theta}_p}{1 + \sigma\left(\frac{n}{k}\right)^{\frac{1}{2n}-\frac{1}{2}} k^{-\frac{1}{2}} Z_{\alpha/2}} \right], \tag{11}$$

where  $Z_{\alpha/2}$  is the  $\alpha/2$ -quantile of a standard normal random variable and  $\sigma$  is defined in Theorem 1 with  $c(x, y) = 2^{\alpha_1/\alpha_2-1}(x \wedge y)^{\alpha_1/\alpha_2}$ . This is a pseudo confidence interval as it is based on the asymptotic normality, and the standard deviation, which depends on  $c(x, 1)$ , is typically unknown in practice. In addition, note that the standard deviation of the limit distribution does not depend on the value of  $p$ . Table 2 reports the coverage fraction for this confidence interval



**FIGURE 3** The scaled mean squared error ( $sMSE$ ) defined by (10) for (left panel) Model 1(a) and (right panel) Model 1(b) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

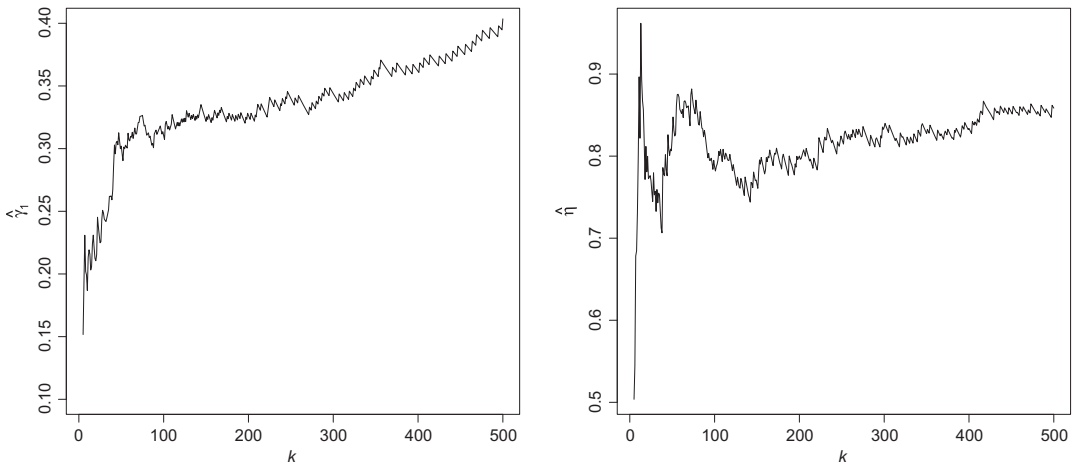
**TABLE 2** Coverage fraction for 95% confidence intervals of  $\theta_p$  given by (11), where  $p_i = 2 \cdot 10^{-i-2}$ , for  $i = 1, 2, 3$

Distribution	$(\alpha_1, \alpha_2)$	$n = 5000$			$n = 10^4$			$n = 10^5$		
		$\theta_{p_1}$	$\theta_{p_2}$	$\theta_{p_3}$	$\theta_{p_1}$	$\theta_{p_2}$	$\theta_{p_3}$	$\theta_{p_1}$	$\theta_{p_2}$	$\theta_{p_3}$
Model 1(a)	(0.4, 0.3)	0.42	0.28	0.21	0.74	0.51	0.40	0.95	0.7	0.46
Model 1(b)	(0.4, 0.35)	0.53	0.34	0.24	0.61	0.38	0.27	0.94	0.62	0.46
Model 1(c)	(0.6, 0.4)	0.51	0.32	0.23	0.54	0.35	0.23	0.77	0.61	0.42

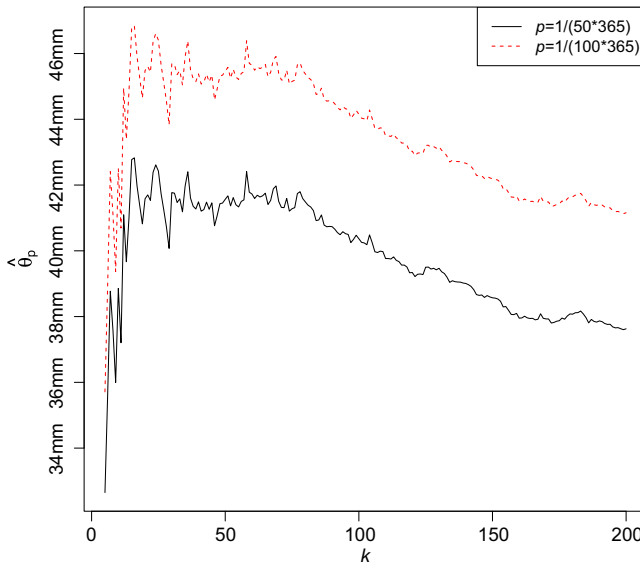
based on 500 simulations. We consider three sample sizes  $n = 5000, 10^4$ , and  $10^5$ . Based on the observation from sensitivity analysis, we choose  $k = k_1 = 200$  for all three sample sizes, but for  $k_2$ , we choose 200, 600, 700, respectively, for  $n = 5000, 10^4$ , and  $10^5$ . Yet, this is far from an optimal choice, which obviously depends on the distribution of the data and the sample size. Nevertheless, the message from Table 2 is rather clear; in order to have the asymptotic limit to hold exactly, it requires a very large sample size. The problem becomes more difficult in the sense that more data are needed if one extrapolates more (smaller  $p$ ).

## 4 | APPLICATION

We apply our estimation to daily precipitation data from two weather stations in the Netherlands, namely, Cabauw and Rotterdam. The distance between these two stations is about 32 km. The station Cabauw is close to the Lek river, whereas the station Rotterdam is close to the river Nieuwe Maas, which is the continuation of Lek. Heavy rainfall at both stations might lead to a severe flood in this region. Thus, the expected amount of rainfall in Cabauw given a heavy rainfall in Rotterdam is an important risk measure for the hydrology safety control. We estimate this quantity based on the data from August 1, 1990, to December 31, 2016. After removing the missing values, there are in total 9605 observations. There is open access to the data at <http://projects.knmi.nl/klimatologie/uurgegevens/selectie.cgi>.



**FIGURE 4** The Hill estimates of  $\gamma_1$  for the daily precipitation at Cabauw (left panel) and the estimates of  $\eta$  for the daily precipitations at Cabauw and Rotterdam (right panel)



**FIGURE 5** The estimates of the conditional expected rainfall at Cabauw given that the rainfall at Rotterdam exceeds the 50 (red) and 100 (black) year return level [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Let  $X$  be the daily rainfall at Cabauw and  $Y$  be the daily rainfall at Rotterdam. Before applying our estimation, we first obtain the estimate of extremal index using block method (cf. Smith & Weissman, 1994) and the result (not presented here) is very close to one for both stations, which indicates no clustering of extremes. Next, we look at the sign of the extreme value index of  $X$  and the extremal dependence of  $X$  and  $Y$ . From the Hill estimates of  $\gamma$  as shown in right panel of Figure 4, we conclude that  $\gamma > 0$ , which is in line with the existing literature. For instance, Buishand, de Haan, and Zhou (2008) obtain 0.1082 as the estimate of  $\gamma$  for the daily rainfall in the Netherlands, and Coles and Tawn (1996) report 0.066 as the estimate of  $\gamma$  for the daily rainfall in the southwest of England.

Next, we compute the Hill estimator of  $\eta$  given by (9). The estimates are above 0.5 as shown in the right panel of Figure 4.

Finally, we apply our estimator to answer the following question. Provided that the amount of rainfall in Rotterdam exceeds the  $M$ -year return level, what is the expected amount of rainfall in Cabauw, respectively, for  $M = 50$  and  $100$ ? Let  $R_M$  denote the  $M$ -year return level. Coles (2001) gives the definition of  $R_M$  as the level expected to be exceeded once every  $M$  years. As we consider the daily precipitation,  $R_M = U_2(365M)$ .

Choosing  $k_1 = k_2 = 200$ , we obtain the following estimates of  $\gamma$  and  $\eta$ :  $\hat{\gamma}_1 = 0.326$  and  $\hat{\eta} = 0.835$ . Figure 5 plots the estimates of  $\theta_p$  against  $k$ , from which we conclude  $k = 50$  lying in the interval where the estimates are stable. We thus report the following estimates of  $\hat{\theta}_p$ : 41.6 mm for  $M = 50$  and 45.5 mm for  $M = 100$ .

## 5 | PROOFS

*Proof of Proposition 1.* We recall that, for any positive random variable  $Z$ , the expectation can be written as

$$\mathbb{E}[Z] = \int_0^\infty \mathbb{P}(Z > x) dx. \quad (12)$$

Then, by definition of  $\theta_p$  and a change of variable, we have

$$\begin{aligned} \theta_{1/t} &= \int_0^\infty t \mathbb{P}(X > x, Y > U_2(t)) dx \\ &= t^{-\frac{1}{\eta}+1} U_1(t) \int_0^\infty t^{\frac{1}{\eta}} \mathbb{P}(X > x U_1(t), Y > U_2(t)) dx. \end{aligned} \quad (13)$$

Define  $f_t(x) := t^{\frac{1}{\eta}} \mathbb{P}(X > x U_1(t), Y > U_2(t))$ ,  $x > 0$ . Then,

$$\frac{\theta_{1/t}}{t^{-\frac{1}{\eta}+1} U_1(t)} = \int_0^\infty f_t(x) dx.$$

For any fixed  $x$ , by (3) and the continuity of the function  $x \mapsto c(x, 1)$ , we have

$$\lim_{t \rightarrow \infty} f_t(x) = c\left(x^{-\frac{1}{\eta}}, 1\right).$$

We shall apply the generalized dominated convergence theorem to validate that  $\lim_{t \rightarrow \infty} \int_0^\infty f_t(x) dx = \int_0^\infty c(x^{-\frac{1}{\eta}}, 1) dx$ .

By Assumption A(1), for any  $\epsilon > 0$ , there exists  $t_0$  such that

$$\left| \frac{U_1(t)}{t^{\gamma_1}} - d \right| < \epsilon, \quad \text{for all } t > t_0.$$

Hence, for  $c_1 = (d + \epsilon)/(d - \epsilon)$  and  $x > c_1(t_0/t)^{\gamma_1}$ , we get

$$\frac{U_1(t)x}{U_1\left(t(x/c_1)^{\frac{1}{\eta}}\right)} = \frac{U_1(t)/t^{\gamma_1}}{U_1\left(t(x/c_1)^{\frac{1}{\eta}}\right)/(t^{\gamma_1}x/c_1)} c_1 > \frac{d - \epsilon}{d + \epsilon} c_1 = 1.$$

Consequently, for  $x > c_1(t_0/t)^{\gamma_1}$ ,

$$f_t(x) \leq t^{\frac{1}{\eta}} \mathbb{P}\left(X > U_1\left(t(x/c_1)^{\frac{1}{\eta}}\right), Y > U_2(t)\right) = c_t\left((x/c_1)^{-\frac{1}{\eta}}, 1\right).$$

On the other hand, for  $0 < x \leq c_1(t_0/t)^{\gamma_1}$ ,  $f_t(x) \leq t^{\frac{1}{n}-1}$ . Define

$$g_t(x) := \begin{cases} c_t \left( (x/c_1)^{-\frac{1}{\gamma_1}}, 1 \right), & \text{if } x > c_1(t_0/t)^{\gamma_1}; \\ t^{\frac{1}{n}-1}, & \text{otherwise.} \end{cases}$$

Then,  $f_t(x) \leq g_t(x)$ . By generalized dominated convergence theorem, it is then sufficient to prove that

$$\lim_{t \rightarrow \infty} \int_0^\infty g_t(x) dx = \int_0^\infty \lim_{t \rightarrow \infty} g_t(x) dx = \int_0^\infty c \left( (x/c_1)^{-\frac{1}{\gamma_1}}, 1 \right) dx.$$

Observe that

$$\begin{aligned} \int_0^\infty g_t(x) dx &= \int_0^{c_1(t_0/t)^{\gamma_1}} t^{\frac{1}{n}-1} dx + c_1 \int_{(t_0/t)^{\gamma_1}}^\infty c_t \left( x^{-\frac{1}{\gamma_1}}, 1 \right) dx \\ &= c_1 t_0^{\gamma_1} t^{\frac{1}{n}-1-\gamma_1} + c_1 \int_{(t_0/t)^{\gamma_1}}^\infty c_t \left( x^{-\frac{1}{\gamma_1}}, 1 \right) dx \\ &\rightarrow 0 + c_1 \int_0^\infty c \left( x^{-\frac{1}{\gamma_1}}, 1 \right) dx, \end{aligned}$$

as  $t \rightarrow \infty$ . The last convergence follows from  $\frac{1}{n} - 1 - \gamma_1 < 0$ ,  $\int_0^{(t_0/t)^{\gamma_1}} c(x^{-\frac{1}{\gamma_1}}, 1) dx \rightarrow 0$ , and the fact that

$$\begin{aligned} &\left| \int_{(t_0/t)^{\gamma_1}}^\infty c_t \left( x^{-\frac{1}{\gamma_1}}, 1 \right) dx - \int_{(t_0/t)^{\gamma_1}}^\infty c \left( x^{-\frac{1}{\gamma_1}}, 1 \right) dx \right| \\ &\leq \int_{(t_0/t)^{\gamma_1}}^1 \left| c_t \left( x^{-\frac{1}{\gamma_1}}, 1 \right) - c \left( x^{-\frac{1}{\gamma_1}}, 1 \right) \right| dx + \int_1^\infty \left| c_t \left( x^{-\frac{1}{\gamma_1}}, 1 \right) - c \left( x^{-\frac{1}{\gamma_1}}, 1 \right) \right| dx \\ &= o(1) \int_{(t_0/t)^{\gamma_1}}^1 x^{-\beta_2/\gamma_1} dx + o(1) \int_1^\infty x^{-\beta_1/\gamma_1} dx \rightarrow 0, \end{aligned}$$

by Assumptions A(2) and A(3). □

Through out the proof section, we denote the speed of the convergence in Theorem 1 by

$$T_n = \sqrt{k} \left( \frac{n}{k} \right)^{-\frac{1}{2\eta} + \frac{1}{2}}. \tag{14}$$

From Assumption B(5),  $T_n \rightarrow \infty$ , as  $n \rightarrow \infty$ . By construction, the asymptotic normality of  $\hat{\theta}_p$  depends on the asymptotic behavior of  $\hat{\theta}_{k/n}$ , which is given in Proposition 2.

**Proposition 2.** *Under the assumptions of Theorem 1, it holds*

$$\frac{T_n \left( \frac{n}{k} \right)^{\frac{1}{\eta}-1}}{U_1(n/k)} \left( \hat{\theta}_{\frac{k}{n}} - \theta_{\frac{k}{n}} \right) \xrightarrow{d} N(0, \sigma_1^2),$$

where  $\sigma_1^2 = - \int_0^\infty c(x, 1) dx^{-2\gamma_1}$ .

The proof of Proposition 2 is postponed to the Appendix.

*Proof of Theorem 1.* Recall that  $d_n = \frac{k}{np}$ . By the definition of  $\hat{\theta}_p$ , we make the following decomposition:

$$\begin{aligned} \frac{\hat{\theta}_p}{\theta_p} &= \frac{d_n^{-\frac{1}{\eta}+1+\hat{\gamma}_1} \hat{\theta}_{k/n}}{\theta_p} = d_n^{\hat{\gamma}_1-\gamma_1} d_n^{\frac{1}{\eta}-\frac{1}{\eta}} \frac{\hat{\theta}_{k/n}}{\theta_{k/n}} \frac{d_n^{-\frac{1}{\eta}+1+\gamma_1} \theta_{k/n}}{\theta_p} \\ &=: I_1 \cdot I_2 \cdot I_3 \cdot I_4. \end{aligned}$$

We shall show that these four terms all converges to unity at certain rates. First, from the assumption that  $\sqrt{k}(\hat{\gamma}_1 - \gamma_1) = O_P(1)$ , it follows that

$$I_1 - 1 = e^{(\hat{\gamma}_1 - \gamma_1) \log d_n} - 1 = (\hat{\gamma}_1 - \gamma_1) \log d_n + o((\hat{\gamma}_1 - \gamma_1) \log d_n) = O_P\left(\frac{\log d_n}{\sqrt{k}}\right) = o_P\left(\frac{1}{T_n}\right).$$

In the last equality, we used the assumption that  $\lim_{n \rightarrow \infty} (n/k)^{1/2-1/(2\eta)} \log d_n = 0$ . Recall that  $T_n$  is defined in (14).

In the same way, we get  $I_2 - 1 = o_P\left(\frac{1}{T_n}\right)$ .

Combining Propositions 1 and 2, we derive that

$$\begin{aligned} T_n(I_3 - 1) &= \frac{T_n}{\theta_{k/n}} (\hat{\theta}_{k/n} - \theta_{k/n}) = \frac{T_n \left(\frac{n}{k}\right)^{\frac{1}{\eta}-1}}{U_1(n/k)} (\hat{\theta}_{k/n} - \theta_{k/n}) \cdot \frac{U_1(n/k) \left(\frac{n}{k}\right)^{-\frac{1}{\eta}+1}}{\theta_{k/n}} \\ &\xrightarrow{d} \left( \int_0^\infty c\left(x^{-\frac{1}{\eta}}, 1\right) dx \right)^{-1} N(0, \sigma_1^2). \end{aligned}$$

That is,  $T_n(I_3 - 1) \xrightarrow{P} \Gamma_1$ , where  $\Gamma_1$  is a normal distribution with mean zero and variance,  $\sigma^2 = -\int_0^\infty c(x, 1) dx^{-2\gamma_1} \left| \int_0^\infty c\left(x^{-\frac{1}{\eta}}, 1\right) dx \right|^{-2}$ , which is the limit distribution in Theorem 1.

Then, we deal with the last term,  $I_4$ . Here we need a rate for the convergence in Proposition 1. Continuing with (13) and using that  $F_1$  is strictly increasing, we get

$$\begin{aligned} \frac{\theta_{k/n}}{U_1(n/k)(n/k)^{-\frac{1}{\eta}+1}} &= \int_0^\infty \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \mathbb{P}(X > xU_1(n/k), Y > U_2(n/k)) dx \\ &= - \int_0^\infty c_{\frac{n}{k}}(s_n(x), 1) dx^{-\gamma_1}, \end{aligned} \quad (15)$$

with

$$s_n(x) = \frac{n}{k} [1 - F_1(U_1(n/k)x^{-\gamma_1})]. \quad (16)$$

By the regular variation of  $1 - F_1$ , we have  $\lim_{n \rightarrow \infty} s_n(x) = x$ , for any  $x > 0$ . By (iii) and (v) in Lemma 1 in the Appendix, we have that

$$\int_0^\infty c_{\frac{n}{k}}(s_n(x), 1) dx^{-\gamma_1} = \int_0^\infty c(x, 1) dx^{-\gamma_1} + o\left(\frac{1}{T_n}\right). \quad (17)$$

It follows from Assumptions B(4) and B(5) that

$$\frac{U_1(1/p)}{U_1(n/k)d_n^{\gamma_1}} - 1 = O(A(n/k)) = o\left(\frac{1}{\sqrt{k}}\right).$$

Combining this result with (15) and (17) leads to

$$\begin{aligned}
 I_4 &= \frac{\theta_{k/n}}{U_1(n/k)(n/k)^{-\frac{1}{\eta}+1}} \frac{U_1(1/p)(1/p)^{-\frac{1}{\eta}+1}}{\theta_p} \frac{U_1(n/k)}{U_1(1/p)} \left(\frac{k}{np}\right)^{\gamma_1} \\
 &= \left(\int_0^\infty c(x, 1) dx^{-\gamma_1} + o\left(\frac{1}{T_n}\right)\right) \left(\int_0^\infty c(x, 1) dx^{-\gamma_1} + o\left(\frac{1}{T_n}\right)\right)^{-1} \left(1 + o\left(\frac{1}{\sqrt{k}}\right)\right)^{-1} \\
 &= 1 + o\left(\frac{1}{T_n}\right).
 \end{aligned}$$

Thus, we obtain

$$\begin{aligned}
 \frac{\hat{\theta}_p}{\theta_p} - 1 &= I_1 I_2 I_3 I_4 - 1 \\
 &= \left[1 + o_p\left(\frac{1}{T_n}\right)\right]^2 \left[1 + \frac{\Gamma_1}{T_n} + o_p\left(\frac{1}{T_n}\right)\right] \left[1 + o\left(\frac{1}{T_n}\right)\right] - 1 \\
 &= \frac{\Gamma_1}{T_n} + o_p\left(\frac{1}{T_n}\right).
 \end{aligned}$$

The proof is completed. □

## ORCID

Eni Musta  <https://orcid.org/0000-0003-3356-4307>

## REFERENCES

- Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. *The Review of Financial Studies*, 30, 2–47.
- Adler, R. J. (1990). *Lecture Notes-Monograph Series. An introduction to continuity, extrema, and related topics for general Gaussian processes*. Rockville, MD: Institute of Mathematical Statistics.
- Beirlant, J., Goegebeur, Y., Segers, J., & Teugels, J. L. (2004). *Statistics of extremes, theory and applications*. Chichester, England: John Wiley & Sons.
- Buishand, T. A., de Haan, L., & Zhou, C. (2008). On spatial extremes: With application to a rainfall problem. *Annals of Applied Statistics*, 2, 624–642.
- Cai, J., & Li, H. (2005). Conditional tail expectations for multivariate phase-type distributions. *Journal of Applied Probability*, 42, 810–825.
- Cai, J.-J. (2012). *Estimation concerning risk under extreme value conditions*. Tilburg, The Netherlands: Tilburg University.
- Cai, J.-J., Einmahl, J. H. J., de Haan, L., & Zhou, C. (2015). Estimation of the marginal expected shortfall: The mean when a related variable is extreme. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77, 417–442.
- Caporin, M., & Santucci de Magistris, P. (2012). On the evaluation of marginal expected shortfall. *Applied Economics Letters*, 19, 175–179.
- Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London, UK: Springer-Verlag.
- Coles, S. G., & Tawn, J. A. (1996). Modelling extremes of the areal rainfall process. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 58, 329–347.
- Cousin, A., & Di Bernardino, E. (2014). On multivariate extensions of conditional-tail-expectation. *Insurance: Mathematics and Economics*, 55, 272–282.
- Das, B., & Fasen-Hartmann, V. (2018). Risk contagion under regular variation and asymptotic tail independence. *Journal of Multivariate Analysis*, 165, 194–215.

- de Haan, L., & Ferreira, A. (2006). *Extreme value theory: An introduction*. Berlin, Germany: Springer.
- Draisma, G., Drees, H., Ferreira, A., & de Haan, L. (2004). Bivariate tail estimation: Dependence in asymptotic independence. *Bernoulli*, 10, 251–280.
- Heffernan, J. E., & Resnick, S. I. (2007). Limit laws for random vectors with an extreme component. *The Annals of Applied Probability*, 17, 537–571.
- Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66, 497–546.
- Kulik, R., & Soulier, P. (2015). Heavy tailed time series with extremal independence. *Extremes*, 18, 273–299.
- 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. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59, 475–499.
- Resnick, S. (2002). Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5, 303–336.
- Smith, R. L., & Weissman, I. (1994). Estimating the extremal index. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56, 515–528.
- van der Vaart, A. W. (1998). *Cambridge Series in Statistical and Probabilistic Mathematics: Vol. 3. Asymptotic statistics*. Cambridge, UK: Cambridge University Press.
- van der Vaart, A. W., & Wellner, J. A. (1996). *Weak convergence and empirical processes with applications to statistics*. New York, NY: Springer-Verlag.

**How to cite this article:** Cai J-J, Musta E. Estimation of the marginal expected shortfall under asymptotic independence. *Scand J Statist.* 2020;47:56–83. <https://doi.org/10.1111/sjos.12397>

## APPENDIX

### Proof of Proposition 2

In this section, all the limit processes involved are defined in the same probability space via the Skorohod construction, that is, they are only equal in distribution to the original processes. If we define

$$e_n = \frac{n}{k} (1 - F_2(Y_{n-k,n})),$$

we have

$$\hat{\theta}_{k/n} = \frac{1}{k} \sum_{i=1}^n X_i \mathbb{1}_{\{Y_i > U_2(\frac{n}{ke_n})\}}.$$

Note that  $e_n \xrightarrow{\mathbb{P}} 1$  because  $1 - F_2(Y_{n-k,n})$  is the  $(k+1)$ th order statistics of a random sample from the standard uniform distribution.

We first investigate the asymptotic behavior of

$$\tilde{\theta}_{\frac{ky}{n}} = \frac{1}{ky} \sum_{i=1}^n X_i \mathbb{1}_{\{Y_i > U_2(\frac{n}{ky})\}}$$

as a random process for  $y \in [1/2, 2]$ .



Let  $W(y)$  denote a mean zero Gaussian process on  $[1/2, 2]$  with covariance structure

$$\mathbb{E}[W(y_1)W(y_2)] = -\frac{1}{y_1 y_2} \int_0^\infty c(x, y_1 \wedge y_2) dx^{-2\gamma_1}, \quad y_1, y_2 \in [1/2, 2].$$

**Proposition 3.** *Suppose conditions B(1)–B(5) hold. Let  $S_n = \left(\frac{n}{k}\right)^{\frac{1}{2p}-\frac{1}{2}} \sqrt{k}$ . Then, as  $n \rightarrow \infty$ ,*

$$\left\{ \frac{S_n}{U_1(n/k)} \left( \tilde{\theta}_{\frac{ky}{n}} - \theta_{\frac{ky}{n}} \right) \right\}_{y \in [1/2, 2]} \xrightarrow{d} \{W(y)\}_{y \in [1/2, 2]}. \tag{A1}$$

The convergence of the process holds in distribution in  $l^\infty(1/2, 2)$ .

Before proving Proposition 3, we first show two lemmas. The first lemma states some properties of the functions  $c_t(x, y)$  and  $c(x, y)$ , which will be used frequently in the proof. The second lemma is established to compute the covariance of the limit process in Proposition 3.

**Lemma 1.**

(i) *If  $\int_0^\infty c(x, y) dx^{-2\gamma_1} < \infty$ , then*

$$\int_0^\infty \int_0^\infty c(x_1 \wedge x_2, y) dx_1^{-\gamma_1} dx_2^{-\gamma_1} = \int_0^\infty c(x, y) dx^{-2\gamma_1}.$$

(ii) *The function  $y \mapsto \int_0^\infty c(x, y) dx^{-\gamma_1}$  is Lipschitz, that is, there exists  $C > 0$  such that, for each  $y_1, y_2 \in [1/2, 2]$ ,*

$$\left| \int_0^\infty c(x, y_1) dx^{-\gamma_1} - \int_0^\infty c(x, y_2) dx^{-\gamma_1} \right| \leq C|y_1 - y_2|.$$

(iii) *Assumptions B(1) and B(3)–B(5) imply that*

$$\sup_{y \in [1/2, 2]} T_n \left| \int_0^\infty c(x, y) dx^{-\gamma_1} - \int_0^\infty c(s_n(x), y) dx^{-\gamma_1} \right| \rightarrow 0.$$

(iv) *If Assumptions B(2)–B(3) hold, then, for  $\rho = 1, 2, 2 + \bar{\delta}$ ,*

$$\left| \int_0^\infty c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, y \right) dx^{-\rho\gamma_1} - \int_0^\infty c(x, y) dx^{-\rho\gamma_1} \right| \rightarrow 0. \tag{A2}$$

(v) *If Assumptions B(2), B(3), and B(5) hold, then*

$$\sup_{y \in [1/2, 2]} T_n \left| \int_0^\infty c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, y \right) dx^{-\gamma_1} - \int_0^\infty c(x, y) dx^{-\gamma_1} \right| \rightarrow 0, \tag{A3}$$

and

$$\sup_{y \in [1/2, 2]} T_n \left| \int_0^\infty c_{\frac{n}{k}}(s_n(x), y) dx^{-\gamma_1} - \int_0^\infty c(s_n(x), y) dx^{-\gamma_1} \right| \rightarrow 0. \tag{A4}$$

*Proof.* The first statement follows from simple transformations of the integral. Indeed, we have

$$\begin{aligned} & \int_0^\infty \int_0^\infty c(x_1 \wedge x_2, y) dx_1^{-\gamma_1} dx_2^{-\gamma_1} \\ &= \int_0^\infty \int_{x_1}^\infty c(x_1, y) dx_2^{-\gamma_1} dx_1^{-\gamma_1} + \int_0^\infty \int_0^{x_1} c(x_2, y) dx_2^{-\gamma_1} dx_1^{-\gamma_1} \\ &= 2 \int_0^\infty x_1^{-\gamma_1} c(x_1, y) dx_1^{-\gamma_1} \\ &= \int_0^\infty c(x, y) dx^{-2\gamma_1}. \end{aligned}$$

By the homogeneity property of  $c(x, y)$ :  $c(kx, ky) = k^{\frac{1}{\eta}} c(x, y)$ , we have

$$\begin{aligned} & \left| \int_0^\infty c(x, y_1) dx^{-\gamma_1} - \int_0^\infty c(x, y_2) dx^{-\gamma_1} \right| \\ &= \left| \int_0^\infty y_1^{1/\eta} c\left(\frac{x}{y_1}, 1\right) dx^{-\gamma_1} - \int_0^\infty y_2^{1/\eta} c\left(\frac{x}{y_2}, 1\right) dx^{-\gamma_1} \right| \\ &= \left| y_1^{1/\eta-2\gamma_1} \int_0^\infty c(x, 1) dx^{-\gamma_1} - y_2^{1/\eta-2\gamma_1} \int_0^\infty c(x, 1) dx^{-\gamma_1} \right| \\ &\leq K|y_1 - y_2|. \end{aligned}$$

(iii) Let  $l_n = \left(\frac{n}{k}\right)^\lambda$  with  $\lambda$  as in Assumption B(5). We start by writing

$$\begin{aligned} & \sup_{y \in [1/2, 2]} T_n \left| \int_0^\infty c(x, y) dx^{-\gamma_1} - \int_0^\infty c(s_n(x), y) dx^{-\gamma_1} \right| \\ &\leq \sup_{y \in [1/2, 2]} T_n \left| \int_0^{l_n} [c(x, y) - c(s_n(x), y)] dx^{-\gamma_1} \right| \\ &\quad + \sup_{y \in [1/2, 2]} T_n \left\{ \left| \int_{l_n}^\infty c(x, y) dx^{-\gamma_1} \right| + \left| \int_{l_n}^\infty c(s_n(x), y) dx^{-\gamma_1} \right| \right\}. \end{aligned} \tag{A5}$$

First, we deal with the first term in the right-hand side of (A5). By the homogeneity and monotonicity property of  $c(x, y)$ , we have

$$|c(x_1, y) - c(x_2, y)| \leq \left| \left(\frac{x_2}{x_1}\right)^{1/\eta} - 1 \right| c(x_1, y).$$

It follows that

$$\begin{aligned} & \sup_{y \in [1/2, 2]} T_n \left| \int_0^{l_n} [c(x, y) - c(s_n(x), y)] dx^{-\gamma_1} \right| \\ &\leq \sup_{y \in [1/2, 2]} T_n \left| \int_0^{l_n} \left| \left(\frac{s_n(x)}{x}\right)^{1/\eta} - 1 \right| c(x, y) dx^{-\gamma_1} \right|. \end{aligned}$$

Note that, for any  $\epsilon_0 > 0$ , for sufficiently large  $n$  and  $x < l_n$  (see Cai, 2012, p. 85),

$$\left| \frac{s_n(x)/x - 1}{A_1(n/k)} - \frac{x^{-\rho_1} - 1}{\gamma_1 \rho_1} \right| \leq x^{-\rho_1} \max(x^{\epsilon_0}, x^{-\epsilon_0}).$$

This implies that

$$\left| \frac{s_n(x)}{x} - 1 \right| \leq |A_1(n/k)| \left\{ \left| \frac{x^{-\rho_1} - 1}{\gamma_1 \rho_1} \right| + x^{-\rho_1} \max(x^{\epsilon_0}, x^{-\epsilon_0}) \right\}.$$

Because, for  $\epsilon_0 < -\rho_1(1 - \lambda)/\lambda$  and  $x < l_n$ ,

$$|A_1(n/k)| \left\{ \left| \frac{x^{-\rho_1} - 1}{\gamma_1 \rho_1} \right| + x^{-\rho_1} \max(x^{\epsilon_0}, x^{-\epsilon_0}) \right\} = o(1)$$

by a Taylor expansion, we obtain

$$\begin{aligned} \left| \left( \frac{S_n(x)}{x} \right)^{1/\eta} - 1 \right| &= |A_1(n/k)| \left\{ \left| \frac{x^{-\rho_1} - 1}{\gamma_1 \rho_1} \right| + x^{-\rho_1} \max(x^{\epsilon_0}, x^{-\epsilon_0}) \right\} \\ &+ o \left( |A_1(n/k)| \left\{ \left| \frac{x^{-\rho_1} - 1}{\gamma_1 \rho_1} \right| + x^{-\rho_1} \max(x^{\epsilon_0}, x^{-\epsilon_0}) \right\} \right). \end{aligned}$$

Consequently,

$$\begin{aligned} \sup_{y \in [1/2, 2]} T_n \left| \int_0^{l_n} \left( \frac{S_n(x)}{x} \right)^{1/\eta} - 1 \right| c(x, y) dx^{-\gamma_1} \Big| & \\ \leq C \sup_{y \in [1/2, 2]} T_n |A_1(n/k)| \left| \int_0^{l_n} x^{-\rho_1} \max(x^{\epsilon_0}, x^{-\epsilon_0}) c(x, y) dx^{-\gamma_1} \right|. & \tag{A6} \end{aligned}$$

Furthermore, using the triangular inequality and Cauchy–Schwarz, we get

$$\begin{aligned} &\left| \int_0^{l_n} x^{-\rho_1} \max(x^{\epsilon_0}, x^{-\epsilon_0}) c(x, y) dx^{-\gamma_1} \right| \\ &\leq \int_0^1 x^{-\rho_1 - \epsilon_0} c(x, y) dx^{-\gamma_1} + \left| \int_1^\infty c(x, y)^2 dx^{-\gamma_1} \right|^{1/2} \left| \int_1^{l_n} x^{-2\rho_1 + 2\epsilon_0} dx^{-\gamma_1} \right|^{1/2} \\ &= O \left( l_n^{-\rho_1 + \epsilon_0 - \frac{\gamma_1}{2}} \right). \end{aligned}$$

Going back to (A6), we obtain

$$\sup_{y \in [1/2, 2]} T_n \left| \int_0^{l_n} \left( \frac{S_n(x)}{x} \right)^{1/\eta} - 1 \right| c(x, y) dx^{-\gamma_1} \Big| = O \left( \sqrt{k} \left( \frac{n}{k} \right)^{-\frac{1}{2\eta} + \frac{1}{2} - \lambda \frac{\gamma_1}{2}} \right) \rightarrow 0$$

because of Assumption B(5).

Next, we deal with the second term in the right-hand side of (A5). By Cauchy–Schwarz and Assumption B(1), we obtain

$$\begin{aligned} \left| \int_{l_n}^\infty c(x, y) dx^{-\gamma_1} \right| &\leq \gamma_1 \left( \int_{l_n}^\infty x^{-\gamma_1 - 1} dx \right)^{1/2} \left( \int_1^\infty c(x, y)^2 x^{-\gamma_1 - 1} dx \right)^{1/2} \\ &\leq C l_n^{-\gamma_1/2} \end{aligned} \tag{A7}$$

for some constant  $C > 0$ . Moreover, because

$$T_n l_n^{-\gamma_1/2} = \sqrt{k} \left( \frac{n}{k} \right)^{-\frac{1}{2\eta} + \frac{1}{2} - \lambda \frac{\gamma_1}{2}},$$

by Assumption B(5), it follows that

$$\sup_y T_n \left| \int_{l_n}^\infty c(x, y) dx^{-\gamma_1} \right| \rightarrow 0.$$

Furthermore, the triangular inequality and B(3) yield

$$\begin{aligned} & \left| \int_{l_n}^{\infty} c(s_n(x), y) dx^{-\gamma_1} \right| \\ & \leq \left| \int_{l_n}^{\infty} c_{\frac{n}{k}}(s_n(x), y) dx^{-\gamma_1} \right| + \left| \int_{l_n}^{\infty} [c(s_n(x), y) - c_{\frac{n}{k}}(s_n(x), y)] dx^{-\gamma_1} \right| \\ & \leq \left| \int_{l_n}^{\infty} c_{\frac{n}{k}}(s_n(x), y) dx^{-\gamma_1} \right| + \sup_{\substack{0 < x < n/k \\ y \in [1/2, 2]}} \frac{|c_{\frac{n}{k}}(x, y) - c(x, y)|}{x^{\beta_2}} \int_{l_n}^{\infty} s_n(x)^{\beta_2} dx^{-\gamma_1}. \end{aligned}$$

Note that, by Assumption B(3),

$$\sup_{\substack{0 < x < n/k \\ y \in [1/2, 2]}} \frac{|c_{\frac{n}{k}}(x, y) - c(x, y)|}{x^{\beta_2}} = O_P\left(\left(\frac{n}{k}\right)^{\tau}\right).$$

Then, using the definition of  $s_n$ , a change of variable, and Jensen inequality, we obtain

$$\begin{aligned} \int_{l_n}^{\infty} s_n(x)^{\beta_2} dx^{-\gamma_1} &= \int_{l_n}^{\infty} \left\{ \frac{n}{k} \mathbb{P}(X > U_1(n/k)x^{-\gamma_1}) \right\}^{\beta_2} dx^{-\gamma_1} \\ &= \left(\frac{n}{k}\right)^{\beta_2} \int_0^{l_n^{-\gamma_1}} \{ \mathbb{P}(X > U_1(n/k)x) \}^{\beta_2} dx \\ &\leq \left(\frac{n}{k}\right)^{\beta_2} l_n^{-\gamma_1} \left\{ l_n^{\gamma_1} \int_0^{l_n^{-\gamma_1}} \mathbb{P}(X > U_1(n/k)x) dx \right\}^{\beta_2} \\ &= \left(\frac{n}{k}\right)^{\beta_2} l_n^{-\gamma_1} \left\{ \frac{l_n^{\gamma_1}}{U_1(n/k)} \int_0^{U_1(n/k)l_n^{-\gamma_1}} \mathbb{P}(X > x) dx \right\}^{\beta_2} \\ &\leq \left(\frac{n}{k}\right)^{\beta_2 - \lambda\gamma_1} \left(\frac{l_n^{\gamma_1}}{U_1(n/k)}\right)^{\beta_2} (\mathbb{E}[X])^{\beta_2} \\ &= o\left(\left(\frac{n}{k}\right)^{\beta_2 - \lambda\gamma_1}\right). \end{aligned} \tag{A8}$$

Hence, Assumption B(5) implies

$$\begin{aligned} & \sup_y T_n \left| \int_{l_n}^{\infty} [c(s_n(x), y) - c_{\frac{n}{k}}(s_n(x), y)] dx^{-\gamma_1} \right| \\ &= O_P\left(\sqrt{k} \left(\frac{n}{k}\right)^{-\frac{1}{2\eta} + \frac{1}{2} + \tau + \beta_2 - \lambda\gamma_1}\right) \rightarrow 0. \end{aligned}$$

On the other hand, using the definition of  $s_n$ , we get

$$\begin{aligned} & \left| \int_{l_n}^{\infty} c_{\frac{n}{k}}(s_n(x), y) dx^{-\gamma_1} \right| \\ &= \left| \int_{l_n}^{\infty} \left(\frac{n}{k}\right)^{1/\eta} \mathbb{P}\left[X > U_1\left(\frac{n}{ks_n(x)}\right), Y > U_2\left(\frac{n}{ky}\right)\right] dx^{-\gamma_1} \right| \\ &\leq \gamma_1 \frac{ky}{n} \left(\frac{n}{k}\right)^{1/\eta} l_n^{-\gamma_1}, \end{aligned}$$

and as a result,

$$\sup_y T_n \left| \int_{l_n}^{\infty} c_{\frac{n}{k}}(s_n(x), y) dx^{-\gamma_1} \right| \leq C\sqrt{k} \left(\frac{n}{k}\right)^{\frac{1}{2\eta} - \frac{1}{2} - \lambda\gamma_1} \rightarrow 0, \tag{A9}$$

because of Assumption B(5).

(iv) We write

$$\begin{aligned} & \sup_{y \in [1/2, 2]} \left| \int_0^{\infty} c_{\frac{n}{k}} \left(x \wedge \frac{n}{k}, y\right) dx^{-\rho\gamma_1} - \int_0^{\infty} c(x, y) dx^{-\rho\gamma_1} \right| \\ & \leq \sup_{y \in [1/2, 2]} \left\{ \left| \int_0^{\frac{n}{k}} c_{\frac{n}{k}}(x, y) dx^{-\rho\gamma_1} - \int_0^{\frac{n}{k}} c(x, y) dx^{-\rho\gamma_1} \right| \right. \\ & \quad \left. + \left| c_{\frac{n}{k}}(n/k, y) \int_{\frac{n}{k}}^{\infty} dx^{-\rho\gamma_1} \right| + \left| \int_{\frac{n}{k}}^{\infty} c(x, y) dx^{-\rho\gamma_1} \right| \right\} \\ & \leq \sup_{\substack{0 < x < n/k \\ y \in [1/2, 2]}} \frac{|c_{\frac{n}{k}}(x, y) - c(x, y)|}{x^{\beta_1} \wedge x^{\beta_2}} \left| \int_0^{\infty} x^{\beta_1} \wedge x^{\beta_2} dx^{-\rho\gamma_1} \right| + \left(\frac{n}{k}\right)^{-\rho\gamma_1} c_{\frac{n}{k}}(n/k, 2) + o(1). \end{aligned} \tag{A10}$$

The first term on the right-hand side of the inequality converges to zero by Assumptions B(2)–B(3). Moreover, by definition (4) and Assumption B(1), we have

$$\left(\frac{n}{k}\right)^{-\rho\gamma_1} c_{\frac{n}{k}}(n/k, 2) \leq \left(\frac{n}{k}\right)^{\frac{1}{\eta} - 1 - \rho\gamma_1} \rightarrow 0.$$

(v) We first give the proof for (A4). By Assumptions B(2) and B(3), we have

$$\begin{aligned} & \sup_{y \in [1/2, 2]} T_n \left| \int_0^{\infty} c_{\frac{n}{k}}(s_n(x), y) - c(s_n(x), y) dx^{-\gamma_1} \right| \\ & \leq T_n \sup_{\substack{0 < x < n/k \\ y \in [1/2, 2]}} \frac{|c_{\frac{n}{k}}(x, y) - c(x, y)|}{x^{\beta_1} \wedge x^{\beta_2}} \int_0^{\infty} s_n(x)^{\beta_1} \wedge s_n(x)^{\beta_2} dx^{-\gamma_1} \\ & = O\left(T_n \left(\frac{n}{k}\right)^{\tau}\right) \int_0^{\infty} s_n(x)^{\beta_1} \wedge s_n(x)^{\beta_2} dx^{-\gamma_1}. \end{aligned}$$

Next, we obtain an upper bound for the integral in the last equality. Because  $s_n(x)$  is monotone and  $s_n(1) = 1$ , we get the following bound for the integral from zero to one:

$$\int_0^1 s_n(x)^{\beta_1} dx^{-\gamma_1} < \int_{\mathbb{R}} s_n(x)^{\beta_1} \wedge 1 dx^{-\gamma_1},$$

which is shown to be  $O(1)$  in Cai et al. (2015, p. 438). Moreover, using the definition of  $s_n$ , a change of variable, and Jensen inequality, we obtain

$$\begin{aligned} \int_1^\infty s_n(x)^{\beta_2} dx^{-\gamma_1} &= \int_1^\infty \left\{ \frac{n}{k} \mathbb{P}(X > U_1(n/k)x^{-\gamma_1}) \right\}^{\beta_2} dx^{-\gamma_1} \\ &= \left( \frac{n}{k} \right)^{\beta_2} \int_0^1 \{ \mathbb{P}(X > U_1(n/k)x) \}^{\beta_2} dx \\ &\leq \left( \frac{n}{k} \right)^{\beta_2} \left\{ \int_0^1 \mathbb{P}(X > U_1(n/k)x) dx \right\}^{\beta_2} \\ &= \left( \frac{n}{k} \right)^{\beta_2} \left\{ \frac{1}{U_1(n/k)} \int_0^{U_1(n/k)} \mathbb{P}(X > x) dx \right\}^{\beta_2} \\ &\approx \left( \frac{n}{k} \right)^{\beta_2 - \gamma_1 \beta_2} (\mathbb{E}[X])^{\beta_2}. \end{aligned}$$

By Assumption B(5),

$$O\left(T_n \left(\frac{n}{k}\right)^{\tau + \beta_2 - \gamma_1 \beta_2}\right) = \sqrt{k} \left(\frac{n}{k}\right)^{-\frac{1}{2n} + \frac{1}{2} + \tau + \beta_2(1 - \gamma_1)} \rightarrow 0. \quad (\text{A11})$$

Thus, (A4) is proved.

The proof for (A3) can be obtained in a similar way. We use the triangular inequality as in (A10) to get

$$\begin{aligned} &\sup_{y \in [1/2, 2]} T_n \left| \int_0^\infty c_{\frac{n}{k}}\left(x \wedge \frac{n}{k}, y\right) dx^{-\gamma_1} - \int_0^\infty c(x, y) dx^{-\gamma_1} \right| \\ &\leq T_n \sup_{\substack{0 < x < n/k \\ y \in [1/2, 2]}} \left| \frac{c_{\frac{n}{k}}(x, y) - c(x, y)}{x^{\beta_1} \wedge x^{\beta_2}} \right| \left| \int_0^\infty x^{\beta_1} \wedge x^{\beta_2} dx^{-\rho \gamma_1} \right| \\ &\quad + T_n \left(\frac{n}{k}\right)^{-\gamma_1} c_{\frac{n}{k}}(n/k, 2) + T_n \left| \int_{\frac{n}{k}}^\infty c(x, 2) dx^{-\rho \gamma_1} \right| \\ &= A_1 + A_2 + A_3. \end{aligned}$$

$A_1$  converges to zero by (A11). Moreover, as in (A7),

$$A_3 = O\left(T_n \left(\frac{n}{k}\right)^{-\gamma_1/2}\right) \rightarrow 0$$

by Assumption B(5). Finally,

$$A_2 = O_P\left(T_n \left(\frac{n}{k}\right)^{-\gamma_1 + \frac{1}{n} - 1}\right) = O_P\left(\sqrt{k} \left(\frac{n}{k}\right)^{\frac{1}{2n} - \frac{1}{2} - \gamma_1}\right) \rightarrow 0$$

(see (A9)). □

**Lemma 2.** Assume B(1)–B(3). For  $y \in [1/2, 2]$  and  $\rho \in \{1, 2, 2 + \bar{\delta}\}$ , define

$$A_n(y, \rho) = \left(\frac{n}{k}\right)^{1/n} \left(- \int_0^\infty \mathbb{1}_{\{1 - F_1(X_1) < \frac{k}{n}x, 1 - F_2(Y_1) < \frac{kx}{n}\}} dx^{-\gamma_1}\right)^\rho.$$

Then,

$$\mathbb{E}[A_n(y, \rho)] \rightarrow - \int_0^\infty c(x, y) dx^{-\rho \gamma_1}.$$

*Proof.* Denote  $W_1 = 1 - F_1(X_1)$  and  $V_1 = 1 - F_2(Y_1)$ . Then, we can write the integral as

$$\int_0^\infty \mathbb{1}_{\{W_1 < \frac{k}{n}x, V_1 < \frac{ky}{n}\}} dx^{-\gamma_1} = -\mathbb{1}_{\{V_1 < \frac{ky}{n}\}} \left(\frac{n}{k} W_1\right)^{-\gamma_1}.$$

As a result, by (12) and a change of variable, we obtain

$$\begin{aligned} \mathbb{E}[A_n(y, \rho)] &= \left(\frac{n}{k}\right)^{1/\eta} \mathbb{E} \left[ \mathbb{1}_{\{V_1 < \frac{ky}{n}\}} \left(\frac{n}{k} W_1\right)^{-\rho\gamma_1} \right] \\ &= \left(\frac{n}{k}\right)^{1/\eta} \int_0^\infty \mathbb{P} \left[ W_1 < \frac{k}{n}x^{-\frac{1}{\rho\gamma_1}}, V_1 < \frac{ky}{n} \right] dx \\ &= -\left(\frac{n}{k}\right)^{1/\eta} \int_0^\infty \mathbb{P} \left[ W_1 < \frac{k}{n}x, V_1 < \frac{ky}{n} \right] dx^{-\rho\gamma_1} \\ &= -\int_0^\infty c_{\frac{n}{k}}^{\frac{n}{k}} \left(x \wedge \frac{n}{k}, y\right) dx^{-\rho\gamma_1}. \end{aligned}$$

The statement follows from (A2). □

*Proof of Proposition.* For  $i = 1, \dots, n$ , let  $W_i = 1 - F_1(X_i)$  and  $V_i = 1 - F_2(Y_i)$ . We write

$$\begin{aligned} \tilde{\theta}_{\frac{ky}{n}} &= \frac{1}{ky} \sum_{i=1}^n \int_0^\infty \mathbb{1}_{\{X_i > x, Y_i > U_2(\frac{n}{ky})\}} dx \\ &= -\frac{U_1(n/k)}{ky} \sum_{i=1}^n \int_0^\infty \mathbb{1}_{\{W_i < \frac{k}{n}s_n(x), V_i < \frac{k}{n}y\}} dx^{-\gamma_1}, \end{aligned}$$

where  $s_n(x)$  is defined in (16).

Similarly, we have

$$\theta_{\frac{ky}{n}} = -\frac{nU_1(n/k)}{ky} \int_0^\infty \mathbb{P} \left( W_1 < \frac{k}{n}s_n(x), V_1 < \frac{k}{n}y \right) dx^{-\gamma_1}. \tag{A12}$$

This means that  $\mathbb{E}[\tilde{\theta}_{\frac{ky}{n}}] = \theta_{\frac{ky}{n}}$ , and it enables us to write the left-hand side of (A1) as

$$\frac{S_n}{U_1(n/k)} \left( \tilde{\theta}_{\frac{ky}{n}} - \theta_{\frac{ky}{n}} \right) = \sum_{i=1}^n \left( Z_{n,i}^*(y) - \mathbb{E} \left[ Z_{n,i}^*(y) \right] \right),$$

where

$$Z_{n,i}^*(y) = -\frac{S_n}{ky} \int_0^\infty \mathbb{1}_{\{W_i < \frac{k}{n}s_n(x), V_i < \frac{k}{n}y\}} dx^{-\gamma_1}. \tag{A13}$$

Recall that we have  $\lim_{n \rightarrow \infty} s_n(x) = x$  by the regular variation of  $1 - F_1$ . We shall study a simpler process obtained by replacing  $s_n(x)$  with  $x$  in (A14):

$$Z_{n,i}(y) = -\frac{S_n}{ky} \int_0^\infty \mathbb{1}_{\{W_i < \frac{k}{n}x, V_i < \frac{k}{n}y\}} dx^{-\gamma_1}. \tag{A14}$$

To prove (A1), it suffices to show that

$$\sup_{y \in [1/2, 2]} n \mathbb{E} \left[ \left| Z_{n,1}^*(y) - Z_{n,1}(y) \right| \right] \rightarrow 0, \tag{A15}$$

and

$$\left\{ \sum_{i=1}^n \left( Z_{n,i}(y) - \mathbb{E}[Z_{n,i}(y)] \right) \right\}_{y \in [1/2, 2]} \xrightarrow{d} \{W(y)\}_{y \in [1/2, 2]}. \tag{A16}$$

Note that (A15) implies that

$$\sup_{y \in [1/2, 2]} \sum_{i=1}^n \left( Z_{n,i}^*(y) - Z_{n,i}(y) \right) \xrightarrow{\mathbb{P}} 0$$

and

$$\sup_{y \in [1/2, 2]} \sum_{i=1}^n \left( \mathbb{E}[Z_{n,i}^*(y)] - \mathbb{E}[Z_{n,i}(y)] \right) \xrightarrow{\mathbb{P}} 0.$$

### Step 1: Proof of (A15)

Using the definitions and the triangular inequality, we write

$$\begin{aligned} & n\mathbb{E} \left[ \left| Z_{n,1}^* - Z_{n,1} \right| \right] \\ &= - \left( \frac{n}{k} \right)^{\frac{1}{2n} - \frac{1}{2}} \frac{n}{\sqrt{ky}} \int_0^\infty \mathbb{P} \left( \frac{k}{n} (x \wedge s_n(x)) < W_1 < \frac{k}{n} (x \vee s_n(x)), V_1 < \frac{k}{n} y \right) dx^{-\gamma_1} \\ &= - \frac{T_n}{y} \int_0^\infty \left[ c_{\frac{n}{k}} \left( \left( x \wedge \frac{n}{k} \right) \vee s_n(x), y \right) - c_{\frac{n}{k}} (x \wedge s_n(x), y) \right] dx^{-\gamma_1} \\ &\leq - \frac{T_n}{y} \int_0^\infty |c(x, y) - c(s_n(x), y)| dx^{-\gamma_1} \\ &\quad - \frac{T_n}{y} \int_0^\infty \left| c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, y \right) - c(x, y) \right| dx^{-\gamma_1} \\ &\quad - \frac{T_n}{y} \int_0^\infty \left| c_{\frac{n}{k}} (s_n(x), y) - c(s_n(x), y) \right| dx^{-\gamma_1}. \end{aligned}$$

All three terms in the left-hand side converge to zero by (iii) and (v) in Lemma 1.

### Step 2: Proof of (A16)

We aim to apply Theorem 2.11.9 in van der Vaart and Wellner (1996). We will prove that the four conditions of this theorem are satisfied. Here,  $(\mathcal{F}, \rho) = \{[1/2, 2], \rho(y_1, y_2) = |y_1 - y_2|\}$ , and  $\|Z\|_{\mathcal{F}} = \sup_{y \in \mathcal{F}} |Z(y)|$ .

(a) Fix  $\epsilon > 0$ . Using that  $\|Z_{n,1}\|_{\mathcal{F}} \leq 4Z_{n,1}(2)$ , we get the following, with  $\bar{\delta}$  as defined in Assumption B(1),

$$\begin{aligned} n\mathbb{E} \left[ \|Z_{n,1}\|_{\mathcal{F}} \mathbb{1}_{\{\|Z_{n,1}\|_{\mathcal{F}} > \epsilon\}} \right] &\leq 4n\mathbb{E} \left[ Z_{n,1}(2) \mathbb{1}_{\{Z_{n,1}(2) > \epsilon\}} \right] \\ &\leq \frac{4n}{\epsilon^{1+\bar{\delta}}} \mathbb{E} \left[ Z_{n,1}^{2+\bar{\delta}}(2) \right] \\ &= \frac{n}{\epsilon^{1+\bar{\delta}}} \frac{S_n^{2+\bar{\delta}}}{2^{\bar{\delta}} k^{2+\bar{\delta}}} \mathbb{E} \left[ \left( - \int_0^\infty \mathbb{1}_{\{W_i < \frac{k}{n} x, V_i < \frac{2k}{n}\}} dx^{-\gamma_1} \right)^{2+\bar{\delta}} \right] \\ &= \frac{1}{\epsilon^{1+\bar{\delta}} 2^{\bar{\delta}}} T_n^{-\bar{\delta}} \mathbb{E} \left[ \left( \frac{n}{k} \right)^{1/\eta} \left( - \int_0^\infty \mathbb{1}_{\{W_i < \frac{k}{n} x, V_i < \frac{2k}{n}\}} dx^{-\gamma_1} \right)^{2+\bar{\delta}} \right] \\ &\rightarrow 0. \end{aligned} \tag{A17}$$

The last convergence follows from  $T_n \rightarrow \infty$  and Lemma 2.



(b) Take a sequence  $\delta_n \rightarrow 0$ . Then, by the triangular inequality and that  $y_1, y_2 \geq 1/2$ , it follows that

$$\begin{aligned} & \sup_{|y_1 - y_2| < \delta_n} \sum_{i=1}^n \mathbb{E} [(Z_{n,i}(y_1) - Z_{n,i}(y_2))^2] \\ & \leq 4 \sup_{|y_1 - y_2| < \delta_n} \sum_{i=1}^n \left(\frac{n}{k}\right)^{\frac{1}{\eta}-1} \frac{1}{k} \mathbb{E} \left[ \left( \int_0^\infty \mathbb{1}_{\{W_i < \frac{k}{n}x, \frac{k}{n}y_2 < V_i < \frac{k}{n}y_1\}} dx^{-\gamma_1} \right)^2 \right] \\ & \quad + \sup_{|y_1 - y_2| < \delta_n} \left| \frac{1}{y_2^2} - \frac{1}{y_1^2} \right| \mathbb{E} \left[ \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \left( \int_0^\infty \mathbb{1}_{\{W_1 < \frac{k}{n}x, V_1 < \frac{k}{n}y_1\}} dx^{-\gamma_1} \right)^2 \right] \\ & = 4 \sup_{|y_1 - y_2| < \delta_n} \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \mathbb{E} \left[ \mathbb{1}_{\{\frac{k}{n}y_1 < V_1 < \frac{k}{n}y_2\}} \left(\frac{n}{k}W_1\right)^{-2\gamma_1} \right] + \sup_{|y_1 - y_2| < \delta_n} \left| \frac{1}{y_2^2} - \frac{1}{y_1^2} \right| \mathbb{E}[A_n(y_1, 2)], \end{aligned} \tag{A18}$$

where  $A_n(y_1, 2)$  is defined as in Lemma 2. Thus, the second summand converges to zero because  $\lim_{n \rightarrow \infty} \mathbb{E}[A_n(y_1, 2)] < \infty$ ,  $\delta_n \rightarrow 0$ , and  $y_1, y_2 \geq 1/2$ . Moreover, by the triangular inequality and by (ii) and (iv) in Lemma 1,

$$\begin{aligned} & \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \mathbb{E} \left[ \mathbb{1}_{\{\frac{k}{n}y < V_1 < \frac{k}{n}(y + \delta_n)\}} \left(\frac{n}{k}W_1\right)^{-2\gamma_1} \right] \\ & = -\left(\frac{n}{k}\right)^{\frac{1}{\eta}} \int_0^\infty \mathbb{P} \left( \frac{k}{n}y < V_1 < \frac{k}{n}(y + \delta_n), W_1 < \frac{k}{n}x \right) dx^{-2\gamma_1} \\ & = \left| \int_0^\infty c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, y + \delta_n \right) dx^{-2\gamma_1} - \int_0^\infty c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, y \right) dx^{-2\gamma_1} \right| \rightarrow 0. \end{aligned} \tag{A19}$$

(c) Let  $N_{\square}(\epsilon, \mathcal{F}, L_2^n)$  be the minimal number of sets  $N_\epsilon$  in a partition  $[1/2, 2] = \cup_{j=1}^{N_\epsilon} I_{n,j}^\epsilon$  such that

$$\sum_{i=1}^n \mathbb{E} \left[ \sup_{y_1, y_2 \in I_{n,j}^\epsilon} |Z_{n,i}(y_1) - Z_{n,i}(y_2)|^2 \right] \leq \epsilon^2, \quad \forall j = 1, \dots, N_\epsilon.$$

Consider the partition given by  $I_{n,j}^\epsilon = [1/2 + (j - 1)\Delta_n, 1/2 + j\Delta_n]$ . Then,  $N_\epsilon = 3/2\Delta_n$ . We aim at finding  $\Delta_n$  such that, for every sequence  $\delta_n \rightarrow 0$ , it holds

$$\int_0^{\delta_n} \sqrt{\log N_{\square}(\epsilon, \mathcal{F}, L_2^n)} d\epsilon \rightarrow 0.$$

By the same reasoning for (A18), we obtain

$$\begin{aligned} & n \mathbb{E} \left[ \sup_{y_1, y_2 \in I_{n,j}^\epsilon} |Z_{n,1}(y_1) - Z_{n,1}(y_2)|^2 \right] \\ & \leq \sup_{y_1, y_2 \in I_{n,j}^\epsilon} \left| \frac{1}{y_1^2} - \frac{1}{y_2^2} \right| \mathbb{E} A_n(y_1, 2) + 4 \sup_{y_1, y_2 \in I_{n,j}^\epsilon} \left(\frac{n}{k}\right)^{\frac{1}{\eta}} \mathbb{E} \left[ \mathbb{1}_{\{\frac{k}{n}y_1 < V_1 < \frac{k}{n}y_2\}} \left(\frac{n}{k}W_1\right)^{-2\gamma_1} \right] \\ & =: B_n + C_n. \end{aligned}$$

For the first term, we have  $B_n \leq K_1 \Delta_n$  for some constant  $K_1 > 0$  by Lemma 2. Let  $\bar{y}_1 = 1/2 + (j - 1)\Delta_n$  and  $\bar{y}_2 = 1/2 + j\Delta_n$ . Next, we derive two different upper bounds for  $C_n$ . First,

by Holder inequality, we obtain

$$\begin{aligned} C_n &\leq 4 \left(\frac{n}{k}\right)^{\frac{1}{q}} \mathbb{E} \left[ \mathbb{1}_{\left\{ \frac{k}{n} \bar{y}_1 < V_1 < \frac{k}{n} \bar{y}_2 \right\}} \left( W_1 \frac{n}{k} \right)^{-2\gamma_1} \right] \\ &\leq 4 \left(\frac{n}{k}\right)^{\frac{1}{q}} \mathbb{E} \left[ \mathbb{1}_{\left\{ \frac{k}{n} \bar{y}_1 < V_1 < \frac{k}{n} \bar{y}_2 \right\}} \right]^{1/p} \mathbb{E} \left[ \mathbb{1}_{\left\{ V_1 < \frac{k}{n} \bar{y}_2 \right\}} \left( W_1 \frac{n}{k} \right)^{-2q\gamma_1} \right]^{1/q} \\ &\leq 4 \left(\frac{n}{k}\right)^{\frac{1}{q} - \frac{1}{p} - \frac{1}{nq}} |\bar{y}_1 - \bar{y}_2|^{\frac{1}{p}} \mathbb{E}[A_n(\bar{y}_2, 2q)] \\ &= K_2 \left(\frac{n}{k}\right)^{\frac{1}{q} - \frac{1}{p} - \frac{1}{nq}} \Delta_n^{\frac{1}{p}}, \end{aligned}$$

for some constant  $K_2$ . The last equality is obtained by applying Lemma 2 and choosing  $q = (2 + \bar{\delta})/2$  and  $1/p + 1/q = 1$ .

Second, by the same reasoning for (A19), the triangular inequality, and (ii) and (iv) in Lemma 1, we get a second bound on  $C_n$ :

$$\begin{aligned} C_n &\leq -8 \int_0^\infty \left[ c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, \bar{y}_2 \right) - c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, \bar{y}_1 \right) \right] dx^{-2\gamma_1} \\ &= -8 \int_0^\infty [c(x, \bar{y}_2) - c(x, \bar{y}_1)] dx^{-2\gamma_1} \\ &\quad - 8 \int_0^\infty \left[ c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, \bar{y}_2 \right) - c(x, \bar{y}_2) \right] dx^{-2\gamma_1} \\ &\quad - 8 \int_0^\infty \left[ c(x, \bar{y}_1) - c_{\frac{n}{k}} \left( x \wedge \frac{n}{k}, \bar{y}_1 \right) \right] dx^{-2\gamma_1} \\ &\leq K_3 \Delta_n + K_4 \left(\frac{n}{k}\right)^\tau \end{aligned}$$

for some constants  $K_3$  and  $K_4$ .

If  $\epsilon^2 < \left(\frac{n}{k}\right)^{\tau^*}$  for some  $\tau^* \in (\tau, 0)$ , we use the first bound on  $C_n$ , that is,

$$B_n + C_n \leq 2K_2 \left(\frac{n}{k}\right)^{\frac{1}{q} - \frac{1}{p} - \frac{1}{nq}} \Delta_n^{\frac{1}{p}},$$

and by choosing

$$\Delta_n = (2K_2)^{-p} \left(\frac{n}{k}\right)^{-\frac{p}{q} + 1 + \frac{p}{nq}} \epsilon^{2p},$$

we get  $B_n + C_n \leq \epsilon^2$ . Hence,

$$N_\epsilon \leq \frac{3(2K_2)^p}{\epsilon^{2p}} \left(\frac{n}{k}\right)^{\frac{p}{q} - 1 - \frac{p}{nq}}.$$

Otherwise, if  $\epsilon^2 > \left(\frac{n}{k}\right)^{\tau^*}$ , for sufficiently large  $n$ ,

$$K_4 \left(\frac{n}{k}\right)^\tau < \frac{1}{2} \left(\frac{n}{k}\right)^{\tau^*} < \frac{1}{2} \epsilon^2,$$

and we use the second bound on  $C_n$  with

$$\Delta_n = \frac{\epsilon^2}{2(K_1 + K_3)},$$

that is, we get

$$B_n + C_n \leq (K_1 + K_3) \Delta_n + K_4 \left(\frac{n}{k}\right)^\tau \leq \epsilon^2.$$

Hence, in this case,

$$N_\epsilon \leq \frac{3(K_1 + K_3)}{\epsilon^2}.$$

Now, we distinguish between two cases. If  $\delta_n \sqrt{\log(n/k)} \rightarrow 0$ , using  $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$  and  $\log(x) \leq x$  for large  $x$ , we get

$$\begin{aligned} \int_0^{\delta_n} \sqrt{\log N_{\square}(\epsilon, \mathcal{F}, L_2^n)} d\epsilon &\leq \int_0^{\delta_n} \sqrt{\left(\frac{p}{\eta} - 1 - \frac{p}{\eta q}\right) \log(n/k) + 2p \log \epsilon^{-1} + \log 3(2K_2)^p} d\epsilon \\ &\leq K \left( \int_0^{\delta_n} \sqrt{\log(n/k)} d\epsilon + \int_0^{\delta_n} \sqrt{\epsilon^{-1}} d\epsilon \right), \end{aligned}$$

and the left-hand side converges to zero as  $\delta_n \rightarrow 0$ .

On the other hand, if  $\delta_n \sqrt{\log(n/k)} \not\rightarrow 0$ , take  $\delta_n^* = (n/k)^{\tau^*}$ . Note that  $\delta_n^* \sqrt{\log(n/k)} \rightarrow 0$ . Hence, we write

$$\begin{aligned} \int_0^{\delta_n} \sqrt{\log N_{\square}(\epsilon, \mathcal{F}, L_2^n)} d\epsilon &= \int_0^{\delta_n^*} \sqrt{\log N_{\square}(\epsilon, \mathcal{F}, L_2^n)} d\epsilon + \int_{\delta_n^*}^{\delta_n} \sqrt{\log N_{\square}(\epsilon, \mathcal{F}, L_2^n)} d\epsilon \\ &\leq o(1) + \int_{\delta_n^*}^{\delta_n} \sqrt{\log(3(K_1 + K_3)/\epsilon^2)} d\epsilon \\ &\leq o(1) + \sqrt{2} \int_0^{\delta_n^*} \sqrt{\epsilon^{-1}} d\epsilon \rightarrow 0. \end{aligned}$$

(d) We have to show that the marginals converge, that is, for each  $M \in \mathbb{N}$  and for each  $y_1, \dots, y_M \in [1/2, 2]$ , the random vector

$$\left( \sum_{i=1}^n (Z_{n,i}(y_1) - \mathbb{E}[Z_{n,i}(y_1)]), \dots, \sum_{i=1}^n (Z_{n,i}(y_M) - \mathbb{E}[Z_{n,i}(y_M)]) \right)$$

converges to a multivariate normal distribution. It suffices to show that, for each  $a_1, \dots, a_M \in \mathbb{R}$ , we have

$$\sum_{j=1}^M a_j \left[ \sum_{i=1}^n (Z_{n,i}(y_j) - \mathbb{E}[Z_{n,i}(y_j)]) \right] =: \sum_{i=1}^n (N_{n,i} - \mathbb{E}[N_{n,i}])$$

converging to a normal distribution, where  $N_{n,i} = \sum_{j=1}^M a_j Z_{n,i}(y_j)$ . This will follow from the Lindeberg–Feller central limit theorem (see, e.g., van der Vaart, 1998), once we show that, for each  $\epsilon > 0$ ,

$$\sum_{i=1}^n \mathbb{E} [ |N_{n,i}|^2 \mathbf{1}_{\{|N_{n,i}| > \epsilon\}} ] \rightarrow 0 \tag{A20}$$

and

$$\sum_{i=1}^n \text{Var} (N_{n,i}) \rightarrow \sigma_N^2. \tag{A21}$$

We proceed with (A20). First,

$$\begin{aligned} \sum_{i=1}^n \mathbb{E} [ |N_{n,i}|^2 \mathbf{1}_{\{|N_{n,i}| > \epsilon\}} ] &= n \mathbb{E} [ |N_{n,1}|^2 \mathbf{1}_{\{|N_{n,1}| > \epsilon\}} ] \\ &\leq \frac{n \mathbb{E} [ |N_{n,1}|^{2+\delta} ]}{\epsilon^\delta} \leq Kn \sum_{j=1}^M |a_j|^{2+\delta} \frac{\mathbb{E} [ |Z_{n,1}(2)|^{2+\delta} ]}{\epsilon^\delta}, \end{aligned}$$

which converges to zero by (A17). For (A21), we write

$$\begin{aligned} \sum_{i=1}^n \text{Var}(N_{n,i}) &= n \left\{ \mathbb{E} \left[ \left( \sum_{j=1}^M a_j Z_{n,1}(y_j) \right)^2 \right] - \left( \mathbb{E} \left[ \sum_{j=1}^M a_j Z_{n,1}(y_j) \right] \right)^2 \right\} \\ &= n \mathbb{E} \left[ \sum_{j,k=1}^M a_j a_k Z_{n,1}(y_j) Z_{n,1}(y_k) \right] - \left( \sqrt{n} \sum_{j=1}^M a_j \mathbb{E} [Z_{n,1}(y_j)] \right)^2 \\ &= n \sum_{j,k=1}^M a_j a_k \mathbb{E} [Z_{n,1}(y_j) Z_{n,1}(y_k)] + o(1) \end{aligned}$$

because it is easy to check that  $\sqrt{n} \mathbb{E} [Z_{n,1}(y_j)] \rightarrow 0$ , for  $j = 1, \dots, M$ . Moreover, observe that

$$\begin{aligned} n \mathbb{E} [Z_{n,1}(y_j) Z_{n,1}(y_k)] &= \left( \frac{n}{k} \right)^{\frac{1}{\eta}} \frac{1}{y_j y_k} \mathbb{E} \left[ \left( \int_0^\infty \mathbb{1}_{\{U_1 < \frac{k}{n} x, V_1 < \frac{k}{n} (y_j \wedge y_k)\}} dx^{-\gamma_1} \right)^2 \right] \\ &= \frac{1}{y_j y_k} \mathbb{E}[A_n(y_j \wedge y_k, 2)]. \end{aligned}$$

Thus, by Lemma 2, it follows that (A21) holds with

$$\sigma_N^2 = - \sum_{j,k=1}^M \frac{a_j a_k}{y_j y_k} \int_0^\infty c(x, y_j \wedge y_k) dx^{-2\gamma_1}. \quad (\text{A22})$$

We have now verified the four conditions required by Theorem 2.11.9 in van der Vaart and Wellner (1996), which leads to the conclusion that  $\sum_{i=1}^n (Z_{n,i} - \mathbb{E}[Z_{n,i}])$  converges in distribution to a Gaussian process. Finally, we compute the covariance structure of the limit process. For each  $y_1, y_2 \in [1/2, 2]$ , by independence, we have

$$\begin{aligned} \mathbb{E}[W(y_1)W(y_2)] &= \lim_{n \rightarrow \infty} \text{Cov} \left( \sum_{i=1}^n Z_{n,i}(y_1), \sum_{i=1}^n Z_{n,i}(y_2) \right) \\ &= \lim_{n \rightarrow \infty} n \text{Cov}(Z_{n,1}(y_1), Z_{n,1}(y_2)) \\ &= \lim_{n \rightarrow \infty} (n \mathbb{E} [Z_{n,1}(y_1) Z_{n,1}(y_2)] - n \mathbb{E} [Z_{n,1}(y_1)] \mathbb{E} [Z_{n,1}(y_2)]) \\ &= - \frac{1}{y_1 y_2} \int_0^\infty c(x, y_1 \wedge y_2) dx^{-2\gamma_1} \\ &= \frac{1}{y_1 y_2} \int_0^\infty c \left( x^{-\frac{1}{2\gamma_1}}, y_1 \wedge y_2 \right) dx. \end{aligned}$$

The fourth equality follows the same reasoning as that for (A22).  $\square$

*Proof of Proposition 2.* Note that the convergence speed in this proposition is  $\frac{S_n}{U_1(n/k)}$ , which is the same as that in Proposition 3. By definition,

$$\hat{\theta}_{\frac{k}{n}} = e_n \tilde{\theta}_{\frac{ke_n}{n}}.$$

Hence,

$$\begin{aligned} \frac{S_n}{U_1(n/k)} \left( \hat{\theta}_{\frac{k}{n}} - \theta_{\frac{k}{n}} \right) &= \frac{S_n}{U_1(n/k)} e_n \left( \tilde{\theta}_{\frac{ke_n}{n}} - \theta_{\frac{ke_n}{n}} \right) + \frac{S_n}{U_1(n/k)} \left( e_n \theta_{\frac{ke_n}{n}} - \theta_{\frac{k}{n}} \right) \\ &=: T_1 + T_2. \end{aligned} \quad (\text{A23})$$

First, we show that  $T_1 \xrightarrow{d} W(1)$ . We start by writing

$$T_1 = e_n \left[ \frac{S_n}{U_1(n/k)} (\tilde{\theta}_{\frac{ke_n}{n}} - \theta_{\frac{ke_n}{n}}) - W(e_n) \right] + e_n W(e_n).$$

Because  $e_n \xrightarrow{\mathbb{P}} 1$ , the first term of the right-hand side, with probability tending to one, is bounded by

$$2 \sup_{y \in [1/2, 2]} \left| \frac{S_n}{U_1(n/k)} (\tilde{\theta}_{\frac{ky}{n}} - \theta_{\frac{ky}{n}}) - W(y) \right|,$$

which is  $o_p(1)$  by Proposition 3 and the continuous mapping theorem. Moreover, by Corollary 1.11 in Adler (1990) and  $e_n \xrightarrow{\mathbb{P}} 1$ ,  $W(e_n) \xrightarrow{d} W(1)$ . Thus,  $T_1 \xrightarrow{d} W(1)$ .

Using (A12), we can write

$$\frac{S_n}{U_1(n/k)} \theta_{\frac{ky}{n}} = -\frac{T_n}{y} \int_0^\infty c_{\frac{n}{k}}(s_n(x), y) dx^{-\gamma_1}.$$

Thus,  $T_2$  can be rewritten as follows:

$$\begin{aligned} T_2 &= T_n \int_0^\infty \left\{ c_{\frac{n}{k}}(s_n(x), e_n) - c_{\frac{n}{k}}(s_n(x), 1) \right\} dx^{-\gamma_1} \\ &= T_n \int_0^\infty \left\{ c_{\frac{n}{k}}(s_n(x), e_n) - c(s_n(x), e_n) \right\} dx^{-\gamma_1} \\ &\quad + T_n \int_0^\infty \left\{ c(s_n(x), e_n) - c(s_n(x), 1) \right\} dx^{-\gamma_1} \\ &\quad + T_n \int_0^\infty \left\{ c(s_n(x), 1) - c_{\frac{n}{k}}(s_n(x), 1) \right\} dx^{-\gamma_1} \\ &= o_p(1) + T_{21} + o(1). \end{aligned}$$

The last equality follows from the fact that  $e_n \xrightarrow{P} 1$  and (A4). Furthermore, we can decompose  $T_{21}$  into three terms as follows:

$$\begin{aligned} T_{21} &= T_n \int_0^\infty \left\{ c(s_n(x), e_n) - c(x, e_n) \right\} dx^{-\gamma_1} + T_n \int_0^\infty \left\{ c(x, 1) - c(s_n(x), 1) \right\} dx^{-\gamma_1} \\ &\quad + T_n \int_0^\infty \left\{ c(x, e_n) - c(x, 1) \right\} dx^{-\gamma_1} \\ &= o_p(1) + o(1) + O_P(T_n |e_n - 1|), \end{aligned}$$

by (ii) and (iii) in Lemma 1. Finally,

$$T_n |e_n - 1| = \left( \frac{n}{k} \right)^{-\frac{1}{2\eta} + \frac{1}{2}} \sqrt{k} |e_n - 1| \rightarrow 0$$

because  $\sqrt{k} |e_n - 1| = O_P(1)$ ; see (26) in Cai et al. (2015). Consequently,  $T_2 = o_p(1)$  and it has no contribution in the limit distribution. □