



**HAL**  
open science

# On the study of extremes with dependent random right-censoring

Gilles Stupfler

► **To cite this version:**

Gilles Stupfler. On the study of extremes with dependent random right-censoring. 2017. hal-01450775v1

**HAL Id: hal-01450775**

**<https://hal.science/hal-01450775v1>**

Preprint submitted on 31 Jan 2017 (v1), last revised 8 Jun 2018 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the study of extremes with dependent random right-censoring

Gilles Stupfler

School of Mathematical Sciences, The University of Nottingham, University Park,  
Nottingham NG7 2RD, United Kingdom

**Abstract.** The study of extremes in missing data frameworks is a recent developing field. In particular, the randomly right-censored case has been receiving a fair amount of attention in the last decade. All studies on this topic, however, work under the usual assumption that the variable of interest and the censoring variable are independent. In this paper, we look instead at the dependent case. More precisely, we examine the extremes of the response variable (that is, the minimum between time-to-event and time-to-censoring) and the probability of censoring in the right tail of this variable when the structure of the dependent censoring mechanism is given by an extreme value copula. Statistically speaking, our results have important consequences on the identifiability problem for the extremes of the variable of interest: most strikingly, and in contrast to the independent censoring case, if the censoring variable has a lighter tail than the variable of interest and the dependence structure is given by a non-independent extreme value copula, then the extremes of the variable of interest appear to be unrecoverable.

**AMS Subject Classifications:** 62N01, 62N02.

**Keywords:** Random right-censoring, dependence, extreme value copula, extreme value index, tail censoring probability.

## 1 Introduction

The problem of missing data, and in particular censoring, is frequently encountered in certain fields of statistical applications. The archetypal example of censoring is arguably the study of the survival times of patients to a given chronic disease in a medical follow-up study lasting up to a fixed time  $t$ .

If a patient is diagnosed with the disease at time  $s$ , then his/her survival time will be known if and only if he/she dies before time  $t$ . If this is not the case, then the only information available is that his/her survival time is not less than the censoring time  $t - s$ . In mathematical terms, the information available to the practitioner is the pair  $(Z, \delta)$ , where  $Z$  is the minimum between the survival time and censoring time, and  $\delta$  is the 0-1 variable equal to 1 if and only if the survival time is actually observed. This situation is one of the most frequent examples of random right-censoring, which shall be our framework in this paper.

Random right-censoring is also found in actuarial science: in non-life insurance, a claim available to a claim adjuster is usually right-censored by the policy limit of the contract, while in life insurance any study that monitors policyholders in a given time period contains right-censored data points, since many of the subjects still live at the end of the monitoring period. Another example is reliability data analysis: if a car company collects failure data during the warranty period, then a failure could happen not only because of a failure in the mechanics of the car, but also because of an accident or driver error. In the latter case, time-to-accident or time-to-driver-error should be treated as the censoring time, see Modarres *et al.* (2009). Random right-censoring should not be confused with other types of missing data mechanisms such as right-truncation, where no information is available at all when the random variable of interest is not actually observed: the earliest reference tackling this problem in the extreme value context is Gardes and Stupfler (2015).

In a random right-censoring framework, a stimulating problem is the estimation of extreme parameters of the underlying distribution of the variable of interest. In the aforementioned examples, this would mean analysing survival times of exceptionally strong/weak patients to a given disease, extreme losses/payouts in insurance, or failure times for highly resistant/unreliable devices. This subfield of extreme value statistics has received a good amount of attention in recent years: we refer to Beirlant *et al.* (2007), Einmahl *et al.* (2008), Beirlant *et al.* (2010), Gomes and Neves (2011), Ndao *et al.* (2014), Sayah *et al.* (2014), Worms and Worms (2014), Brahimy *et al.* (2015), Beirlant *et al.* (2016) and Stupfler (2016). All these papers work under the hypothesis that the variable of interest  $Y$  and the censoring variable  $T$  are independent random variables. Among others, this allows for a very convenient expression of the cumulative distribution function of the observed variable  $Z$ , as well as for a simple discussion of its extreme value properties, which leads to simple estimators of the extreme parameters of  $Y$ . It should be said that since the pioneering paper of Kaplan and Meier (1958) on the product-limit estimator for the survival function, the assumption of independent censoring is arguably the standard assumption in the context of random right-censoring.

And yet, cases where there are strong suspicions of dependence between the variable of interest and the

censoring time have been reported several times over the last decades. An early reference is Lagakos (1979). In medical studies especially, common causes of the probable violation of the independence hypothesis include a sizeable number of patient dropouts (Huang and Zhang, 2008 and Jackson *et al.*, 2014) and the presence of ignored but valuable covariate information (see Zeng, 2004 and Li *et al.*, 2007). Crucially, using traditional estimators such as the Kaplan-Meier estimator when there is dependence may yield to invalid inferences, see Fisher and Kanarek (1974), Klein and Moeschberger (1987), and the introduction of Ebrahimi and Molefe (2003). Moreover, there is the additional issue of identifiability, in the sense that if the dependence structure is not specified then the distribution of  $(Y, T)$  cannot be recovered from that of the pair  $(Z, \delta)$ , see Tsiatis (1975). A number of authors have suggested partial solutions to the problem: some recent efforts include working on covariate models (Zeng, 2004), fitting specific types of known copulas (Li *et al.*, 2007, Huang and Zhang, 2008), or assuming weaker assumptions than independence on the pair  $(X, Y)$  (Ebrahimi *et al.*, 2003).

One may therefore wonder what kind of effect dependence has on the estimation of extreme parameters of  $Y$ . As noted above, there are numerous ways to specify dependence; in this paper, we assume that the dependence structure of the pair  $(Y, T)$  is given by an extreme value copula, which is equivalent to assuming that  $(Y, T)$  has a bivariate extreme value distribution in the sense of Tawn (1988). The construction and early development of extreme value copulas date back to Galambos (1978) and Deheuvels (1984), and a recent account is Gudendorf and Segers (2010). This type of copula is particularly adapted to the description of joint extreme events, *i.e.* of situations when both  $Y$  and  $T$  are extreme, which is precisely the kind of events one has to consider in order to understand the extremes of the observed variable  $Z = \min(Y, T)$ , as this is the only way to get back to the extremes of  $Y$ . The focus of this paper is to, assuming standard extreme value conditions on the distribution of  $Y$  and  $T$  together with an extreme value copula dependence model, analyse the extreme value properties of  $Z$  first and then of  $\delta$  given that  $Z$  is large, the latter variable indicating how much censoring there is in the extremes of the sample. We will examine the consequences of our results on the identifiability of the extremes of  $Y$ : in particular, it shall be shown that when the dependence structure of  $(Y, T)$  is given by a non-independent extreme value copula, then the extremes of  $Y$  are exactly those of the observed variable  $Z$  when  $T$  has a heavier tail than  $Y$ , therefore making the problem trivial from the extreme value perspective. More worryingly from the practical point of view, we shall also show that the extremes of  $Y$  cannot be recovered when  $T$  has a lighter tail than  $Y$ , although we know from Einmahl *et al.* (2008) that this problem has a solution indeed in the independent censoring case.

Let us highlight that the basic assumption of a purely extreme value copula model may appear restrictive at first. It is indeed similar in spirit to assuming that, in the univariate case, the underlying distribution is a Generalised Extreme Value distribution (see de Haan and Ferreira, 2006). This as-

sumption shall nevertheless prove very useful in identifying several problems that arise when there is dependence in the censoring mechanism. The idea is that any issue appearing in the present context can only get worse in a more general model, such as, for instance, the Archimax copula model of Capéràa *et al.* (2000). Let us also point out that the very popular bivariate model of Ledford and Tawn (1996, 1997) would not be appropriate here, because this model assumes that  $Y$  and  $T$  are unit Fréchet distributed. To use this model in practice, one therefore has to transform  $Y$  and  $T$  to a unit Fréchet distribution, which implies that the distributions of  $Y$  and  $T$  have at least been estimated beforehand. This would be an issue in our setting, since we know that consistency of estimators of the distribution functions of  $Y$  and  $T$ , such as the Kaplan-Meier estimator, is a difficult problem when there is dependence in the censoring mechanism. In this paper, our focus is rather that nothing is known about the distributions of  $Y$  and  $T$ , and we would like to know whether the extreme value characteristics of  $Y$  can be recovered.

The outline of our paper is as follows. Section 2 gives further details about our assumptions and especially about our dependence framework. Section 3 gives our main results, first about the extremes of the observed variable  $Z$  and then about the tail censoring probability. Section 4 concludes with some statistical considerations. Proofs of the main results are deferred to Appendix A, and auxiliary results and their proofs are relegated to Appendix B.

## 2 Framework

We assume throughout that the variable of interest  $Y$  is partially unobserved, due to the existence of a right-censoring random variable  $T$ . In other words, we only observe the pair  $(Z, \delta)$ , where  $Z = \min(Y, T)$  and  $\delta = \mathbb{I}_{\{Y \leq T\}}$ . Contrary to the standard setup, we also assume that  $Y$  and  $T$  are *not independent*. We describe here the dependence structure of the pair  $(Y, T)$  by the means of a copula function. The key result in order to do so is Sklar's theorem (Sklar, 1959), which says that there exists a copula  $C$  with

$$\forall y, t \in \mathbb{R}, \mathbb{P}(Y \leq y, T \leq t) = C(F_Y(y), F_T(t)),$$

in which  $F_Y$  and  $F_T$  denote the respective cumulative distribution functions of  $Y$  and  $T$ . A copula  $C$  is simply, in our case, a bivariate distribution function of a pair of standard uniform random variables. We assumed here that  $Y$  and  $T$  are not independent, so that the copula  $C$  cannot be the independent copula  $(u, v) \mapsto uv$ .

Since we want to analyse the extremes of  $Y$ , we should do so in a relevant extreme value framework. The condition we shall introduce, on an arbitrary distribution function  $F$ , is one of the many equivalent versions of the classical extreme value condition saying that the distribution should belong to the

domain of attraction of some extreme value distribution. As we work in a randomly right-censored situation, it is convenient to write our extreme value condition in terms of the survival function, which leads to the following formulation (see Theorem 1.1.6 in de Haan and Ferreira, 2006):

**Condition  $\mathcal{C}_1(\gamma)$ :** There is  $\gamma \in \mathbb{R}$ , called the extreme value index of  $F$ , and a positive function  $a$  such that the distribution function  $F$  satisfies:

$$\lim_{t \uparrow \tau^*} \frac{1 - F(t + xa(t))}{1 - F(t)} = (1 + \gamma x)^{-1/\gamma}$$

for all  $x$  with  $1 + \gamma x > 0$ , where  $\tau^* = \sup\{x \in \mathbb{R}, F(x) < 1\}$  is the right endpoint of  $F$ .

Since the observed variable is  $Z = \min(Y, T)$ , it makes sense to assume that the distributions of  $Y$  and  $T$  can be included within this extreme value framework, and our main results will then be stated under the assumption that the distributions of  $Y$  and  $T$  satisfy conditions  $\mathcal{C}_1(\gamma_Y)$  and  $\mathcal{C}_1(\gamma_T)$  respectively. Because actually

$$\mathbb{P}(Z > z) = \mathbb{P}(\min(Y, T) > z) = \mathbb{P}(Y > z, T > z)$$

it follows that the study of the extremes of  $Z$  will require a study of the situation when  $Y$  and  $T$  are jointly large. A very convenient assumption on the copula  $C$  in this context is then to suppose that  $C$  is an extreme value copula (see *e.g.* Gudendorf and Segers, 2010):

$$\forall (u, v) \in (0, 1]^2 \setminus \{(1, 1)\}, C(u, v) = (uv)^{A(\log(v)/\log(uv))}$$

where  $A$  is the so-called Pickands dependence function related to  $C$ , *i.e.* it is a function that maps  $[0, 1]$  into  $[1/2, 1]$ , is convex and satisfies the inequalities  $\max(t, 1 - t) \leq A(t) \leq 1$  for all  $t \in [0, 1]$ . The function  $A$  characterises the copula  $C$ ; the case  $A(t) = 1$  corresponds to the independent copula  $C(u, v) = uv$  (which we therefore exclude), and the case  $A(t) = \max(t, 1 - t)$  is that of the perfect dependence copula  $C(u, v) = \min(u, v)$ . In theoretical terms, these copulae arise naturally as limiting copulae of suitably normalised sequences of componentwise maxima of independent and identically distributed bivariate pairs (Joe, 1997). In this sense, assuming that  $(Y, T)$  follows a bivariate extreme value distribution is analogous to, in a univariate context, assuming that the random variable of interest has a Generalised Extreme Value distribution. Working in such a context, which is a simplified version of the general case, can help to identify statistical issues, such as non-identifiability or inconsistency of certain estimators, that would arise in a more general context as well. Our plan here is precisely to reveal such problems and then draw a couple of statistical consequences of them.

Our first step is to, compared to the independent censoring case, quantify the influence that the dependence structure induced by  $C$  has on the random variable  $Z$ . We do this by using a simple identity linking the survival function  $\bar{F}_Z(z) = \mathbb{P}(Z > z)$  of  $Z$  to the survival functions  $\bar{F}_Y$  and  $\bar{F}_T$  of  $Y$  and  $T$  and to the copula  $C$ :

**Proposition 1.** *The survival function of  $Z$  is*

$$\bar{F}_Z(z) = \bar{F}_Y(z)\bar{F}_T(z) + \varphi(\bar{F}_Y(z), \bar{F}_T(z))$$

where the function  $\varphi$  is defined by

$$\forall u, v \in [0, 1], \varphi(u, v) = C(1 - u, 1 - v) - (1 - u)(1 - v).$$

In other words, we can write the survival function of  $Z$  as what it would be under independence of  $Y$  and  $T$ , plus the term  $\varphi(\bar{F}_Y(z), \bar{F}_T(z))$  measuring the effect that the dependence structure in  $C$  has on  $Z$ . Since the behaviour of  $\bar{F}_Y(z)\bar{F}_T(z)$  for large  $z$  is easy to analyse, Proposition 1 suggests that, to analyse the extremes of  $Z$ , it is enough to understand the behaviour of  $\varphi(\bar{F}_Y(z), \bar{F}_T(z))$  for large  $z$ .

We now explain clearly what “large  $z$ ” means here. Let  $\tau_Y$  and  $\tau_T$  be the right endpoints of  $Y$  and  $T$ , respectively. In the case when  $\tau_T < \tau_Y$  then the extremes of  $Y$  cannot be recovered because the distribution of  $Y$  is not identifiable past the point  $\tau_T$ . In the case  $\tau_Y < \tau_T$  the extremes of  $Y$  are, under a mild regularity assumption, those of  $Z$ , *i.e.* they can be recovered by standard techniques. The following result makes this statement precise.

**Proposition 2.** *Assume that  $\tau_Y < \tau_T$  and  $C$  is an extreme value copula whose Pickands dependence function is continuously differentiable on  $[0, 1]$ . Then  $\bar{F}_Z(z)/\bar{F}_Y(z)$  has a positive and finite limit as  $z \uparrow \tau_Z = \tau_Y$ .*

Note that outside of the extreme value copula model, this result is still true under the regularity condition on  $C$  that the limit

$$\lim_{u \downarrow 0} \frac{C(1 - u, 1 - v) - (1 - u)(1 - v)}{u}$$

exist and be nonnegative.

We therefore assume in what follows that  $Y$  and  $T$  have a common right endpoint  $\tau = \tau_Y = \tau_T$ . Note that distributions with positive extreme value index have an infinite right endpoint, while distributions with negative extreme value index have a finite right endpoint, see Theorem 1.2.1 in de Haan and Ferreira (2006). It therefore follows from our basic assumption  $\tau_Y = \tau_T$  that the extreme value indices  $\gamma_Y$  and  $\gamma_T$  should have the same sign.

We may now summarise our hypotheses about the joint behaviour of  $Y$  and  $T$  in the following condition:

**Condition  $(\mathcal{M})$ :**  $Y$  and  $T$  have a common right endpoint  $\tau = \tau_Y = \tau_T$  and their joint distribution function is given by

$$\forall y, t \in \mathbb{R}, \mathbb{P}(Y \leq y, T \leq t) = C(F_Y(y), F_T(t))$$

where  $C$  is an extreme value copula whose Pickands dependence function  $A$  is not equal to the constant function 1.

In their pioneering paper, Einmahl *et al.* (2008) use the right-tail behaviour of the observed variable  $Z$ , and later an analysis of the censoring probability given that  $Z$  is large, to show that the extremes of  $Y$  can be recovered from those of  $(Z, \delta)$  in a very simple manner under classical extreme value assumptions. All later works on randomly right-censored extremes follow this line of thought to some extent. Our aim in this paper is to carry out a similar study and see in particular what influence the introduction of dependence in the censoring mechanism has on the identifiability of the extremes of  $Y$ . In other words, we shall try to obtain conditions under which we know whether the problem of recovering the extremes of  $Y$  from  $(Z, \delta)$  has a solution or not.

### 3 Main results

#### 3.1 The extremes of the response variable

Our first step is to analyse the behaviour of  $Z$  in its extremes, in terms of the extreme value indices  $\gamma_Y$  and  $\gamma_T$  of  $Y$  and  $T$ . This is straightforward in the independent censoring case, because then the survival function of  $Z$  is the product of those of  $Y$  and  $T$ . Our aim here is to state an analogue result under dependence model  $(\mathcal{M})$ .

**Theorem 1.** *Assume that condition  $(\mathcal{M})$  holds, that  $A$  is continuously differentiable on  $[0, 1]$ , and suppose that the distributions of  $Y$  and  $T$  satisfy conditions  $\mathcal{C}_1(\gamma_Y)$  and  $\mathcal{C}_1(\gamma_T)$  respectively; if moreover  $\gamma_Y = \gamma_T$ , assume that the ratio  $\bar{F}_T(z)/\bar{F}_Y(z)$  has a finite and positive limit as  $z \uparrow \tau$ . We have that:*

- (i) *If  $\gamma_Y, \gamma_T \geq 0$  with  $\gamma_Y + \gamma_T > 0$ , then  $Z$  satisfies condition  $\mathcal{C}_1(\min(\gamma_Y, \gamma_T))$ .*
- (ii) *If  $\gamma_Y, \gamma_T \leq 0$  with  $\gamma_Y + \gamma_T < 0$ , then  $Z$  satisfies condition  $\mathcal{C}_1(\max(\gamma_Y, \gamma_T))$ .*

The extreme behaviour of  $Z$  is therefore that of the variable with the lightest tail in the pair  $(Y, T)$ , *i.e.* that of  $Y$  in the case when  $Y$  has a lighter tail than (or a similar tail to that of)  $T$ , or that of  $T$  when  $T$  has a lighter tail than  $Y$ . Before drawing some interesting consequences from Theorem 1, let us consider a simple illustrative example.

**Example 1.** Suppose that a random variable  $Y$  has a unit Pareto distribution, so that  $F_Y(y) = 1 - y^{-1}$  for  $y > 1$ . Suppose also that the censoring variable  $T$  has a Pareto distribution with tail index  $\gamma_T$ , that is,  $F_T(t) = 1 - t^{-1/\gamma_T}$  for  $t > 1$ . Finally, assume that the dependence structure of the pair  $(Y, T)$  is described by a logistic copula (also called Gumbel-Hougaard copula; see Gumbel, 1960):

$$\forall u, v \in (0, 1], C_\theta(u, v) = \exp \left\{ - \left[ (-\log u)^\theta + (-\log v)^\theta \right]^{1/\theta} \right\}$$



where  $\theta \geq 1$  is a constant; here we choose  $\theta > 1$ , in order to ensure that  $Y$  and  $T$  are not independent. Theorem 1 predicts that  $\bar{F}_Z$  should be regularly varying with index  $\min(1, \gamma_T)$ ; we check this by analysing the asymptotic behaviour, as  $z \uparrow \infty$ , of

$$C_\theta(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z)) = \exp \left\{ - \left[ (-\log(1 - z^{-1}))^\theta + (-\log(1 - z^{-1/\gamma_T}))^\theta \right]^{1/\theta} \right\}.$$

If  $0 < \gamma_T < 1$  then

$$\begin{aligned} \frac{C_\theta(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z))}{1 - z^{-1}} &= \exp \left\{ \log(1 - z^{-1}) \left( \left[ 1 + \left( \frac{\log(1 - z^{-1/\gamma_T})}{\log(1 - z^{-1})} \right)^\theta \right]^{1/\theta} - 1 \right) \right\} \\ &= \exp \left\{ -\frac{1}{\theta} z^{\theta(1-\gamma_T^{-1})-1} (1 + o(1)) \right\} \\ &= 1 + o\left(z^{-1/\gamma_T}\right) \quad \text{because } \theta > 1 \text{ and } \gamma_T < 1. \end{aligned}$$

If  $\gamma_T = 1$  then

$$\frac{C_\theta(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z))}{1 - z^{-1}} = \exp \left\{ \log(1 - z^{-1}) (2^{1/\theta} - 1) \right\} = 1 - (2^{1/\theta} - 1)z^{-1}(1 + o(1)).$$

If  $\gamma_T > 1$  then

$$\begin{aligned} \frac{C_\theta(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z))}{1 - z^{-1/\gamma_T}} &= \exp \left\{ \log(1 - z^{-1/\gamma_T}) \left( \left[ 1 + \left( \frac{\log(1 - z^{-1})}{\log(1 - z^{-1/\gamma_T})} \right)^\theta \right]^{1/\theta} - 1 \right) \right\} \\ &= \exp \left\{ -\frac{1}{\theta} z^{\theta(\gamma_T^{-1}-1)-\gamma_T^{-1}} (1 + o(1)) \right\} \\ &= 1 + o\left(z^{-1}\right) \quad \text{because } \theta > 1 \text{ and } \gamma_T > 1. \end{aligned}$$

Together with the identity

$$\begin{aligned} \bar{F}_Z(z) &= \bar{F}_Y(z)\bar{F}_T(z) + C(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z)) - (1 - \bar{F}_Y(z))(1 - \bar{F}_T(z)) \\ &= C(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z)) - 1 + \bar{F}_Y(z) + \bar{F}_T(z) \end{aligned}$$

this entails

$$\bar{F}_Z(z) = \begin{cases} z^{-1/\gamma_T}(1 + o(1)) & \text{if } \gamma_T < 1, \\ (2 - 2^{1/\theta})z^{-1}(1 + o(1)) & \text{if } \gamma_T = 1, \\ z^{-1}(1 + o(1)) & \text{if } \gamma_T > 1. \end{cases}$$

It follows from this computation that  $\bar{F}_Z$  is indeed regularly varying with index  $\min(1, \gamma_T)$ .

Let us now highlight a couple of consequences of Theorem 1 about the tail behaviour of the observed variable  $Z$  in our setup. For ease of exposition, we assume until the end of this section that  $\gamma_Y \gamma_T > 0$ , *i.e.*  $Y$  and  $T$  both belong to the same max-domain of attraction, that can be either the Fréchet or

Weibull domain of attraction. It has been shown (see Einmahl *et al.*, 2008) that in the independent censoring case,  $Z$  then belongs to the common max-domain of attraction of  $Y$  and  $T$ , with extreme value index

$$\gamma_Z = \frac{\gamma_Y \gamma_T}{\gamma_Y + \gamma_T}; \text{ in particular, } |\gamma_Z| < \min(|\gamma_Y|, |\gamma_T|).$$

In the independent case, the absolute value of the extreme value index of the observed variable  $Z$  is therefore strictly lower than what it is in the dependent case. This means that in the dependent case, the right tail of  $Z$  is heavier than it is in the independent case. Qualitatively, this could have been inferred from the positive tail dependence between  $Y$  and  $T$ , which is due to the dependence structure being described by an extreme value copula (see Gudendorf and Segers, 2010). What is surprising is that the expression of the dependence function itself does not affect the extreme value index of  $Z$  at all.

Theorem 1 also has an important corollary relative to the identifiability of the extremes of  $Y$ : whereas the extremes of  $Z$  always contain information about those of  $Y$  and those of  $T$  in the independent case, they are driven either solely by those of  $Y$  or those of  $T$  in the dependent case considered here, no matter how close to independence the dependence structure is. It should be especially emphasised that in the case when  $\gamma_Y \gamma_T > 0$  and  $|\gamma_T| < |\gamma_Y|$  then the extreme value index of  $Z$  is exactly that of  $T$ , which seems to indicate that the extremes of  $Y$  are unrecoverable. The only cases when this type of behaviour is observed in the independent censoring case are when either the right tail of  $T$  is much lighter than the tail of  $Y$  (*e.g.*  $T$  is light-tailed while  $Y$  is heavy-tailed) or when  $\tau_T < \tau_Y \leq \infty$ , both situations clearly being cases when the problem of recovering the extremes of  $Y$  has no solution.

Our next step is to analyse how much censoring there is in the right tail of  $Z$ , depending on the extreme behaviour of  $Y$  and  $T$ , to confirm whether or not our tentative conclusions are true.

### 3.2 Tail censoring probability

The second part of our study focuses on the information available in the censoring indicator  $\delta$ . Since we are interested in how censoring affects the right tail of  $Y$ , we should look at how  $\delta$  behaves given that the observed variable  $Z$  is large, or equivalently evaluate the probability

$$P(z) := \mathbb{P}(\delta = 1 | Z > z) = \mathbb{P}(Y \leq T | Z > z).$$

When  $z$  is close to  $\tau$ , the probability  $1 - P(z)$  gives an idea of the probability of censoring in the extremes of  $Z$ . In particular, if  $P(z)$  converges to a limit  $P(\tau)$  as  $z \uparrow \tau$ , the probability  $1 - P(\tau)$  shall be called the *tail censoring probability*.

It should be noted that Einmahl *et al.* (2008) achieve the study of censoring in the right tail by slightly

different means, as they assume that  $Y$  and  $T$  have continuous distributions and they consider

$$p(z) := \mathbb{P}(Y \leq T | Z = z) := \frac{\frac{d}{dz} [\mathbb{P}(Y \leq T, Z > z)]}{\frac{d}{dz} [\mathbb{P}(Z > z)]} = \frac{\frac{d}{dz} [\mathbb{P}(Y \leq T, Z > z)]}{\frac{d}{dz} [\bar{F}_Z(z)]}.$$

They mention (without proof) that under independent censoring and suitable extreme value conditions, this function has a limit as  $z \uparrow \tau$ ; it is actually straightforward to show that this is the limit of  $P$  as well. Their statistical arguments, however, use the quantity  $P(z)$  instead: in particular, they develop an estimator of the extreme value index of  $Y$  using the sample counterpart of  $P(z)$  for large  $z$ . We nonetheless shall consider both the functions  $P(z)$  and  $p(z)$  here and show that, as in the independent case, they both give rise to a common notion of censoring probability in the tail, but their common limit is different from what it would be in the independent case. This quantity plays a pivotal role in our context, just as the classical censoring probability  $\mathbb{P}(Y > T)$  does for the estimation of central quantities. A positive (and less than 1) censoring probability means that the problem has a solution and that traditional estimators have to be corrected in some way in order to retain consistency. A censoring probability equal to 0 means that standard, uncorrected techniques will still be consistent: we will call this case a *totally uncensored* situation. Finally, a censoring probability equal to 1 gives rise to a *totally censored* case, which is a situation when the estimation problem does not have a solution. The condition that the censoring probability belongs to the interval  $(0, 1)$  is thus crucial for the problem to be both nontrivial and workable.

Our first step to prove that the limiting tail censoring probability exists is to obtain a convenient representation of the functions  $p$  and  $P$ . In all what follows,  $\partial_1 C$  (resp.  $\partial_2 C$ ) denotes the partial derivative of  $C$  with respect to its first (resp. second) argument: it is easy to show that when  $A$  is continuously differentiable,

$$\begin{aligned} \partial_1 C(u, v) &= \frac{C(u, v)}{u} \left[ A \left( \frac{\log v}{\log u + \log v} \right) - \frac{\log v}{\log u + \log v} A' \left( \frac{\log v}{\log u + \log v} \right) \right] \\ \text{and } \partial_2 C(u, v) &= \frac{C(u, v)}{v} \left[ A \left( \frac{\log v}{\log u + \log v} \right) + \frac{\log u}{\log u + \log v} A' \left( \frac{\log v}{\log u + \log v} \right) \right]. \end{aligned}$$

The next result uses these partial derivatives to provide a convenient expression of  $p$  and  $P$ .

**Proposition 3.** *Assume that condition  $(\mathcal{M})$  holds, that  $A$  is twice continuously differentiable on  $[0, 1]$ , and that  $Y$  and  $T$  have continuous distributions with respective probability density functions  $f_Y$  and  $f_T$ . Then we have that:*

$$\begin{aligned} P(z) &= \frac{1}{\bar{F}_Z(z)} \int_z^\tau [1 - \partial_1 C(F_Y(t), F_T(t))] f_Y(t) dt \\ \text{and } p(z) &= \frac{[1 - \partial_1 C(F_Y(z), F_T(z))] f_Y(z)}{[1 - \partial_1 C(F_Y(z), F_T(z))] f_Y(z) + [1 - \partial_2 C(F_Y(z), F_T(z))] f_T(z)}. \end{aligned}$$

Interestingly, while  $P(z) \in [0, 1]$  by definition, it was far less obvious from its definition (and despite its convenient notation) that  $p(z) \in [0, 1]$ . It turns out that, in view of Proposition 3 and Lemma 1(iii), this is indeed the case. Therefore, computing the limit of  $p(z)$  as  $z \uparrow \tau$  is, as in the independent censoring case, indeed worthwhile pursuing in order to obtain a measure of censorship given that the observed variable is extreme.

Proposition 3 above is the key to our analysis of the asymptotic behaviour of  $p$  and  $P$ . We may now state our second main result, on the existence and value of the limit of the function  $P$ .

**Theorem 2.** *Assume that condition  $(\mathcal{M})$  holds, that  $A$  is twice continuously differentiable on  $[0, 1]$ , and suppose that  $Y$  and  $T$  have continuous distributions that satisfy conditions  $\mathcal{C}_1(\gamma_Y)$  and  $\mathcal{C}_1(\gamma_T)$  respectively. Then, as  $z \uparrow \tau$ , we have that:*

- (i)  $P(z) \rightarrow 0$  if  $\gamma_Y \gamma_T \geq 0$  and  $|\gamma_Y| > |\gamma_T|$ ;
- (ii)  $P(z) \rightarrow 1$  if  $\gamma_Y \gamma_T \geq 0$  and  $|\gamma_T| > |\gamma_Y|$ ;
- (iii)  $P(z) \rightarrow P_c$  if  $\gamma_Y \gamma_T \geq 0$ ,  $|\gamma_T| = |\gamma_Y|$  and  $\bar{F}_T(z)/\bar{F}_Y(z) \rightarrow c \in (0, \infty)$  as  $z \uparrow \tau$ , where:

$$\forall c > 0, P_c = \frac{1 - \frac{c}{c+1}}{1 - A\left(\frac{c}{c+1}\right)} \left\{ 1 - A\left(\frac{c}{c+1}\right) + \frac{c}{c+1} A'\left(\frac{c}{c+1}\right) \right\} \in [0, 1].$$

If moreover  $A$  is strictly convex, then  $P_c \in (0, 1)$ .

Before commenting upon this result, we show that the function  $p$  shares the asymptotic behaviour of the function  $P$ . Because  $p$  is computed using the probability density functions  $f_Y$  and  $f_T$  of  $Y$  and  $T$ , we should think of extending our extreme value condition in order to be able to control  $f_Y(z)$  and  $f_T(z)$  for large  $z$ . To keep things as simple as possible, we introduce the following refined extreme value condition.

**Condition  $\mathcal{C}_2(\gamma)$ :** The function  $F$  is continuously differentiable in a neighbourhood of its right endpoint  $\tau^*$ . Moreover, one of the following two conditions holds:

- $\tau^* = \infty$  and there is  $\gamma > 0$  such that

$$\lim_{x \uparrow \infty} \frac{x F'(x)}{1 - F(x)} = \frac{1}{\gamma};$$

- $\tau^* < \infty$  and there is  $\gamma < 0$  such that

$$\lim_{x \uparrow \tau^*} \frac{(\tau^* - x) F'(x)}{1 - F(x)} = -\frac{1}{\gamma}.$$

It is a consequence of Theorems 1.1.11 and 1.1.13 in de Haan and Ferreira (2006) that for  $\gamma \neq 0$ , condition  $\mathcal{C}_2(\gamma)$  implies condition  $\mathcal{C}_1(\gamma)$ . The former condition excludes the Gumbel max-domain of attraction, for which  $\gamma = 0$ ; it should be noted that the study of the behaviour of the function  $p$  when the distributions of  $Y$  and  $T$  belong to this domain is much harder, and this is also emphasised in Einmahl *et al.* (2008) who, in this case and assuming independent censoring, simply suppose that the limit of  $p$  exists. Outside of the Gumbel domain of attraction, condition  $\mathcal{C}_2(\gamma)$  is a fairly mild regularity hypothesis, which is true for all usual examples of distributions used in statistical aspects of extreme value theory, see p.59 and p.68 in Beirlant *et al.* (2004) and further examples in de Haan and Ferreira (2006).

When  $Y$  and  $T$  satisfy this refined extreme value condition, we can examine the asymptotic behaviour of the function  $p$ . This is the focus of our third main result.

**Theorem 3.** *Assume that condition  $(\mathcal{M})$  holds, that  $A$  is twice continuously differentiable on  $[0, 1]$ , and suppose that  $Y$  and  $T$  have continuous distributions that satisfy conditions  $\mathcal{C}_2(\gamma_Y)$  and  $\mathcal{C}_2(\gamma_T)$  respectively, with  $\gamma_Y\gamma_T > 0$ . Then, as  $z \uparrow \tau$ , we have that:*

(i)  $p(z) \rightarrow 0$  if  $|\gamma_Y| > |\gamma_T|$ ;

(ii)  $p(z) \rightarrow 1$  if  $|\gamma_T| > |\gamma_Y|$ ;

(iii)  $p(z) \rightarrow P_c$  if  $|\gamma_T| = |\gamma_Y|$  and  $\bar{F}_T(z)/\bar{F}_Y(z) \rightarrow c \in (0, \infty)$  as  $z \uparrow \tau$ , with the notation of Theorem 2.

As a conclusion, the two functions  $p$  and  $P$  yield, in the limit  $z \uparrow \tau$ , the same limit, as observed in the independent censoring case. The limit itself, however, is very different compared to what it is in the latter case:

- When  $T$  has a lighter tail than  $Y$ , then the tail censoring probability is 1, and the extremes of  $Y$  are totally censored;
- When  $Y$  has a lighter tail than  $T$ , then the tail censoring probability is 0, and the extremes of  $Y$  are totally uncensored;
- The only nontrivial case arises when  $A$  is strictly convex and  $Y$  and  $T$  have proportional tails, the tail censoring probability then being between 0 and 1 depending on the function  $A$ .

By comparison, in the independent case then, provided  $Y$  and  $T$  have continuous distributions that satisfy conditions  $\mathcal{C}_2(\gamma_Y)$  and  $\mathcal{C}_2(\gamma_T)$  respectively, with  $\gamma_Y\gamma_T > 0$  and equal right endpoints, the tail censoring probability is

$$1 - P(\tau) = \frac{\gamma_Y}{\gamma_Y + \gamma_T} \in (0, 1),$$

see Einmahl *et al.* (2008). In other words, in the independent censoring case and when the right tails of  $Y$  and  $T$  are of the same type, the problem of recovering the extremes of  $Y$  always has a solution. We illustrate these remarks further in the example below.

**Example 2.** Suppose that a random variable  $Y$  has a unit Pareto distribution, so that  $F_Y(y) = 1 - y^{-1}$  for  $y > 1$ . Suppose also that the censoring variable  $T$  has a Pareto distribution with tail index  $\gamma_T$ , that is,  $F_T(t) = 1 - t^{-1/\gamma_T}$  for  $t > 1$ . In the independent case, the tail censoring probability would be

$$1 - P(\tau) = \frac{1}{1 + \gamma_T} \in (0, 1),$$

decreasing smoothly from 1 to 0 as the tail of  $T$  gets heavier, as we would expect. By contrast, in the dependent case with Pickands dependence function  $A(t) = 1 - t(1 - t)$ , the tail censoring probability is

$$1 - P(\tau) = \begin{cases} 1 & \text{if } \gamma_T < 1, \\ 1/2 & \text{if } \gamma_T = 1, \\ 0 & \text{if } \gamma_T > 1, \end{cases}$$

which, although still a decreasing function of  $\gamma_T$ , is discontinuous at 1 and piecewise constant. In particular, any case in which  $\gamma_T < 1$  is totally censored and recovering the extremes of  $Y$  is then impossible.

Due to the nature of the limit  $P(\tau)$  in the dependent case, this example is very much representative of the different behaviours that can occur, the only degrees of liberty here being the position of the discontinuity (*i.e.* the value of  $\gamma_Y$ , taken to be equal to 1 in the above example) and the value of  $P(\tau)$  at this discontinuity (which depends on the function  $A$ ). Again, it should be emphasised that the Pickands dependence function plays a largely insignificant role in the results, in that it does not affect the value of  $P(\tau)$  except at the discontinuity. In particular, the coefficient of upper tail dependence:

$$\lambda := \lim_{t \downarrow 0} \frac{1}{t} (2t - 1 + C(1 - t, 1 - t)) = 2(1 - A(1/2))$$

(see Gudendorf and Segers, 2010), only appears when  $Y$  and  $T$  have exactly equivalent tails, and in that case

$$1 - P(\tau) = 1 - \frac{1}{\lambda} \left\{ \frac{\lambda}{2} + \frac{1}{2} A' \left( \frac{1}{2} \right) \right\} = \frac{1}{2} - \frac{A'(1/2)}{2\lambda}.$$

This actually makes sense: recall that the upper tail dependence coefficient is

$$\lambda = \lim_{u \uparrow 1} \mathbb{P}(U > u | V > u)$$

where  $(U, V)$  is a random pair with distribution function  $C$  (see again Gudendorf and Segers, 2010). In other words, the upper tail dependence coefficient measures extremal dependence in the direction

of the 45-degree line. Now  $(U, V) := (F_Y(Y), F_T(T))$  is such a random pair and

$$\begin{aligned}\bar{F}_Z(z) &= \mathbb{P}(Y > z, T > z) = \mathbb{P}(U > F_Y(z), V > F_T(z)) \\ \text{and } P(z) &= \mathbb{P}(Y \leq T | Y > z, T > z) = \mathbb{P}(Y \leq T | U > F_Y(z), V > F_T(z)).\end{aligned}$$

We should therefore only expect the upper tail dependence coefficient to appear in this problem if  $F_Y(z)$  and  $F_T(z)$  are equivalent for large  $z$ . More generally though, the fact that no parameters quantifying the amount of dependence in the extremes of  $(Y, T)$  have an influence in our results outside of the very specific case of similar tails can seem pretty surprising.

## 4 Conclusion and statistical consequences

In this paper, we analysed the extreme properties of the observed pair in a randomly right-censored model when there is a dependent censoring mechanism induced by an extreme value copula. The range of situations where the extremes of the response variable can be recovered was found to be significantly more limited than in the independent censoring case. It is also interesting to note that the actual dependence structure in the extreme value copula model only plays a very marginal role in the determination of the extremes of the observed variable.

Our results imply that a dependent censoring mechanism entails statistical consequences, that can be as serious in the study of extremes of the variable of interest as they are in the study of its central parameters. To illustrate this point further, we focus on the estimators that have been developed up to now in the literature. For ease of exposition, we work here under the extreme value conditions of Theorem 2. The estimators of Beirlant *et al.* (2007), Einmahl *et al.* (2008), Gomes and Neves (2011), Ndao *et al.* (2014), Brahimy *et al.* (2015), Beirlant *et al.* (2016) and Stupfler (2016) are all based on the fact that in the independent censoring case,

$$\frac{\gamma_Z}{P(\tau)} = \frac{\gamma_Y \gamma_T}{\gamma_Y + \gamma_T} \times \left\{ \frac{\gamma_T}{\gamma_Y + \gamma_T} \right\}^{-1} = \gamma_Y.$$

The common idea these authors use is then to plug a consistent estimator of  $\gamma_Z$  and a consistent estimator of  $P(\tau)$  in the left-hand side above in order to obtain a consistent estimator of  $\gamma_Y$ . In particular, an estimator designed this way shall be consistent whenever  $\gamma_Y, \gamma_T > 0$ , or whenever  $\gamma_Y, \gamma_T < 0$  with the right endpoints  $\tau_Y$  and  $\tau_T$  being equal. In the dependent case however, the equality

$$\frac{\gamma_Z}{P(\tau)} = \gamma_Y$$

is only true when the tail of  $\gamma_Y$  is strictly lighter than that of  $T$ , in which case we actually have  $\gamma_Z = \gamma_Y$ . This leads to a considerable restriction in the applicability of the aforementioned estimators.

Moreover, although  $P(\tau)$  has to be equal to 1 for these estimators to stay consistent, their finite-sample behaviour might be problematic because of the finite-sample bias in the estimation of  $P(\tau)$ . To illustrate this, suppose that the available data is made of independent pairs  $(Z_i, \delta_i)$ ,  $1 \leq i \leq n$ . Denote by  $Z_{1,n} \leq \dots \leq Z_{n,n}$  the order statistics of  $(Z_1, \dots, Z_n)$  and by  $\delta_{[n-i+1,n]}$  the  $\delta$ -indicator corresponding to  $Z_{n-i+1,n}$ . The standard estimator of  $P(\tau)$  is then

$$\hat{P}_k(\tau) := \frac{1}{k} \sum_{j=1}^k \delta_{[n-j+1,n]}$$

where  $k = k(n) \rightarrow \infty$  and  $k/n \rightarrow 0$ , see Einmahl *et al.* (2008). The issue here is that  $\hat{P}_k(\tau)$  is actually the sample counterpart of  $\mathbb{P}(Y \leq T | Z \geq q_Z(1 - k/n))$ , where  $q_Z(1 - k/n)$  is the quantile of  $Z$  exceeded with probability  $k/n$ . This is not equal to  $P(\tau)$  in general. As a consequence, if  $k$  is chosen too large, then the estimate  $\hat{P}_k(\tau)$  could suffer from some serious under-estimation bias, leading to over-estimation of the parameter  $\gamma_Y$ . While this is also a concern in the independent censoring case, it is all the more inconvenient in the present context since if we knew that we were in the dependent case discussed here, it would be sufficient to simply use an estimate of  $\gamma_Z$  as an estimate of  $\gamma_Y$ !

Finally, let us point out that the estimators of Beirlant *et al.* (2010), Sayah *et al.* (2014) and Worms and Worms (2014) do not directly use the ratio of  $\gamma_Z$  and  $P(\tau)$  in order to design a consistent estimator. They are however based on a Kaplan-Meier estimate of  $F_Y$  in its right tail, so that they would also be seriously affected by a violation of the independent censoring assumption, as dependence can cause the Kaplan-Meier estimator to become inconsistent (see Fisher and Kanarek, 1974, Klein and Moeschberger, 1987 and Ebrahimi and Molefe, 2003). Furthermore, the construction of the methods suggested by these papers, and/or the proofs of their consistency, crucially use the independence assumption. For instance, both estimators in Worms and Worms (2014) are based on the computation of a certain expectation of a function of  $Z$ , for which the relationship  $\mathbb{P}(Z > z) = \mathbb{P}(Y > z)\mathbb{P}(T > z)$  is essential, see equations (2) and (11) therein. Consistency of these methods is therefore unclear as well in a dependent censoring framework.

## Appendix A: Proofs of the main results

**Proof of Proposition 1.** This result is well-known under one of its many equivalent forms; we include a proof for the sake of completeness. Note that

$$\begin{aligned} \bar{F}_Z(z) = \mathbb{P}(Z > z) &= \mathbb{P}(Y > z, T > z) \\ &= 1 - \mathbb{P}(Y \leq z, T \leq z) - \mathbb{P}(Y \leq z, T > z) - \mathbb{P}(Y > z, T \leq z). \end{aligned}$$



We now rewrite each term as

$$\mathbb{P}(Y \leq z, T \leq z) = C(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z))$$

$$\mathbb{P}(Y \leq z, T > z) = \mathbb{P}(Y \leq z) - \mathbb{P}(Y \leq z, T \leq z) = 1 - \bar{F}_Y(z) - C(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z))$$

$$\text{and } \mathbb{P}(Y > z, T \leq z) = \mathbb{P}(T \leq z) - \mathbb{P}(Y \leq z, T \leq z) = 1 - \bar{F}_T(z) - C(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z)).$$

Rearranging yields

$$\begin{aligned} \bar{F}_Z(z) &= \bar{F}_Y(z) + \bar{F}_T(z) + C(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z)) - 1 \\ &= \bar{F}_Y(z)\bar{F}_T(z) + C(1 - \bar{F}_Y(z), 1 - \bar{F}_T(z)) - (1 - \bar{F}_Y(z))(1 - \bar{F}_T(z)) \end{aligned}$$

which concludes the proof. ■

**Proof of Proposition 2.** Use Proposition 1 to get

$$\frac{\bar{F}_Z(z)}{\bar{F}_Y(z)} = \bar{F}_T(z) + \frac{\varphi(\bar{F}_Y(z), \bar{F}_T(z))}{\bar{F}_Y(z)}.$$

Set  $v_0 = \lim_{z \uparrow \tau_Y} \bar{F}_T(z) > 0$  and apply Lemma 2 to get

$$\lim_{z \uparrow \tau_Y} \frac{\varphi(\bar{F}_Y(z), \bar{F}_T(z))}{\bar{F}_Y(z)} = (1 - v_0)A'(1).$$

This is a positive constant, by Lemma 1. Thus

$$\lim_{z \uparrow \tau_Y} \frac{\bar{F}_Z(z)}{\bar{F}_Y(z)} = v_0 + (1 - v_0)A'(1) > 0$$

which completes the proof. ■

**Proof of Theorem 1.** Apply Lemma 4 to get:

- (i) If  $\gamma_Y > \gamma_T \geq 0$ , then  $\tau = \infty$  and  $\bar{F}_T(z)/\bar{F}_Y(z) \rightarrow 0$  as  $z \rightarrow \infty$ .
- (ii) If  $\gamma_T > \gamma_Y \geq 0$ , then  $\tau = \infty$  and  $\bar{F}_T(z)/\bar{F}_Y(z) = [\bar{F}_Y(z)/\bar{F}_T(z)]^{-1} \rightarrow \infty$  as  $z \rightarrow \infty$ .
- (iii) If  $0 \geq \gamma_Y > \gamma_T$ , then  $\tau < \infty$  and  $\bar{F}_T(z)/\bar{F}_Y(z) = [\bar{F}_Y(z)/\bar{F}_T(z)]^{-1} \rightarrow \infty$  as  $z \uparrow \tau$ .
- (iv) If  $0 \geq \gamma_T > \gamma_Y$ , then  $\tau < \infty$  and  $\bar{F}_T(z)/\bar{F}_Y(z) \rightarrow 0$  as  $z \uparrow \tau$ .

In each of these four cases, the result is then a direct corollary of Lemma 6. ■

**Proof of Proposition 3.** The statement on  $P(z)$  is a direct consequence of Lemma 7 and of the equality

$$P(z) = \frac{\mathbb{P}(Y \leq T, Z > z)}{\bar{F}_Z(z)}.$$

To obtain the corresponding statement on  $p(z)$ , just note that by Proposition 1,

$$\bar{F}_Z(z) = \bar{F}_Y(z) + \bar{F}_T(z) - [1 - C(F_Y(z), F_T(z))]$$

after rearranging terms. The result then follows directly by differentiating numerator and denominator in  $P(z)$ . ■

**Proof of Theorem 2.** By Proposition 3, we can write

$$P(z) = \frac{\int_z^\tau [1 - \partial_1 C(F_Y(t), F_T(t))] f_Y(t) dt}{\bar{F}_Z(z)}.$$

To prove (i), remark that if  $\gamma_Y \gamma_T \geq 0$  and  $|\gamma_Y| > |\gamma_T|$ , then  $\bar{F}_T(z)/\bar{F}_Y(z) \rightarrow 0$  as  $z \uparrow \tau$  by Lemma 4. In that case, Lemma 6(i) entails together with Lemma 8(i) that

$$P(z) = \frac{\int_z^\tau \left[ (1 + A'(0) + o(1)) \bar{F}_T(t) + \frac{1}{2} (A''(0) + o(1)) (\bar{F}_T(t)/\bar{F}_Y(t))^2 \right] f_Y(t) dt}{-A'(0) \bar{F}_T(z) (1 + o(1))}.$$

The first part of the integral in the numerator is controlled by noting that

$$0 \leq \frac{\int_z^\tau \bar{F}_T(t) f_Y(t) dt}{\bar{F}_T(z)} \leq \int_z^\tau f_Y(t) dt = \bar{F}_Y(z) \rightarrow 0 \text{ as } z \rightarrow \tau.$$

To control the second one, we first write

$$0 \leq \frac{1}{\bar{F}_T(z)} \int_z^\tau \left( \frac{\bar{F}_T(t)}{\bar{F}_Y(t)} \right)^2 f_Y(t) dt \leq \int_z^\tau \left\{ \frac{\bar{F}_T(t)}{\bar{F}_Y(t)} \right\} \left\{ \frac{f_Y(t)}{\bar{F}_Y(t)} \right\} dt.$$

When  $\gamma_Y > 0$ , and therefore  $\tau = \infty$ , we apply Lemma 4 to obtain that there is  $\varepsilon > 0$  such that

$$\begin{aligned} 0 \leq \frac{1}{\bar{F}_T(z)} \int_z^\infty \left( \frac{\bar{F}_T(t)}{\bar{F}_Y(t)} \right)^2 f_Y(t) dt &\leq \int_z^\infty \left\{ \frac{f_Y(t)}{\bar{F}_Y(t)} \right\} t^{-\varepsilon} dt \\ &= \int_z^\infty \frac{d}{dt} [-\log(\bar{F}_Y(t)) t^{-\varepsilon}] - \varepsilon \log(\bar{F}_Y(t)) t^{-1-\varepsilon} dt. \end{aligned}$$

Now, by Theorem 1.2.1 in de Haan and Ferreira (2006), the function  $\bar{F}_Y$  is regularly varying with index  $-1/\gamma_Y$ . In other words, we can write this function as

$$\bar{F}_Y(t) = t^{-1/\gamma_Y} L_Y(t)$$

where  $L_Y$  is a slowly varying function at infinity. Since  $\log(L_Y(t))/\log t \rightarrow 0$  as  $t \rightarrow \infty$  (see Proposition 1.3.6 p.16 in Bingham *et al.*, 1987) we get  $\log(\bar{F}_Y(t)) t^{-\varepsilon/2} \rightarrow 0$  as  $t \rightarrow \infty$ . In particular:

$$|\log(\bar{F}_Y(t)) t^{-\varepsilon}| \rightarrow 0 \text{ and } |\log(\bar{F}_Y(t)) t^{-1-\varepsilon}| = o(t^{-1-\varepsilon/2}).$$

Therefore

$$0 \leq \frac{1}{\overline{F}_T(z)} \int_z^\infty \left( \frac{\overline{F}_T(t)}{\overline{F}_Y(t)} \right)^2 f_Y(t) dt \leq \log(\overline{F}_Y(z)) z^{-\varepsilon} - \varepsilon \int_z^\infty \log(\overline{F}_Y(t)) t^{-1-\varepsilon} dt \rightarrow 0 \text{ as } z \rightarrow \infty$$

which concludes the proof in this case. When  $\gamma_Y < 0$  and  $\tau < \infty$ , we apply Lemma 4 again to obtain that there is  $\varepsilon > 0$  such that

$$0 \leq \frac{1}{\overline{F}_T(z)} \int_z^\tau \left( \frac{\overline{F}_T(t)}{\overline{F}_Y(t)} \right)^2 f_Y(t) dt \leq \int_z^\tau \left\{ \frac{f_Y(t)}{\overline{F}_Y(t)} \right\} (\tau - t)^\varepsilon dt.$$

Use the change of variables  $T = (\tau - t)^{-1}$  to obtain

$$0 \leq \frac{1}{\overline{F}_T(z)} \int_z^\tau \left( \frac{\overline{F}_T(t)}{\overline{F}_Y(t)} \right)^2 f_Y(t) dt \leq \int_{(\tau-z)^{-1}}^\infty \left\{ \frac{f_Y(\tau - T^{-1})}{\overline{F}_Y(\tau - T^{-1})} \frac{d}{dT} (\tau - T^{-1}) \right\} T^{-\varepsilon} dT.$$

Since by Theorem 1.2.1 in de Haan and Ferreira (2006) the function  $T \mapsto \overline{F}_Y(\tau - T^{-1})$  is regularly varying with index  $1/\gamma_Y < 0$ , we can exactly mimic the proof of the case  $\gamma_Y > 0$  to obtain

$$\frac{1}{\overline{F}_T(z)} \int_z^\tau \left( \frac{\overline{F}_T(t)}{\overline{F}_Y(t)} \right)^2 f_Y(t) dt \rightarrow 0 \text{ as } z \uparrow \tau.$$

The proof of (i) is then complete.

To prove (ii), we note that in this case  $\overline{F}_T(z)/\overline{F}_Y(z) = [\overline{F}_Y(z)/\overline{F}_T(z)]^{-1} \rightarrow \infty$  as  $z \uparrow \tau$  by Lemma 4. Apply then Lemma 6(ii) together with Lemma 8(ii) to get

$$P(z) = \frac{\int_z^\tau (A'(1) + o(1)) f_Y(t) dt}{A'(1) \overline{F}_Y(z) (1 + o(1))}.$$

Since  $A'(1) > 0$  by Lemma 1(ii), the numerator is clearly equivalent to  $A'(1) \overline{F}_Y(z)$ , which entails that  $P(z) \rightarrow 1$  as  $z \rightarrow \tau$  and concludes the proof of (ii).

Finally, to show (iii), apply Lemma 6(iii) together with Lemma 8(iii) to obtain

$$P(z) = \frac{\int_z^\tau \left( 1 - \left[ A \left( \frac{c}{c+1} \right) - \frac{c}{c+1} A' \left( \frac{c}{c+1} \right) \right] + o(1) \right) f_Y(t) dt}{(c+1) \left[ 1 - A \left( \frac{c}{c+1} \right) \right] \overline{F}_Y(z) (1 + o(1))}.$$

If the constant  $1 - \left[ A \left( \frac{c}{c+1} \right) - \frac{c}{c+1} A' \left( \frac{c}{c+1} \right) \right]$  is zero then

$$P(z) = \frac{o(\overline{F}_Y(z))}{(c+1) \left[ 1 - A \left( \frac{c}{c+1} \right) \right] \overline{F}_Y(z) (1 + o(1))} = o(1) \rightarrow 0 = P_c \text{ as } z \uparrow \tau$$

as required. Otherwise we clearly have

$$P(z) \rightarrow \frac{1 - \left[ A \left( \frac{c}{c+1} \right) - \frac{c}{c+1} A' \left( \frac{c}{c+1} \right) \right]}{(c+1) \left[ 1 - A \left( \frac{c}{c+1} \right) \right]} = P_c \text{ as } z \uparrow \tau,$$

and this completes the proof. ■

**Proof of Theorem 3.** Recall that Proposition 3 yields

$$p(z) = \frac{[1 - \partial_1 C(F_Y(z), F_T(z))]f_Y(z)}{[1 - \partial_1 C(F_Y(z), F_T(z))]f_Y(z) + [1 - \partial_2 C(F_Y(z), F_T(z))]f_T(z)}.$$

We first prove (i). In that case,  $\bar{F}_T(z)/\bar{F}_Y(z) \rightarrow 0$  as  $z \uparrow \tau$  by Lemma 4 and thus

$$\begin{aligned} p(z) &= \frac{\left[ (1 + A'(0) + o(1))\bar{F}_T(z) + \frac{1}{2}(A''(0) + o(1))(\bar{F}_T(z)/\bar{F}_Y(z))^2 \right] f_Y(z)}{\left[ (1 + A'(0) + o(1))\bar{F}_T(z) + \frac{1}{2}(A''(0) + o(1))(\bar{F}_T(z)/\bar{F}_Y(z))^2 \right] f_Y(z) - (A'(0) + o(1))f_T(z)} \\ &= \frac{(1 + A'(0) + o(1))\bar{F}_T(z) + \frac{1}{2}(A''(0) + o(1))(\bar{F}_T(z)/\bar{F}_Y(z))^2}{(1 + A'(0) + o(1))\bar{F}_T(z) + \frac{1}{2}(A''(0) + o(1))(\bar{F}_T(z)/\bar{F}_Y(z))^2 - (A'(0) + o(1))f_T(z)/f_Y(z)} \end{aligned}$$

by Lemma 8(i) and Lemma 9(i). Moreover, it follows from conditions  $\mathcal{C}_2(\gamma_Y)$  and  $\mathcal{C}_2(\gamma_T)$  that  $f_T(z)/f_Y(z)$  is equivalent to  $\gamma_Y \bar{F}_T(z)/[\gamma_T \bar{F}_Y(z)]$  when  $z \uparrow \tau$ . Therefore

$$\begin{aligned} p(z) &= \frac{(1 + A'(0) + o(1))\bar{F}_T(z) + \frac{1}{2}(A''(0) + o(1))(\bar{F}_T(z)/\bar{F}_Y(z))^2}{-(A'(0) + o(1))\gamma_Y \bar{F}_T(z)/[\gamma_T \bar{F}_Y(z)](1 + o(1))} \\ &= \frac{\gamma_T}{\gamma_Y} \frac{(1 + A'(0) + o(1))\bar{F}_Y(z) + \frac{1}{2}(A''(0) + o(1))\bar{F}_T(z)/\bar{F}_Y(z)}{-(A'(0) + o(1))} (1 + o(1)) \rightarrow 0 \text{ as } z \uparrow \tau \end{aligned}$$

as required.

Let us now show (ii). In that case,  $\bar{F}_T(z)/\bar{F}_Y(z) = [\bar{F}_Y(z)/\bar{F}_T(z)]^{-1} \rightarrow \infty$  as  $z \uparrow \tau$  by Lemma 4. By Lemma 8(ii) and Lemma 9(ii), we then get

$$\begin{aligned} p(z) &= \frac{(A'(1) + o(1))f_Y(z)}{(A'(1) + o(1))f_Y(z) + \left[ (1 - A'(1) + o(1))\bar{F}_Y(z) + \frac{1}{2}(A''(1) + o(1))(\bar{F}_Y(z)/\bar{F}_T(z))^2 \right] f_T(z)} \\ &= \frac{(A'(1) + o(1))}{(A'(1) + o(1)) + \left[ (1 - A'(1) + o(1))\bar{F}_Y(z) + \frac{1}{2}(A''(1) + o(1))(\bar{F}_Y(z)/\bar{F}_T(z))^2 \right] f_T(z)/f_Y(z)}. \end{aligned}$$

Again, since  $f_T(z)/f_Y(z)$  is equivalent to  $\gamma_Y \bar{F}_T(z)/[\gamma_T \bar{F}_Y(z)]$  when  $z \uparrow \tau$ , we obtain

$$\begin{aligned} &\left[ (1 - A'(1) + o(1))\bar{F}_Y(z) + \frac{1}{2}(A''(1) + o(1))(\bar{F}_Y(z)/\bar{F}_T(z))^2 \right] f_T(z)/f_Y(z) \\ &= O(\bar{F}_T(z)) + O(\bar{F}_Y(z)/\bar{F}_T(z)) \rightarrow 0 \end{aligned}$$

and therefore  $p(z) \rightarrow 1$  as  $z \uparrow \tau$ , which is (ii).

We finally show (iii). In that case, we know that by conditions  $\mathcal{C}_2(\gamma_Y)$  and  $\mathcal{C}_2(\gamma_T)$ ,  $f_T(z)/f_Y(z)$  is equivalent to  $\bar{F}_T(z)/\bar{F}_Y(z)$  when  $z \uparrow \tau$ , and therefore converges to the constant  $c$ . Apply then

Lemma 8(iii) and Lemma 9(iii) to get, as  $z \uparrow \tau$ :

$$\begin{aligned}
p(z) &\rightarrow \frac{1 - \left[ A\left(\frac{c}{c+1}\right) - \frac{c}{c+1} A'\left(\frac{c}{c+1}\right) \right]}{1 - \left[ A\left(\frac{c}{c+1}\right) - \frac{c}{c+1} A'\left(\frac{c}{c+1}\right) \right] + c \left( 1 - \left[ A\left(\frac{c}{c+1}\right) + \frac{1}{c+1} A'\left(\frac{c}{c+1}\right) \right] \right)} \\
&= \frac{1 - \left[ A\left(\frac{c}{c+1}\right) - \frac{c}{c+1} A'\left(\frac{c}{c+1}\right) \right]}{(c+1) \left[ 1 - A\left(\frac{c}{c+1}\right) \right]} = P_c.
\end{aligned}$$

The proof is complete. ■

## Appendix B: Auxiliary results and proofs

In this Appendix,  $A$  denotes a Pickands dependence function,  $C$  is the corresponding extreme value copula

$$C(u, v) = (uv)^{A(\log(v)/\log(uv))}$$

and  $\varphi$  is the function defined by:

$$\varphi(u, v) = C(1-u, 1-v) - (1-u)(1-v).$$

The first lemma gathers a few results about Pickands dependence functions.

**Lemma 1.** *Let  $A$  be a Pickands dependence function, and assume that the related copula  $C$  is not the independent copula. Then:*

(i) *It holds that  $A(t) < 1$  for all  $t \in (0, 1)$ .*

*If moreover  $A$  is continuously differentiable on  $[0, 1]$ , then:*

(ii) *We have  $A'(0) < 0$  and  $A'(1) > 0$ .*

(iii) *We have*

$$\forall x \in [0, 1], \quad 0 \leq A(x) - xA'(x) \leq 1 \quad \text{and} \quad 0 \leq A(x) + (1-x)A'(x) \leq 1.$$

*These inequalities are all strict on  $(0, 1)$  should  $A$  be strictly convex.*

**Proof of Lemma 1.** To show (i), assume that there is  $t_0 \in (0, 1)$  such that  $A(t_0) = 1$ . Recall the increasing slopes inequality:

$$\forall x, y, z \in [0, 1], \quad x < y < z \Rightarrow \frac{A(y) - A(x)}{y - x} \leq \frac{A(z) - A(x)}{z - x} \leq \frac{A(z) - A(y)}{z - y}.$$

Apply this first with  $x = t \in (0, t_0)$ ,  $y = t_0$  and  $z = 1$  to get

$$\frac{1 - A(t)}{1 - t} \leq 0$$

so that  $A(t) \geq 1$  and therefore  $A(t) = 1$  for all  $t \in (0, t_0)$ . Apply now the above set of inequalities with  $x = 0$ ,  $y = t_0$  and  $z = t \in (t_0, 1)$  to obtain

$$0 \leq \frac{A(t) - 1}{t}$$

so that again  $A(t) = 1$  for all  $t \in (t_0, 1)$ . Consequently,  $A$  is the constant function 1, which is impossible since  $C$  is not the independent copula.

We turn to the proof of (ii). Since  $A$  is continuously differentiable on  $[0, 1]$  and convex, its derivative  $A'$  is nondecreasing. Consequently, if  $A'(0)$  were nonnegative, then so would be  $A'(x)$  for any  $x \in (0, 1]$ , and thus  $A$  would be nondecreasing on  $[0, 1]$ . Since  $A(0) = A(1) = 1$ , this would entail that  $A$  is the constant function 1, which is a contradiction. Similarly,  $A'(1)$  cannot be nonpositive since if it were, then  $A$  would be nonincreasing on  $[0, 1]$ , which is a contradiction in virtue of the equality  $A(0) = A(1) = 1$ .

We now show (iii). Notice that since  $A(x) \geq x$ , we have

$$A(x) - xA'(x) \geq x(1 - A'(x)).$$

Assume that there is  $x_0 \in (0, 1)$  such that  $A'(x_0) > 1$ . Since  $A$  is convex, the function  $A'$  is nondecreasing and therefore  $A'(x) > 1$  on  $[x_0, 1]$ . Integrating yields

$$1 - A(x_0) = A(1) - A(x_0) = \int_{x_0}^1 A'(x) dx > 1 - x_0$$

which entails  $A(x_0) < x_0$  and a contradiction. As a consequence,  $A'(x) \leq 1$  for all  $x \in (0, 1)$ , and this should also be true on the closed interval  $[0, 1]$  by continuity of  $A'$ . Finally,  $A(x) - xA'(x) \geq x(1 - A'(x)) \geq 0$  for all  $x \in [0, 1]$ . Use now the increasing slopes inequality to write

$$\forall x, t \in (0, 1), x < t \Rightarrow \frac{A(x) - 1}{x} = \frac{A(x) - A(0)}{x - 0} \leq \frac{A(t) - A(x)}{t - x}.$$

Let  $t \downarrow x$  to get

$$\forall x \in (0, 1), \frac{A(x) - 1}{x} \leq A'(x)$$

which entails  $A(x) - xA'(x) \leq 1$  for all  $x \in (0, 1)$ . Using the continuity of  $A$  and  $A'$ , this inequality is of course also true for  $x \in \{0, 1\}$ , which concludes the proof of the first set of desired inequalities if  $A$  is convex.

If  $A$  is moreover strictly convex, assume that there is  $x_0 \in (0, 1)$  such that  $A'(x_0) \geq 1$ . Then  $A'(x) \geq 1$  on  $[x_0, 1]$  and integrating yields

$$\forall t \in [x_0, 1] \cap [1/2, 1], \quad 1 - A(t) = A(1) - A(t) = \int_t^1 A'(x) dx \geq 1 - t$$

which entails  $A(t) \leq t$  and therefore  $A(t) = t$  on an interval with nonempty interior, which is a contradiction by the strict convexity of  $A$ . This entails  $A(x) - xA'(x) \geq x(1 - A'(x)) > 0$  for all  $x \in (0, 1)$ . Finally, to prove that  $A(x) - xA'(x) < 1$  for all  $x \in (0, 1)$ , we apply again the increasing slopes inequality: for all  $x, t \in (0, 1)$ ,

$$x < t \Rightarrow \frac{A(x) - 1}{x} = \frac{A(x) - A(0)}{x - 0} < \frac{A(x) - A(x/2)}{x/2} < \frac{A(t) - A(x/2)}{t - (x/2)} < \frac{A(t) - A(x)}{t - x}$$

where all inequality signs are strict because of the strict convexity of  $A$ . Let  $t \downarrow x$  to get

$$\forall x \in (0, 1), \quad \frac{A(x) - 1}{x} < \frac{A(x) - A(x/2)}{x/2} \leq A'(x)$$

which entails  $A(x) - xA'(x) < 1$  for all  $x \in (0, 1)$  as required.

The second set of inequalities is a straightforward consequence of the first one applied to the Pickands dependence function  $\tilde{A}(x) = A(1 - x)$ . We omit the details for the sake of brevity.  $\blacksquare$

The second lemma is an equivalent of  $\varphi(u, v)$  as  $u \rightarrow 0$  and  $v \rightarrow v_0 > 0$ .

**Lemma 2.** *Assume that  $A$  is continuously differentiable on  $[0, 1]$ . Then, if  $u \rightarrow 0$  and  $v \rightarrow v_0 > 0$ , we have that:*

$$\varphi(u, v) = (1 - v_0)A'(1)u(1 + o(1)).$$

**Proof of Lemma 2.** Note that

$$\begin{aligned} \varphi(u, v) &= (1 - u)(1 - v) \left\{ [(1 - u)(1 - v)]^{A(\log(1-v)/\log([1-u][1-v]))-1} - 1 \right\} \\ &= (1 - v_0 + o(1)) \left\{ [(1 - u)(1 - v)]^{A(\log(1-v)/\log([1-u][1-v]))-1} - 1 \right\}. \end{aligned}$$

Since  $u \rightarrow 0$  and  $v \rightarrow v_0 > 0$ , we get

$$\frac{\log(1 - v)}{\log([1 - u][1 - v])} \rightarrow 1 \quad \text{and therefore} \quad A\left(\frac{\log(1 - v)}{\log([1 - u][1 - v])}\right) \rightarrow A(1) = 1.$$

In particular:

$$\begin{aligned} A\left(\frac{\log(1 - v)}{\log([1 - u][1 - v])}\right) - 1 &= \left[ \frac{\log(1 - v)}{\log([1 - u][1 - v])} - 1 \right] A'(1) + o\left(\left| \frac{\log(1 - v)}{\log([1 - u][1 - v])} - 1 \right|\right) \\ &= \left[ \frac{-\log(1 - u)}{\log([1 - u][1 - v])} \right] A'(1) + o\left(\frac{\log(1 - u)}{\log([1 - u][1 - v])}\right). \end{aligned}$$

Since as  $u \rightarrow 0$  and  $v \rightarrow v_0 > 0$  we have:

$$\frac{1}{u} \times \frac{-\log(1-u)}{\log([1-u][1-v])} \rightarrow \frac{1}{\log(1-v_0)} < 0,$$

we get

$$\log([1-u][1-v]) \left[ A \left( \frac{\log(1-v)}{\log([1-u][1-v])} \right) - 1 \right] = uA'(1) + o(u).$$

Because this quantity converges to 0 and

$$\varphi(u, v) = (1 - v_0 + o(1)) \left\{ \exp \left( \log([1-u][1-v]) \left[ A \left( \frac{\log(1-v)}{\log([1-u][1-v])} \right) - 1 \right] \right) - 1 \right\}$$

the conclusion readily follows by using a Taylor expansion of the exponential function in a neighbourhood of 0. ■

The next lemma is a simple result of real analysis we shall use in the proof of our next lemma.

**Lemma 3.** *Let  $\eta$  be a continuously differentiable function on a neighbourhood of infinity such that  $\eta'(t) \rightarrow 0$  as  $t \rightarrow \infty$ . Then  $\eta(t)/t \rightarrow 0$  as  $t \rightarrow \infty$ .*

**Proof of Lemma 3.** Because  $\eta'$  converges to 0 at infinity one may, for any  $\varepsilon > 0$ , choose  $t_0$  such that  $|\eta'(s)| \leq \varepsilon/2$  for all  $s > t_0$ . Therefore, for  $t > t_0$  sufficiently large,

$$\left| \frac{\eta(t)}{t} \right| \leq \frac{1}{t} \left( |\eta(t_0)| + \int_{t_0}^t |\eta'(s)| ds \right) \leq \frac{|\eta(t_0)|}{t} + \sup_{s>t_0} |\eta'(s)| \leq \varepsilon,$$

or equivalently,  $\eta(t)/t \rightarrow 0$  as  $t \rightarrow \infty$  as required. ■

Lemma 4 is a very useful result about the asymptotic interactions between the survival functions of  $Y$  and  $T$ .

**Lemma 4.** *Assume that condition  $(\mathcal{M})$  holds and that the distributions of  $Y$  and  $T$  satisfy conditions  $\mathcal{C}_1(\gamma_Y)$  and  $\mathcal{C}_1(\gamma_T)$  respectively. If in addition  $\gamma_Y \gamma_T \geq 0$  and  $|\gamma_Y| > |\gamma_T|$ , then when  $\tau = \infty$  we have*

$$\exists \varepsilon > 0, \frac{\bar{F}_T(z)}{\bar{F}_Y(z)} = O(z^{-\varepsilon}) \quad \text{as } z \rightarrow \infty.$$

When otherwise  $\tau < \infty$ , then

$$\exists \varepsilon > 0, \frac{\bar{F}_T(z)}{\bar{F}_Y(z)} = O((\tau - z)^\varepsilon) \quad \text{as } z \uparrow \tau.$$

**Proof of Lemma 4.** Although this result appears to be mostly well-known, we prove it for the sake of self-containedness. We start with the case  $\gamma_Y > 0$ , in which  $\tau = \infty$ . We may write for large  $z$  that:

$$\bar{F}_Y(z) = c_Y(z) \exp \left( - \int_{z_0}^z \frac{ds}{\eta_Y(s)} \right) \quad \text{and} \quad \bar{F}_T(z) = c_T(z) \exp \left( - \int_{z_0}^z \frac{ds}{\eta_T(s)} \right)$$



where  $c_Y$  and  $c_T$  converge to positive constants at infinity, and  $\eta_Y$  and  $\eta_T$  are two continuously differentiable functions on a neighbourhood of infinity such that  $\eta_Y(t)/t \rightarrow \gamma_Y$  and  $\eta_T(t)/t \rightarrow \gamma_T$  as  $t \rightarrow \infty$ . This is a direct consequence of Theorem 1.2.6 in de Haan and Ferreira (2006) when  $\gamma_T > 0$ , and of this same result together with Lemma 3 when  $\gamma_T = 0$ . Then

$$\frac{\bar{F}_T(z)}{\bar{F}_Y(z)} = O\left(\exp\left[-\int_{z_0}^z \left(\frac{s}{\eta_T(s)} - \frac{s}{\eta_Y(s)}\right) \frac{ds}{s}\right]\right).$$

If  $\gamma_T > 0$  then the bracketed term in the above integral converges to  $1/\gamma_T - 1/\gamma_Y$ , and therefore

$$-\int_{z_0}^z \left(\frac{s}{\eta_T(s)} - \frac{s}{\eta_Y(s)}\right) \frac{ds}{s} = -\left(\frac{1}{\gamma_T} - \frac{1}{\gamma_Y}\right) \log(z)(1 + o(1)).$$

Thus, in this case:

$$\frac{\bar{F}_T(z)}{\bar{F}_Y(z)} = O(z^{-\varepsilon}) \quad \text{with } \varepsilon = \frac{1}{2} \left(\frac{1}{\gamma_T} - \frac{1}{\gamma_Y}\right) > 0.$$

If  $\gamma_T = 0$  then  $s/\eta_T(s)$  tends to infinity, so that

$$-\frac{1}{\log z} \int_{z_0}^z \left(\frac{s}{\eta_T(s)} - \frac{s}{\eta_Y(s)}\right) \frac{ds}{s} \rightarrow -\infty$$

and therefore

$$\frac{\bar{F}_T(z)}{\bar{F}_Y(z)} = O(z^{-\varepsilon}) \quad \text{for all } \varepsilon > 0.$$

This concludes the proof in the case  $\gamma_Y > 0$ . When  $\gamma_Y < 0$  and thus  $\tau < \infty$ , the idea is consider the survival functions  $T \mapsto \bar{F}_T(\tau - T^{-1})$  and  $T \mapsto \bar{F}_Y(\tau - T^{-1})$ , which are regularly varying with indices  $1/\gamma_Y$  and  $1/\gamma_T$ , and to apply the result we have just shown to get

$$\exists \varepsilon > 0, \frac{\bar{F}_T(\tau - T^{-1})}{\bar{F}_Y(\tau - T^{-1})} = O(T^{-\varepsilon}) \quad \text{as } T \rightarrow \infty.$$

We can then substitute  $z = \tau - T^{-1}$  to obtain the desired result. ■

The next result is an equivalent of  $\varphi(u, v)$  as  $u, v \rightarrow 0$ .

**Lemma 5.** *Assume that  $A$  is continuously differentiable on  $[0, 1]$ . Then, if  $u, v \rightarrow 0$ , we have that:*

- $\varphi(u, v) = -A'(0)v(1 + o(1))$  if  $v/u \rightarrow 0$ ;
- $\varphi(u, v) = A'(1)u(1 + o(1))$  if  $v/u \rightarrow \infty$ ;
- $\varphi(u, v) = (c + 1) \left[1 - A\left(\frac{c}{c+1}\right)\right] u(1 + o(1))$  if  $v/u \rightarrow c \in (0, \infty)$ .

**Proof of Lemma 5.** Start by writing

$$\begin{aligned} \varphi(u, v) &= (1-u)(1-v) \left\{ [(1-u)(1-v)]^{A(\log(1-v)/\log([1-u][1-v]))-1} - 1 \right\} \\ &= \left\{ [(1-u)(1-v)]^{A(\log(1-v)/\log([1-u][1-v]))-1} - 1 \right\} (1 + o(1)). \end{aligned}$$

Since  $u, v \rightarrow 0$ , we get

$$A\left(\frac{\log(1-v)}{\log([1-u][1-v])}\right) = A\left(\frac{v(1+o(1))}{u(1+o(1)) + v(1+o(1))}\right) \rightarrow \begin{cases} A(0) = 1 & \text{if } v/u \rightarrow 0, \\ A(1) = 1 & \text{if } v/u \rightarrow \infty, \\ A(c/(c+1)) & \text{if } v/u \rightarrow c \in (0, \infty). \end{cases}$$

In particular:

$$A\left(\frac{\log(1-v)}{\log([1-u][1-v])}\right) - 1 = \begin{cases} \frac{v}{u}A'(0)(1+o(1)) & \text{if } v/u \rightarrow 0, \\ -\frac{u}{v}A'(1)(1+o(1)) & \text{if } v/u \rightarrow \infty, \\ A(c/(c+1)) - 1 + o(1) & \text{if } v/u \rightarrow c \in (0, \infty). \end{cases}$$

Meanwhile

$$\log([1-u][1-v]) = \begin{cases} -u(1+o(1)) & \text{if } v/u \rightarrow 0, \\ -v(1+o(1)) & \text{if } v/u \rightarrow \infty, \\ -(c+1)u(1+o(1)) & \text{if } v/u \rightarrow c \in (0, \infty) \end{cases}$$

so that

$$\begin{aligned} & \log([1-u][1-v]) \left[ A\left(\frac{\log(1-v)}{\log([1-u][1-v])}\right) - 1 \right] \\ &= \begin{cases} -vA'(0)(1+o(1)) & \text{if } v/u \rightarrow 0, \\ uA'(1)(1+o(1)) & \text{if } v/u \rightarrow \infty, \\ (c+1)[1 - A(c/(c+1))]u(1+o(1)) & \text{if } v/u \rightarrow c \in (0, \infty). \end{cases} \end{aligned}$$

Since this quantity converges to 0 and

$$\varphi(u, v) = (1+o(1)) \left\{ \exp\left(\log([1-u][1-v]) \left[ A\left(\frac{\log(1-v)}{\log([1-u][1-v])}\right) - 1 \right]\right) - 1 \right\}$$

the conclusion readily follows by using a Taylor expansion of the exponential function in a neighbourhood of 0. ■

Lemma 6, which provides an asymptotic equivalent of  $\bar{F}_Z$ , is a direct corollary of Lemma 5. Note that it gives indeed a true asymptotic equivalent of  $\bar{F}_Z$ , since  $A'(0) < 0$ ,  $A'(1) > 0$  and  $A(t) < 1$  for any  $t \in (0, 1)$ , see Lemma 1(i) and (ii).

**Lemma 6.** *Assume that condition  $(\mathcal{M})$  holds and that  $A$  is continuously differentiable on  $[0, 1]$ . Then, as  $z \uparrow \tau$ , we have that:*

$$(i) \quad \bar{F}_Z(z) = -A'(0)\bar{F}_T(z)(1+o(1)) \text{ if } \bar{F}_T(z)/\bar{F}_Y(z) \rightarrow 0 \text{ as } z \uparrow \tau;$$

(ii)  $\bar{F}_Z(z) = A'(1)\bar{F}_Y(z)(1 + o(1))$  if  $\bar{F}_T(z)/\bar{F}_Y(z) \rightarrow \infty$  as  $z \uparrow \tau$ ;

(iii)  $\bar{F}_Z(z) = (c+1) \left[ 1 - A \left( \frac{c}{c+1} \right) \right] \bar{F}_Y(z)(1 + o(1))$  if  $\bar{F}_T(z)/\bar{F}_Y(z) \rightarrow c \in (0, \infty)$ .

**Proof of Lemma 6.** Just use Proposition 1 together with Lemma 5. ■

Lemma 7 below is the key to the proof of Proposition 3.

**Lemma 7.** *Assume that condition  $(\mathcal{M})$  holds, that  $A$  is twice continuously differentiable on  $[0, 1]$ , and that  $Y$  and  $T$  have continuous distributions with respective probability density functions  $f_Y$  and  $f_T$ . Then we have that:*

$$\mathbb{P}(Y \leq T, Z > z) = \int_z^\tau [1 - \partial_1 C(F_Y(t), F_T(t))] f_Y(t) dt.$$

**Proof of Lemma 7.** The conditions ensure that  $(y, t) \mapsto C(F_Y(y), F_T(t))$  is twice continuously differentiable. Denoting by  $\partial_{1,2}^2$  the mixed partial derivative, it follows that the pair  $(Y, T)$  has the probability density function

$$(y, t) \mapsto \partial_{1,2}^2 [C(F_Y(y), F_T(t))] = f_Y(y) f_T(t) \partial_{1,2}^2 C(F_Y(y), F_T(t)).$$

Write then

$$\begin{aligned} \mathbb{P}(Y \leq T, Z > z) &= \mathbb{P}(Y \leq T, Y > z) \\ &= \iint \mathbb{I}_{\{y \leq t, y > z\}} f_Y(y) f_T(t) \partial_{1,2}^2 C(F_Y(y), F_T(t)) dy dt \\ &= \int_{t=z}^{t=\tau} \left( \int_{y=z}^{y=t} \partial_{1,2}^2 C(F_Y(y), F_T(t)) f_Y(y) dy \right) f_T(t) dt \\ &= \int_z^\tau [\partial_2 C(F_Y(t), F_T(t)) - \partial_2 C(F_Y(z), F_T(t))] f_T(t) dt. \end{aligned}$$

Note now that since  $C$  is a copula, we have  $C(u, v) \leq \min(u, v)$ . Because the functions  $x \mapsto A(x) - xA'(x)$  and  $x \mapsto A(x) + (1-x)A'(x)$  are bounded on  $[0, 1]$  (see Lemma 1), it follows that the partial derivatives of  $C$  are bounded. We may then break the above integral into two well-defined parts, to get

$$\mathbb{P}(Y \leq T, Z > z) = \int_z^\tau \partial_2 C(F_Y(t), F_T(t)) f_T(t) dt - \int_z^\tau \partial_2 C(F_Y(z), F_T(t)) f_T(t) dt.$$

The second integral on the right-hand side is

$$\int_z^\tau \partial_2 C(F_Y(z), F_T(t)) f_T(t) dt = C(F_Y(z), 1) - C(F_Y(z), F_T(z)) = F_Y(z) - C(F_Y(z), F_T(z)) \quad (1)$$

since  $C$  is a copula. Moreover,

$$\partial_2 C(F_Y(t), F_T(t)) f_T(t) = \frac{d}{dt} [C(F_Y(t), F_T(t))] - \partial_1 C(F_Y(t), F_T(t)) f_Y(t)$$

so that

$$\begin{aligned} \int_z^\tau \partial_2 C(F_Y(t), F_T(t)) f_T(t) dt &= C(1, 1) - C(F_Y(z), F_T(z)) - \int_z^\tau \partial_1 C(F_Y(t), F_T(t)) f_Y(t) dt \\ &= 1 - C(F_Y(z), F_T(z)) - \int_z^\tau \partial_1 C(F_Y(t), F_T(t)) f_Y(t) dt. \end{aligned} \quad (2)$$

Combine (1) and (2) to get

$$\mathbb{P}(Y \leq T, Z > z) = \bar{F}_Y(z) - \int_z^\tau \partial_1 C(F_Y(t), F_T(t)) f_Y(t) dt = \int_z^\tau [1 - \partial_1 C(F_Y(t), F_T(t))] f_Y(t) dt$$

which is the desired result. ■

The next lemma contains an equivalent of the quantity  $1 - \partial_1 C(1 - u, 1 - v)$  as  $u, v \rightarrow 0$ .

**Lemma 8.** *Assume that  $A$  is twice continuously differentiable on  $[0, 1]$ . Then, if  $u, v \rightarrow 0$ , we have that:*

$$(i) \quad 1 - \partial_1 C(1 - u, 1 - v) = (1 + A'(0))v + \frac{1}{2}A''(0)(v/u)^2 + o(v) + o((v/u)^2) \text{ if } v/u \rightarrow 0;$$

$$(ii) \quad 1 - \partial_1 C(1 - u, 1 - v) \rightarrow A'(1) \text{ if } v/u \rightarrow \infty;$$

$$(iii) \quad 1 - \partial_1 C(1 - u, 1 - v) \rightarrow 1 - \left[ A\left(\frac{c}{c+1}\right) - \frac{c}{c+1} A'\left(\frac{c}{c+1}\right) \right] \text{ if } v/u \rightarrow c \in (0, \infty).$$

**Proof of Lemma 8.** Recall the identity

$$\partial_1 C(1 - u, 1 - v) = \frac{C(1 - u, 1 - v)}{1 - u} B\left(\frac{\log(1 - v)}{\log(1 - u) + \log(1 - v)}\right)$$

where  $B(x) = A(x) - xA'(x)$ . We first prove (i). It is straightforward to show that

$$B(x) = 1 - \frac{x^2}{2}A''(0) + o(x^2).$$

Since in this case we assume  $v/u \rightarrow 0$ , we get

$$\frac{\log(1 - v)}{\log(1 - u) + \log(1 - v)} = \frac{v}{u}(1 + o(1)).$$

Therefore

$$\partial_1 C(1 - u, 1 - v) = \frac{C(1 - u, 1 - v)}{1 - u} \left[ 1 - \frac{A''(0)}{2} \left(\frac{v}{u}\right)^2 + o\left(\left[\frac{v}{u}\right]^2\right) \right]. \quad (3)$$

It remains to compute an asymptotic expansion of  $C(1 - u, 1 - v)/(1 - u)$ . To this end, we rewrite this term as

$$\frac{C(1 - u, 1 - v)}{1 - u} = \exp \left\{ A\left(\frac{\log(1 - v)}{\log(1 - u) + \log(1 - v)}\right) [\log(1 - u) + \log(1 - v)] - \log(1 - u) \right\}.$$

Note now that

$$\begin{aligned}\frac{\log(1-v)}{\log(1-u) + \log(1-v)} &= \frac{\log(1-v)}{\log(1-u)} \left(1 + \frac{\log(1-v)}{\log(1-u)}\right)^{-1} \\ &= \frac{\log(1-v)}{\log(1-u)} \left(1 - \frac{\log(1-v)}{\log(1-u)} + o\left(\frac{v}{u}\right)\right).\end{aligned}$$

Plugging this into the Taylor expansion

$$A(x) = A(0) + xA'(0) + \frac{x^2}{2}A''(0) + o(x^2) = 1 + xA'(0) + \frac{x^2}{2}A''(0) + o(x^2)$$

and rearranging yields

$$A\left(\frac{\log(1-v)}{\log(1-u) + \log(1-v)}\right) = 1 + \frac{\log(1-v)}{\log(1-u)}A'(0) + \left(\frac{\log(1-v)}{\log(1-u)}\right)^2\left(\frac{A''(0)}{2} - A'(0)\right) + o\left(\left[\frac{v}{u}\right]^2\right).$$

Now

$$\log(1-u) + \log(1-v) = \log(1-u) \left[1 + \frac{\log(1-v)}{\log(1-u)}\right]$$

so that

$$\begin{aligned}&A\left(\frac{\log(1-v)}{\log(1-u) + \log(1-v)}\right) [\log(1-u) + \log(1-v)] \\ &= \log(1-u) \left(1 + \frac{\log(1-v)}{\log(1-u)}(1 + A'(0)) + \left(\frac{\log(1-v)}{\log(1-u)}\right)^2 \frac{A''(0)}{2} + o\left(\left[\frac{v}{u}\right]^2\right)\right)\end{aligned}$$

and therefore

$$\begin{aligned}&A\left(\frac{\log(1-v)}{\log(1-u) + \log(1-v)}\right) [\log(1-u) + \log(1-v)] - \log(1-u) \\ &= (1 + A'(0)) \log(1-v) + \frac{[\log(1-v)]^2}{\log(1-u)} \frac{A''(0)}{2} + o\left(\frac{v^2}{u}\right).\end{aligned}$$

We combine this with a Taylor expansion of the exponential function in a neighbourhood of 0 to get

$$\begin{aligned}\frac{C(1-u, 1-v)}{1-u} &= 1 + (1 + A'(0)) \log(1-v) + \frac{[\log(1-v)]^2}{\log(1-u)} \frac{A''(0)}{2} + o\left(\frac{v^2}{u}\right) \\ &= 1 - (1 + A'(0))v + o(v).\end{aligned}\tag{4}$$

Combining (3) and (4) and rearranging terms concludes the proof of (i).

We now turn to the proof of (ii) and (iii): as  $u$  and  $v \rightarrow 0$ , we have

$$\partial_1 C(1-u, 1-v) = B\left(\frac{\log(1-v)}{\log(1-u) + \log(1-v)}\right) (1 + o(1)).$$

Besides,

$$B\left(\frac{\log(1-v)}{\log(1-u) + \log(1-v)}\right) = B\left(\frac{v(1 + o(1))}{u(1 + o(1)) + v(1 + o(1))}\right).$$

The argument of  $B$  here converges to 1 if  $v/u \rightarrow \infty$ , and to  $c/(c+1)$  if  $v/u \rightarrow c \in (0, \infty)$ . By continuity of  $B$ , it follows that  $1 - \partial_1 C(1-u, 1-v) \rightarrow 1 - B(1) = A'(1)$  if  $v/u \rightarrow \infty$ , and

$$1 - \partial_1 C(1-u, 1-v) \rightarrow 1 - B\left(\frac{c}{c+1}\right) = 1 - \left[ A\left(\frac{c}{c+1}\right) - \frac{c}{c+1} A'\left(\frac{c}{c+1}\right) \right]$$

if  $v/u \rightarrow c \in (0, \infty)$ . This completes the proof. ■

The final lemma, similar to the previous one, contains an equivalent of the quantity  $1 - \partial_2 C(1-u, 1-v)$  as  $u, v \rightarrow 0$ .

**Lemma 9.** *Assume that  $A$  is twice continuously differentiable on  $[0, 1]$ . Then, if  $u, v \rightarrow 0$ , we have that:*

(i)  $1 - \partial_2 C(1-u, 1-v) \rightarrow -A'(0)$  if  $v/u \rightarrow 0$ ;

(ii)  $1 - \partial_2 C(1-u, 1-v) = (1 - A'(1))u + \frac{1}{2}A''(1)(u/v)^2 + o(u) + o((u/v)^2)$  if  $v/u \rightarrow \infty$ ;

(iii)  $1 - \partial_2 C(1-u, 1-v) \rightarrow 1 - \left[ A\left(\frac{c}{c+1}\right) + \frac{1}{c+1}A'\left(\frac{c}{c+1}\right) \right]$  if  $v/u \rightarrow c \in (0, \infty)$ .

**Proof of Lemma 9.** Note that  $C(u, v) = \tilde{C}(v, u)$  where  $\tilde{C}$  is the copula whose Pickands dependence function is  $\tilde{A}(t) = A(1-t)$ . In particular,  $\partial_2 C(u, v) = \partial_1 \tilde{C}(v, u)$ . The result is then a direct consequence of Lemma 8. ■

## References

- Beirlant, J., Bardoutsos, A., de Wet, T., Gijbels, I. (2016). Bias reduced tail estimation for censored Pareto type distributions, *Statistics and Probability Letters* **109**: 78–88.
- Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. (2004). *Statistics of extremes: Theory and applications*, Wiley.
- Beirlant, J., Guillou, A., Dierckx, G., Fils-Villetard, A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring, *Extremes* **10**: 151–174.
- Beirlant, J., Guillou, A., Toulemonde, G. (2010). Peaks-Over-Threshold modeling under random censoring, *Communications in Statistics – Theory and Methods* **39**: 1158–1179.
- Bingham, N.H., Goldie, C.M., Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press, Cambridge, UK.

- Brahimi, B., Meraghni, D., Necir, A. (2015). Gaussian approximation to the extreme value index estimator of a heavy-tailed distribution under random censoring, *Mathematical Methods of Statistics* **24**: 266–279.
- Capéràa, P., Fougères, A.-L., Genest, C. (2000). Bivariate distributions with given extreme value attractor, *Journal of Multivariate Analysis* **72**: 30–49.
- Deheuvels, P. (1984). Probabilistic aspects of multivariate extremes. In Tiago de Oliveira, J. (ed.) *Statistical extremes and applications*: 117–130.
- Ebrahimi, N., Molefe, D. (2003). Survival function estimation when lifetime and censoring time are dependent, *Journal of Multivariate Analysis* **87**: 101–132.
- Ebrahimi, N., Molefe, D., Ying, Z. (2003). Identifiability and censored data, *Biometrika* **90**: 724–727.
- Einmahl, J.H.J., Fils-Villetard, A., Guillou, A. (2008). Statistics of extremes under random censoring, *Bernoulli* **14**: 207–227.
- Fisher, L., Kanarek, P. (2008). Presenting censored survival data when censoring and survival times may not be independent. In Proschan, F., Serfling, R. (eds.) *Reliability and Biometry: Statistical Analysis of Life Length*, *Society for Industrial and Applied Mathematics*: 303–326.
- Galambos, J. (1978). *The asymptotic theory of extreme order statistics*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York-Chichester-Brisbane.
- Gardes, L., Stupfler, G. (2015). Estimating extreme quantiles under random truncation, *TEST* **24**: 207–227.
- Gomes, M.I., Neves, M.M. (2011). Estimation of the extreme value index for randomly censored data, *Biometrical Letters* **48**: 1–22.
- Gudendorf, G., Segers, J. (2010). Extreme-value copulas. In Jaworski, P., Durante, F., Härdle, W.K., Rychlik, T. (eds.) *Copula Theory and Its Applications*, *Lecture Notes in Statistics* **198**: 127–145.
- Gumbel, E.J. (1960). Bivariate exponential distributions, *Journal of the American Statistical Association* **55**: 698–707.
- de Haan, L., Ferreira, A. (2006). *Extreme value theory: an introduction*, Springer, New York.
- Huang, X., Zhang, N. (2008). Regression survival analysis with an assumed copula for dependent censoring: a sensitivity analysis approach, *Biometrics* **64**: 1090–1099.

- Jackson, D., White, I.R., Seaman, S., Evans, H., Baisley, K., Carpenter, J. (2014). Relaxing the independent censoring assumption in the Cox proportional hazards model using multiple imputation, *Statistics in Medicine* **33**: 4681–4694.
- Joe, H. (1997). *Multivariate models and dependence concepts*, Chapman and Hall/CRC, Boca Raton, Florida.
- Kaplan, E.L., Meier, P. (1958). Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association* **53**: 457–481.
- Klein, J.P., Moeschberger, M.L. (1987). Independent or dependent competing risks: does it make a difference? *Communication in Statistics – Simulation and Computation* **16**: 507–533.
- Lagakos, W. (1979). General right censoring and its impact on the analysis of survival data, *Biometrics* **44**: 529–538.
- Ledford, A.W., Tawn, J.A. (1996). Statistics for near independence in multivariate extreme values, *Biometrika* **83**(1): 169–187.
- Ledford, A.W., Tawn, J.A. (1997). Modelling dependence within joint tail regions, *Journal of the Royal Statistical Society: Series B* **59**(2): 475–499.
- Li, Y., Tiwari, R.C., Guha, S. (2007). Mixture cure survival models with dependent censoring, *Journal of the Royal Statistical Society: Series B* **69**: 285–306.
- Modarres, M., Kaminskiy, M.P., Krivtsov, V. (2009). *Reliability engineering and risk analysis: a practical guide, Second edition*, CRC Press.
- Ndao, P., Diop, A., Dupuy, J.-F. (2014). Nonparametric estimation of the conditional tail index and extreme quantiles under random censoring, *Computational Statistics and Data Analysis* **79**: 63–79.
- Sayah, A., Yahia, D., Brahim, B. (2014). On robust tail index estimation under random censorship, *Afrika Statistika* **9**(1): 671–683.
- Sklar, A. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris* **8**: 229–231.
- Stupfler, G. (2016). Estimating the conditional extreme-value index under random right-censoring, *Journal of Multivariate Analysis* **144**: 1–24.
- Tawn, J.A. (1988). Bivariate extreme value theory: Models and estimation, *Biometrika* **75**(3): 397–



415.

Tsiatis, A. (1975). A non-identifiability aspect of the problem of competing risks, *Proceedings of the National Academy of Sciences* **72**: 20–22.

Worms, J., Worms, R. (2014). New estimators of the extreme value index under random right censoring, for heavy-tailed distributions, *Extremes* **17**(2): 337–358.

Zeng, D. (2004). Estimating marginal survival function by adjusting for dependent censoring using many covariates, *The Annals of Statistics* **32**: 1533–1555.