Approximation and estimation of very small probabilities of multivariate extreme events

Cees de Valk

Feb 16, 2015

Abstract This article discusses the multivariate generalisation of the GW (Generalised Weibull) and log-GW tail limits introduced in de Valk (2014), and its application to estimation of the probability \( p_n \) of a multivariate extreme event from a sample of \( n \) iid random vectors, with \( p_n \in [n^{-\tau_2}, n^{-\tau_1}] \) for some \( \tau_1 > 1 \) and \( \tau_2 > \tau_1 \). As a log-GW limit can be reduced to a GW limit by taking the logarithm of the random variable concerned, only the latter is considered. Its multivariate generalisation is a tail large deviation principle (LDP), which can be regarded as the analogue for very small probabilities of classical multivariate extreme value theory based on weak convergence of measures. After standardising the marginals to a distribution function \( G \) with a Weibull tail limit, dependence is represented by a homogeneous rate function \( I_G \). A connection is established between the tail LDP and residual tail dependence (RTD), originally proposed as refinement of the classical bivariate tail limit in the case of asymptotic independence. Furthermore, a new limit for probabilities of a wide class of tail events is derived which implies the recent extension of RTD in Wadsworth & Tawn (2013) as a special case. The tail LDP is extended to events which may be extreme “in any direction”, and simple estimators for very small probabilities of such events are formulated. These avoid estimation of \( I_G \) by making use of its homogeneity, and employ marginal tail estimation and “stretching” of the data cloud following a normalisation of their sample marginals. Strong consistency of the estimators is proven.

1 Introduction

In this article, we will consider estimation of very small probabilities \( p_n \) of multivariate extreme events from a sample of size \( n \), with

\[
p_n \in [n^{-\tau_2}, n^{-\tau_1}] \quad \text{with} \quad \tau_2 > \tau_1 > 1.
\]

In a univariate context, de Valk (2014) previously considered estimation of high quantiles with probabilities of exceedance \( p_n \) as in (1.1), motivated by applications requiring quantile estimates for \( p_n \ll 1/n \) in e.g. flood protection and more generally,
natural hazard assessment, and in operational risk assessment for financial institutions. Multivariate events with such low probabilities are also relevant to these fields of application. Examples are breaching of a flood protection consisting of multiple sections differing in exposure, design and maintenance along a shoreline or river bank (Vrouwenvelder & Struik (1990)), damage to an offshore structure caused by the combined effects of multiple environmental loads like water level, wave height, etc. (ISO (2005)), and operational losses suffered by banks in different business lines and due to various types of events (Embrechts & Puccetti (2007)).

Most research on estimation of probabilities of multivariate extreme events has been based on the regularity assumption that the multivariate distribution function $F$ is in the domain of attraction of some extreme value distribution function (de Haan & Ferreira (2006); Resnick (1987)), employing the exponent measure or its properties to formulate estimators; see Coles & Tawn (1994), Bruun & Tawn (1998), de Haan & Sinha (1999), Drees & de Haan (2013), and Ch. 8 of de Haan & Ferreira (2006). If some components of the random vector $X$ under consideration are asymptotically independent, these estimators may produce invalid results. To alleviate this problem, residual tail dependence was introduced as an additional regularity assumption on the tail of the multivariate survival function $F^c$ defined by $F^c(x) = P(X_i > x_i \forall i \in \{1,...,m\})$; see e.g. Ledford & Tawn (1996), Ledford & Tawn (1997), Ledford & Tawn (1998) and Draisma et al (2004). This approach was recently extended in Wadsworth & Tawn (2013).

Just as in the univariate case, the basic regularity assumptions above only allow estimation of probabilities $p_n$ vanishing slowly enough that $p_nn \to \infty$ as $n \to \infty$. Since these probabilities may be estimated without bias and consistently by the sample fraction $\hat{p}_n$ of an iid sample (as $\hat{p}_n/p_n \to 1$ in probability), estimators based on such regularity assumptions can at best reduce variance for these $p_n$. Therefore, additional assumptions are introduced in de Haan & Sinha (1999), Drees & de Haan (2013), Draisma et al (2004) and de Haan & Ferreira (2006) which allow tail extrapolation to be carried further. For the marginals, these assumptions are identical to or somewhat stronger than those for univariate quantile estimation, which are rather restrictive for probabilities (1.1) (see de Valk (2014), Proposition 1). For example, they exclude the normal and the lognormal distribution. To overcome this limitation, the GW (Generalised Weibull) and the more general log-GW tail limits were introduced in de Valk (2014) as alternative regularity assumptions to replace the familiar Generalised Pareto (GP) limit (de Haan & Ferreira (2006)). For a univariate distribution function $F$, let

$$q_F := F^{-1}(1 - e^{-1d}) = U \circ \exp,$$  \hspace{1cm} (1.2)

with $U$ the left-continuous inverse of $1/(1 - F)$. A univariate distribution function $F$ satisfies a GW (Generalised Weibull) tail limit if for some positive function $g$,

$$\lim_{y \to \infty} \frac{q_F(y\lambda) - q_F(y)}{g(y)} = \varphi(\lambda) \hspace{0.5cm} \forall \lambda > 0$$  \hspace{1cm} (1.3)

for some nonconstant function $\varphi$. Then $g$ can be chosen such that for some real $\rho$,

$$\varphi(\lambda) = h_\rho(\lambda) := \begin{cases} (\lambda^\rho - 1)/\rho & \text{if } \rho \neq 0 \\ \log \lambda & \text{if } \rho = 0 \end{cases}$$  \hspace{1cm} (1.4)

\footnote{In de Valk (2014), $q_F$ was simply written as $q$; here, we will need to keep track of the distribution function it is related to.}
A condition equivalent to (1.3) with \( q \) is expressed as \( q_F \in ERV \), with \( ERV \) the extended regularly varying functions; see de Haan & Ferreira (2006). We will write
\[
\lim_{y \to \infty} \frac{\log(1 - F(q_F(y) + yg(y)))}{y} = -h^{-1}_\rho(x) \quad \forall x \in h_\rho((0, \infty)). 
\] (1.5)

Commonly, (1.3) is expressed as \( q_F \in ERV \), with \( ERV \) the extended regularly varying functions; see de Haan & Ferreira (2006). We will write \( q_F \in ERV_S \) with \( S \subset \mathbb{R} \) to express that (1.3) holds for some \( \rho \in S \). If \( q_F \in ERV(\rho) \) for some \( \rho > 0 \), then \( \lim_{y \to \infty} q_F(y\lambda)/q_F(y) = \lambda^\rho \) for all \( \lambda > 0 \), i.e. \( q_F \in RV(\rho) \), with \( RV \) the regularly varying functions (de Haan & Ferreira (2006), Theorem B.2.2). This is equivalent to (\textit{e.g.} de Haan & Ferreira (2006), Lemma 1.1.1)
\[
\lim_{y \to \infty} \frac{\log(1 - F(xq_F(y)))}{y} = -x^{1/\rho} \quad \forall x > 0, 
\] (1.6)
i.e., \( F \) satisfies a Weibull tail limit with index \( \rho \) (Klüppelberg (1991)).

If \( F \) satisfies a GW tail limit and \( q_F(\infty) > 1 \frac{2}{3} \) then the GW limit implies the log-GW limit defined by replacing \( q_F \) in (1.3) by \( \log q_F \), i.e., \( \log q_F \in ERV \) (see de Haan & Ferreira (2006), Theorem 2(b)). The log-GW limit is a natural condition for estimation of quantiles with probabilities in the range (1.1): if \( U(\infty) > 0 \), then the GW limit implies the log-GW limit defined by replacing \( q_F \) in (1.3) by \( \log q_F \), i.e., \( \log q_F \in ERV \) (see de Haan & Ferreira (2006), Theorem 2(b)). The log-GW limit is a natural condition for estimation of quantiles with probabilities in the range (1.1): if \( U(\infty) > 0 \), then the GW limit implies the

To keep the notation simple, only GW tail limits will be considered. This does not restrict applicability in any way, since a log-GW limit can always be converted to a GW limit by replacing the random variable \( X \sim F \) under consideration by \( \log X \). In Section 7 we will return to this issue for a brief discussion of practical aspects.

The remainder of this paper is organised as follows. As preparation, (1.5) is generalised to arbitrary Borel sets in \( \mathbb{R} \); this results in a simple large deviation principle, which can take different forms. These are generalised to the multivariate setting in Section 5 in order to provide bounds and limits for probabilities of multivariate extreme events. Section 4 makes the connection to classical multivariate extreme value theory and in particular, to residual tail dependence and related assumptions. Section 5 discusses how the representation of tail dependence can be changed by changing the standardisation of the marginal distributions, and in particular, how we can approximate probabilities of events which are extreme “in any direction” in this manner. Section 6 applies the theory to formulate and analyse estimators for probabilities of multivariate extreme events in the range (1.1), and Section 7 closes with a discussion of the results and of potential applications.

\(^2\) This can always be ensured by adding a constant to \( q_F \), which does not affect the left-hand side of (1.3).

\(^3\) This condition allows \( U(1/p_n) \) with \( p_n \) to be approximated from \( U(n/k_n) \) with \( (k_n) \) an intermediate sequence, so \( U(n/k_n) \) can be estimated consistently by the empirical quantile \( X_{n-k_n+1:n} \) under certain conditions (\textit{e.g.} de Haan & Ferreira (2006), Theorem 2.2.1).
2 The univariate case

Consider a scalar random variable \( X \sim F \). To avoid distraction by technicalities, let \( F \) be continuous; this assumption does not seem restrictive for typical applications of extreme value analysis. The infimum of an (extended) real function \( f \) over a set \( S \) will be written as \( \inf_{x \in S} f(x) \) instead of the conventional but cumbersome \( \inf_{x \in S} f(x) \). Furthermore, \( \inf f(\emptyset) := \infty \). The following holds without further assumptions on \( F \):

\[
\inf(A^c) \leq \liminf_{y \to \infty} y^{-1} \log P(X \in q_F(yA)) \leq \limsup_{y \to \infty} y^{-1} \log P(X \in q_F(yA)) \leq -\inf(\bar{A}). \tag{2.1}
\]

**Proposition 1** For every Borel set \( A \subset [0, \infty) \),

\[
-\inf(A^c) \leq \liminf_{y \to \infty} y^{-1} \log P(X \in q_F(yA)) \leq \limsup_{y \to \infty} y^{-1} \log P(X \in q_F(yA)) \leq -\inf(\bar{A}). \tag{2.1}
\]

**Proof** If \( A^c \) is empty, the lower bound is \(-\infty\). Else, with \( Y := -\log(1 - F(X)) \) and \( \alpha := \inf(A^c) \geq 0 \) and \( \delta > 0 \) such that \((\alpha, \alpha + \delta) \subset A^c, P(Y/y \in A) \geq P(Y/y \in (\alpha, \alpha + \delta)) = e^{-\alpha} \), hence the lower bound in (2.1). The upper bound is proven similarly. \( \square \)

Equation (2.1) is a somewhat trivial example of a large deviation principle (LDP) (e.g. [Dembo & Zeitouni (1998)]). In a similar manner, the GW limit (1.5) is readily generalised as well:

**Proposition 2** Suppose that \( F \) satisfies the GW limit (1.5). Then for every \( \theta > 0 \) and every Borel set \( A \subset [0, \infty) \),

\[
-\inf(A^c) \leq \liminf_{y \to \infty} y^{-1} \log P(X \in q_F(y) + g(y)h_\rho(A)) \leq \limsup_{y \to \infty} y^{-1} \log P(X \in q_F(y) + g(y)h_\rho(A)) = -\inf(\bar{A}). \tag{2.2}
\]

**Proof** The proof is similar to that of Proposition 1 above. In particular for the lower bound, for nonempty \( A^c, \alpha := \inf(A^c) \geq \theta, \delta > 0 \) such that \((\alpha, \alpha + \delta) \subset A^c\) and \( \varepsilon \in (0, \delta/2) \), \( P((X - q_F(y))/g(y) \in \bar{h}_\rho(A^c)) \geq P((X - q_F(y))/g(y) \in \bar{h}_\rho((\alpha, \alpha + \delta)) \geq e^{-\alpha(1 - \varepsilon)}(1 - e^{-\varepsilon(\delta - 2\varepsilon)}) \) provided that \( y \) is large enough, with the last inequality a consequence of (1.5). \( \square \)

For an extension to the multivariate setting, it would be desirable to replace (2.2) by bounds valid for all \( A \subset [0, \infty) \), since a multivariate event we might be concerned with could be extreme in one variable, but not in some other variable. Therefore, defining the quantile approximation \( \tilde{q}_{F,y} \) for \( y > 0 \) by

\[
\tilde{q}_{F,y}(z) := \begin{cases} q_F(z) & \text{if } z \leq y \\ q_F(y) + g(y)h_\rho(z/y) & \text{if } z > y. \end{cases} \tag{2.3}
\]

with \( \rho \) a real number and \( g \) a positive function, the following combination of (2.2) and (2.1) would be useful:

\[
-\inf(A^c) \leq \liminf_{y \to \infty} y^{-1} \log P(X \in \tilde{q}_{F,y}(Ay)) \leq \limsup_{y \to \infty} y^{-1} \log P(X \in \tilde{q}_{F,y}(Ay)) \leq -\inf(\bar{A}). \tag{2.4}
\]

\[
\tilde{q}_{F,y}(z) := \begin{cases} q_F(z) & \text{if } z \leq y \\ q_F(y) + g(y)h_\rho(z/y) & \text{if } z > y. \end{cases} \tag{2.3}
\]
Proposition 3 Suppose that $F$ satisfies the GW limit \((1.5)\). Then \((2.4)\) holds for every Borel set $A \subset [0, \infty)$. 

A proof is omitted, as it is similar to the proofs of the previous propositions. It is straightforward to derive that \((2.4)\) is equivalent to the GW limit, as it implies \((1.5)\) and \((1.3)\).

3 Asymptotic bounds and limits for tail probabilities

Let $X$ be a random vector on $\mathbb{R}^m$ and assume that its marginal distribution functions $F_1, \ldots, F_m$ are continuous. Define the random vector $Y := (Y_1, \ldots, Y_m)$ by

$$Y_j := -\log(1 - F_j(X_j)); \quad (3.1)$$

$Y$ has standard exponential marginals. For $j = 1, \ldots, m$, $qF_j$ (see \((1.2)\)) is injective, by continuity of $F_j$. Therefore, $X$ can be represented as $X = Q(Y)$ with for all $x \in \mathbb{R}^m$,

$$Q(x) := (qF_1(x_1), \ldots, qF_m(x_m)). \quad (3.2)$$

Define a family of probability measures $\{\mu_y; y > 0\}$ on the Borel $\sigma$-algebra of subsets of $\mathbb{R}^m$ by

$$\mu_y(A) := P(X \in Q(yA)) = P(Y \in yA) \quad (3.3)$$

for every Borel set $A \subset \mathbb{R}^m$; for every $y > 0$, the support of $\mu_y$ is in $[0, \infty)^m$.

With $B_\varepsilon(x) := \{x' \in \mathbb{R}^m : \|x - x'\|_\infty < \varepsilon\}$ the open ball of radius $\varepsilon > 0$ around $x \in \mathbb{R}^m$, we will assume that

$$I(x) := \inf_{\varepsilon>0} \inf_{y \to \infty} y^{-1} \log \mu_y(B_\varepsilon(x)) = \inf_{\varepsilon>0} \sup_{y \to \infty} y^{-1} \log \mu_y(B_\varepsilon(x)) \quad (3.4)$$

for all $x \in \mathbb{R}^m$. The function $I$ is known as the rate function; it is nonnegative and lower-semicontinuous. Additional properties follow from \((3.3)\): for every $\varepsilon > 0$ and $x \in \mathbb{R}^m$ with $x_j = \lambda > 0$ for some $j \in \{1, \ldots, m\}$, $y^{-1} \log \mu_y(B_\varepsilon(x)) \leq y^{-1} \log P(Y_j > (\lambda - \varepsilon)y) = \varepsilon - \lambda$, so

$$I(x) \geq \max_{j \in \{1, \ldots, m\}} x_j \quad \forall x \in \mathbb{R}^m. \quad (3.5)$$

This implies that $I$ is a good rate function, meaning that $I^{-1}(0, a]$ is compact for every $a \in [0, \infty)$. Also, since $B_\varepsilon(x\lambda) = \lambda B_{\varepsilon/\lambda}(x)$,

$$I(x\lambda) = \lambda I(x) \quad \forall \lambda > 0, \ x \in \mathbb{R}^m. \quad (3.6)$$

Furthermore, $I(0) = 0$, since $\mu_y(B_\varepsilon(0)) = P(\|Y\|_\infty \leq y\varepsilon) \geq 1 - mP(Y_1 > \varepsilon y) = 1 - me^{-\varepsilon y}$ in \((3.4)\).

Remark 1 By \((3.4)\), $I(x) = \varphi(x)I(x/\varphi(x))$ for every $x \in \mathbb{R}^m \setminus \{0\}$ and every norm $\varphi$ on $\mathbb{R}^m$. This gives for every norm a “spectral representation” of $I$, analogous to the spectral measures in classical extreme value theory (e.g. de Haan & Ferreiró (2006), Section 6.1.4). For example, in the bivariate case, the rate function can be represented as $I(x) = (x_1 + x_2)\varphi(x_2/(x_1 + x_2))$, with the real function $\varphi$ satisfying $\varphi(t) \geq \max(t, 1-t)$ for all $t \in [0, 1]$ because of \((3.5)\).

Remark 2 In the special case that $I$ is subadditive, then by \((3.6)\), it is convex, and furthermore, by \((3.3)\), it is a norm.
Since \( P(Y_1 > y_0) \leq P(\|Y\|_\infty > y_0) \leq mP(Y_1 > y_0) \),
\[
\lim_{y \to \infty} y^{-1} \log P(\|Y\|_\infty > y_0) = -\alpha \quad \forall \alpha \geq 0,
\]
so the maximum of \( Y \) satisfies the same limit as each component of \( Y \). This implies that \( \{\mu_y; y > 0\} \) is exponentially tight (Dembo & Zeitouni (1998)): for every \( \alpha < \infty \), a compact \( E_\alpha \subset \mathbb{R}^m \) exists such that
\[
\limsup_{y \to \infty} y^{-1} \log \mu_y(E_\alpha^c) < -\alpha,
\]
which follows from (3.7) when taking \( E_\alpha = \{x \in \mathbb{R}^m : \|x\|_\infty \leq \alpha + \varepsilon\} \) for some \( \varepsilon > 0 \). Therefore, since \( \{\mu_y; y > 0\} \) satisfies (3.4), it satisfies a large deviation principle (LDP) by Dembo & Zeitouni (1998) (Theorem 4.1.11 and Lemma 1.2.18): for every Borel set \( A \subset \mathbb{R}^m \),
\[
- \inf I(A^c) \leq \liminf_{y \to \infty} y^{-1} \log \mu_y(A) \leq \limsup_{y \to \infty} y^{-1} \log \mu_y(A) \leq - \inf I(A). \tag{3.9}
\]
With (3.6), (3.9) implies the marginal condition
\[
\inf_{x \in \mathbb{R}^m : x_j > \lambda} I(x) = \lambda \quad \forall \lambda \geq 0, \ j \in \{1, \ldots, m\}. \tag{3.10}
\]
In the univariate case, (3.9) reduces to (4.1). This suggests the following multivariate generalisation of (2.4), the large deviation representation of the GW limit: for every Borel set \( A \subset \mathbb{R}^m \),
\[
- \inf \tilde{I}(A^c) \leq \liminf_{y \to \infty} y^{-1} \log \tilde{\mu}_y(A) \leq \limsup_{y \to \infty} y^{-1} \log \tilde{\mu}_y(A) \leq - \inf \tilde{I}(A). \tag{3.11}
\]
with \( \tilde{\mu}_y(A) := P(X \in \tilde{Q}_y(yA)) \) and \( \tilde{Q}_y(x) := (\tilde{q}_{F_{1,y}}(x_1), \ldots, \tilde{q}_{F_{m,y}}(x_m)) \) for all \( x \in \mathbb{R}^m \) and \( y > 0 \) with
\[
\tilde{q}_{F_{j,y}}(z) := \begin{cases} q_{F_{j}}(z) & \text{if } z \leq y \\ q_{F_{j}}(y) + g_{j}(y)h_{\rho_{j}}(z/y) & \text{if } z > y \end{cases} \tag{3.12}
\]
for some real numbers \( \rho_1, \ldots, \rho_m \) and positive functions \( g_1, \ldots, g_m \), and with \( \tilde{I} \) satisfying the same marginal constraints as \( I \) in (3.10):
\[
\inf_{x \in \mathbb{R}^m : x_j > \lambda} \tilde{I}(x) = \lambda \quad \forall \lambda \geq 0, \ j \in \{1, \ldots, m\}. \tag{3.13}
\]

Theorem 1 below addresses the relationship between (4.11) and (3.9). For its proof, recall that two families of random vectors \( \{\beta_y; y > 0\} \) and \( \{\zeta_y; y > 0\} \) on \( \mathbb{R}^m \) defined on the same probability space are said to be exponentially equivalent (Dembo & Zeitouni (1998)) if \( y^{-1} \log P(\|\beta_y - \zeta_y\|_\infty > \varepsilon) \to -\infty \) as \( y \to \infty \) for every \( \varepsilon > 0 \).

**Theorem 1** (a) If \( \{\mu_y; y > 0\} \) satisfies (3.4) and the marginals satisfy GW tail limits, so
\[
\lim_{y \to \infty} \frac{q_{F_j}(y\lambda) - q_{F_j}(y)}{g_j(y)} = h_{\rho_j}(\lambda) \quad \forall \lambda > 0, \ j \in \{1, \ldots, m\} \tag{3.14}
\]
for some real numbers \( \rho_1, \ldots, \rho_m \) and positive functions \( g_1, \ldots, g_m \), then the LDP (3.11) holds with good rate function \( I = \tilde{I} \), with \( I \) satisfying (3.9), (3.13) and \( I(0) = 0 \).
(b) If \( \{ \tilde{\mu}_y; y > 0 \} \) satisfies the LDP \((3.11)\) with \((3.12)\) for real numbers \( p_1, \ldots, p_m \) and positive functions \( g_1, \ldots, g_m \) and \( \bar{I} \) satisfies \((3.12)\), then the marginals satisfy GW tail limits \((3.13)\), and \( \{ \mu_y; y > 0 \} \) satisfies the LDP \((3.9)\) with good rate function \( I = \bar{I} \) satisfying \((3.13)\), \((3.14)\), \((3.15)\), \((3.16)\), \((3.17)\) and \( I(0) = 0 \).

**Proof** Convergence in \((3.13)\) is locally uniform in \( \lambda \) (e.g. de Haan & Ferreira, 2006), B.1.4 and B.2.9, so

\[
\lim_{y \to \infty} \sup_{\lambda \in [A^{-1}, A]} \max_{j \in \{1, \ldots, m\}} \left| h_{p_j}^{-1} \left( \frac{q_{F_j}(y\lambda) - q_{F_j}(y)}{g_j(y)} \right) - \lambda \right| = 0 \quad \forall A > 1. \tag{3.15}
\]

For every \( y > 0 \), \( \tilde{Q}_y \) is injective, so we can define the random vector

\[
\tilde{Y}_y := \tilde{Q}_y^{-1}(X) = \tilde{Q}_y^{-1}Q(Y). \tag{3.16}
\]

For every \( A > 1 \) and \( \delta > 0 \), there exists some \( y_{A, \delta} > 0 \) such that \( \| \tilde{Y}_y - Y \|_{\infty} > \delta y \) implies \( \| Y \|_{\infty} > A y \) for all \( y \geq y_{A, \delta} \), due to \((3.15)\). Therefore, by \((3.14)\),

\[
\limsup_{y \to \infty} y^{-1} \log P(\| \tilde{Y}_y - Y \|_{\infty} > \delta y) \leq -A.
\]

and since \( A > 1 \) is arbitrary, \( \{ \tilde{Y}_y/y, y > 0 \} \) and \( \{ Y/y, y > 0 \} \) are exponentially equivalent. Therefore, since \( \{ \mu_y; y > 0 \} \) satisfies the LDP \((3.9)\) with good rate function \( I \), Dembo & Zeitouni (1998) (Theorem 4.2.13) implies that \( \{ \tilde{\mu}_y; y > 0 \} \) satisfies the same LDP, i.e., \((3.11)\) with \( \bar{I} = I \).

To prove (b), note that \( \lim_{y \to \infty} y^{-1} \log(1 - F_j(y) + g_j(y)h_{p_j}(\lambda)) = -\lambda \) for all \( \lambda \geq 1 \) and \( j \in \{1, \ldots, m\} \) by \((3.11)\) and \((3.12)\), so \((3.13)\) holds. As in the proof of (a), this implies exponential equivalence of \( \{ \tilde{Y}_y/y, y > 0 \} \) and \( \{ Y/y, y > 0 \} \). Moreover, \((3.13)\) implies that \( \bar{I}(x) \geq \max_{j \in \{1, \ldots, m\}} x_j \), so \( \bar{I} \) is a good rate function. An application of Dembo & Zeitouni (1998) (Theorem 4.2.13) completes the proof. \( \square \)

The relationship between the LDP’s \((3.11)\) and \((3.12)\) is the large deviations analogue of a similar relationship in classical extreme value theory (e.g. Resnick, 1987, Proposition 5.10 and 5.15). Theorem 1 justifies the view of \((3.9)\) as representation of tail dependence within the context of the LDP \((3.11)\), which also represents the marginal tails.

\( A \) is called a continuity set of \( I \) (Dembo & Zeitouni 1998) if \( \inf(\bar{I}) = \inf I(A^c) \), so \((3.9)\) reduces to

\[
\lim_{y \to \infty} y^{-1} \log \tilde{\mu}_y(A) = - \inf I(A). \tag{3.17}
\]

\( A \) is a continuity set of \( I \) if \( I \) is continuous and \( A \subset A^c \), for example. Homogeneity \((3.10)\) of \( I \) relaxes the conditions for being a continuity set. For example, without assuming continuity of \( I \), \( A \) is a continuity set if \( \inf I(\bar{A}) = I(x) \) for some \( x \in \overline{A^c} \cap \cup \lambda > 0 \lambda A^c \).

In the remainder of this article, we will discuss continuity sets of rate functions without considering particular conditions which make them continuity sets.
4 Connection to residual tail dependence, and short-range approximation of probabilities of tail events

There is an interesting connection between the theory of Section 3 and residual tail dependence (RTD), introduced in Ledford & Tawn (1996, 1997, 1998); see also de Haan & Ferreira (2006), Section 7.6. In the bivariate case, RTD offers a model of tail dependence within the classical domain of asymptotic independence of component-wise maxima (de Haan & Ferreira (2006), Section 7.6). For a random vector \( X \) on \( \mathbb{R}^m \) with continuous marginals \( F_1, \ldots, F_m \), it amounts to the assumption that for some positive function \( S \) on \( (0, \infty)^m \), for all \( x \in (0, \infty)^m \),

\[
\lim_{t \to \infty} \frac{1}{P((1 - F_j(X_j))^{-1} \geq t \forall j \in \{1, \ldots, m\})} = S(x) > 0. \tag{4.1}
\]

The limit \( S \) satisfies \( S(I) = 1 \), with \( I \) the vector in \( \mathbb{R}^m \) with all its components equal to 1. Furthermore, the denominator in (4.1) must be regularly varying, so \( S(1\lambda) = \lambda^{-1/\eta} \) for all \( \lambda > 0 \) with \( \eta \in (0, 1] \) the residual dependence index, and by (4.1),

\[
S(x\lambda) = \lambda^{-1/\eta} S(x) \quad \forall x \in (0, \infty)^m, \lambda > 0. \tag{4.2}
\]

RTD is an example of short-range tail regularity, dealing with extrapolation of the survival function over marginal distances corresponding to fixed factors \( 1/x_1, \ldots, 1/x_m \) in probability of exceedance. This stands in contrast to long-range tail regularity represented by the tail LDP, which concerns extrapolation over a fixed factor in the logarithm of the probability of a multivariate event (e.g. for a continuity set \( A \) of \( I \) satisfying \( \inf I(A) \in (0, \infty), \lim_{y \to \infty} \mu_y(A)/\mu_y(A) = \lambda \) for all \( \lambda > 0 \) by (3.17)). Define

\[
Y_{\lambda} := \min_{j \in \{1, \ldots, m\}} Y_j, \tag{4.3}
\]

with \( Y \) defined by (3.1), and let \( H_{\lambda} \) be the distribution function of \( Y_{\lambda} \). Define for all \( a \in \mathbb{R}^m \)

\[
A_a := \{x \in \mathbb{R}^m : x_j > a_j \forall j \in \{1, \ldots, m\}\}. \tag{4.4}
\]

For real functions, a one-sided smoothness condition \( \mathcal{L} \) is defined as follows (cf. Bingham et al (1987), Subsection 1.7.6): \( f \in \mathcal{L} \) if \( f \) is nondecreasing, absolutely continuous, and its derivative \( f' \) satisfies

\[
\max \left( \lim_{y \to \infty} \liminf_{\lambda \to 1} \inf_{y \in [1, \lambda]} \frac{f'(y\lambda)}{f'(y)}, \lim_{y \to \infty} \liminf_{\lambda \to 1} \inf_{y \in [1, \lambda]} \frac{f'(y)}{f'(y\lambda)} \right) \geq 1. \tag{4.5}
\]

RTD and the tail LDP (3.9) are related as follows.

Proposition 4 (a) RTD (4.4) implies

\[
\lim_{y \to \infty} y^{-1} \log P(X \in Q(yA_{1\lambda})) = \lim_{y \to \infty} y^{-1} \log P(Y \in yA_{1\lambda}) = -\lambda/\eta \quad \forall \lambda > 0, \tag{4.6}
\]

with \( \eta \) the residual dependence index of \( X \).

(b) If \( X \) satisfies an LDP (3.9) and also RTD, then \( I(1\lambda) = \lambda/\eta \) for all \( \lambda > 0 \).

(c) If \( X \) satisfies an LDP (3.9) and in addition, \(- \log(1 - H_{1\lambda}) \in \mathcal{L} \), then (4.4) holds for \( x = 1\lambda \) for all \( \lambda > 0 \) with \( \eta = 1/I(1) \).
Proof By (4.1), the survival function $1 - H_\lambda \circ \log$ of the random variable $\exp Y_\lambda$ is regularly varying with index $-1/\eta$. Therefore, $f := 1/(1 - H_\lambda \circ \log) \in RV_{1/\eta}$, so by the Potter bounds (Bingham et al. (1987)), for every $\varepsilon \in (0, 1/\eta)$, there is $z_\varepsilon > 0$ such that $(1 - \varepsilon)(x/z)^{1/\eta - \varepsilon} \leq f(z)/f(x) \leq (1 + \varepsilon)(x/z)^{1/\eta + \varepsilon}$ for all $z \geq z_\varepsilon$ and $x \geq z$. Taking the logarithm and substituting $e^{\theta x}$ for $x$ gives
\[
\lim_{y \to \infty} y^{-1} \log f(e^{\theta x}) \to \lambda/\eta \quad \forall \lambda > 0,
\]
so
\[
\lim_{y \to \infty} y^{-1} \log(1 - H_\lambda(y\lambda)) = \lim_{y \to \infty} y^{-1} \log P(Y_\lambda > y\lambda) = -\lambda/\eta \quad \forall \lambda > 0,
\]
(4.7)
and (a) follows. Assertion (b) follows from (a) and the uniqueness of the rate function of an LDP (Dembo & Zeitouni (1998), Lemma 4.1.4). For (c), note that due to (4.6), the LDP (4.9) implies (4.7), so $w(y) := -\log(1 - H_\lambda(y)) \sim y/\eta$ as $y \to \infty$ for $\eta = 1/(1)$. Therefore, by Bingham et al. (1987) (Theorem 1.7.5, using (1.7.10)), the condition $w \in L$ implies $w(y) \to 1/\eta$ and averaging, $w(y + r) - w(y) \to r/\eta$ as $y \to \infty$ for every $r \in \mathbb{R}$. This implies (4.1) for $x = 1\lambda$ for all $\lambda > 0$. □

In Proposition 4(c), the condition $-\log(1 - H_\lambda) \in L$ (see (4.5)) can be replaced by alternative conditions of similar nature as in Bingham et al. (1987) (Theorem 1.7.5).

Proposition 4 shows that RTD implies a limited LDP-like condition and in turn, that the LDP (5.9) with an additional smoothness condition implies an RTD-like condition; the smoothness makes it possible to derive short-range regularity from the long-range regularity represented by the LDP.

RTD provides only a partial description of the tail, since (4.1) only describes the joint survival function, and this only within domains of the form \( \{z \in \mathbb{R}^m : (1 - F_j(z_j))^{-1} > t \min(x_j, 1) \ \forall j \in \{1, \ldots, m\} \} \) as $t \to \infty$. In terms of $Y$ with standard exponential marginals, (4.1) characterises the limiting behaviour of the joint survival function of $Y$ within zones of finite width around the diagonal: for all $z \in \mathbb{R}^m$,
\[
\lim_{y \to \infty} P(Y_j > y + z_j \ \forall j \in \{1, \ldots, m\}) = S(e^z)
\]
(4.8)
with $e^z := (e^{z_1}, \ldots, e^{z_m})$. This leaves a gap in the description between these zones and the boundary of $[0, \infty)^m$ where the marginals are specified. In contrast, the LDP provides a complete characterisation of a multivariate tail which includes the marginal tails. Moreover, under a generalisation of the smoothness condition $-\log(1 - H_\lambda) \in L$ in Proposition 4(c), the LDP implies an alternative representation of short-range tail dependence:

Theorem 2 Assume (4.4), so the LDP (5.9) applies. To any Borel set $A \subset \mathbb{R}^m$ which is a continuity set of $I$ and satisfies the smoothness condition
\[
(y \mapsto -\log \mu_y(A)) \in L,
\]
(see (4.3)), the following limit applies:
\[
\lim_{t \to \infty} \frac{P(X \in Q(A \log(t\lambda)))}{P(X \in Q(A \log t))} = \lambda^{-\inf I(A)} \quad \forall \lambda > 0,
\]
(4.9)
with $I$ satisfying (3.7), (3.8), (3.10) and $I(0) = 0$. In particular, for every $a \in [0, \infty)^m$ such that

$$\lim_{t \to \infty} \frac{P((1 - F_j(X_j))^{-1} > (t\lambda)^a_j, \forall j \in \{1, \ldots, m\})}{P((1 - F_j(X_j))^{-1} > t^\lambda, \forall j \in \{1, \ldots, m\})} = \lambda^{-\inf I(A_\alpha)} \quad \forall \lambda > 0. \quad (4.10)$$

**Proof** For a a continuity set of $I$, (3.9) implies (3.17) and (4.9) is obtained in the same manner as in the proof of Proposition 4(c). In particular, $A_\alpha$ is a continuity set of $I$ for every $a \in [0, \infty)^m$, so (4.10) follows directly from (4.9). \( \square \)

Due to (3.9), application of (4.5) or (4.10) only requires knowledge of $I$ on $\{x \in [0, \infty)^m : \rho(x) = 1\}$ for some norm $\rho$ on $\mathbb{R}^m$. The marginals are described by (4.9) as well, as can be seen from (4.10) and (3.10). In the special case of $a = I$, (4.10) becomes equivalent to (4.11) with $x = I\lambda$ and $\eta = 1/I(t)$, so on the diagonal, the limit (4.10) and RTD (4.1) agree; elsewhere, they differ.

Defining a function $\kappa$ by $\kappa(a) := \inf I(A_\alpha)$ for every $a \in [0, \infty)^m$, (4.10) becomes identical to an extension of RTD recently introduced in Wadsworth & Tawn (2013). Wadsworth & Tawn (2013) proposed this extension to cover the gap between the boundary of $[0, \infty)^m$ and the fixed-width zones around the diagonal where RTD in the form (4.9) may apply. Furthermore, in the bivariate case and under appropriate smoothness conditions, Wadsworth & Tawn (2013) show that for a fixed $a \in [0, \infty)^2 \setminus (0, 0)$, (4.10) implies for every Borel set $B \subset [1, \infty)^2$, with $V_j := (1 - F_j(X_j))^{-1} = \exp{Y_j}$,

$$\lim_{t \to \infty} \frac{P((V_1 t^{-\alpha_1}, V_2 t^{-\alpha_2}) \in B)}{P((V_1 t^{-\alpha_1}, V_2 t^{-\alpha_2}) \in [1, \infty)^2)} = \int_B d(x(1)^{-\kappa_1(a)x(2)^{-\kappa_2(a)}}) \quad (4.11)$$

with $\kappa_j(a) := \frac{\partial \kappa_j(a)}{\partial \alpha_j}$, which by (3.7) must satisfy that $\kappa_j(a) = \kappa_j(a\lambda)$ for all $\lambda > 0$, $a \in [0, \infty)^2 \setminus (0, 0)$ and $j \in \{1, 2\}$. Eq. (4.11) expresses a weak limit of probability measures on $[1, \infty)^2$ corresponding to two independent Pareto variables. For $a \in [0, \infty)^2 \setminus (0, 0)$ such that $\kappa_1(a) > 0$ and $\kappa_2(a) > 0$, (4.11) is nonzero and therefore, since $\kappa(a) = a_1\kappa_1(a) + a_2\kappa_2(a)$, it implies

$$\lim_{t \to \infty} P((V_1 t(\lambda)^{-\alpha_1}, V_2 t(\lambda)^{-\alpha_2}) \in B) = \lambda^{-\kappa(a)}. \quad (4.12)$$

Wadsworth & Tawn (2013) suggest that expressions equivalent to (4.11) and (4.12) might be used to estimate probabilities for sets not of the form (4.4). However, they also note that in an approximation or estimation context, $a$ would have to be chosen, so it is ambiguous. Therefore, it remains unclear how a consistent estimator could be developed from (4.12).

However, it appears that the limit (4.12), inherited from the LDP (3.9), provides a solution to this problem. It offers a straightforward recipe for approximation of probabilities $P(X \in Q(A \log(\lambda))) = P(V \in (t\lambda)^A)$ for all $\lambda > 1$ and all Borel sets $A \subset \mathbb{R}^m$ satisfying the conditions of Theorem 2 which is readily turned into an estimator (we will not pursue this further, focusing instead on estimation based on the LDP (3.9) later in Section 9).

In (4.11), it is not $\kappa$, but the rate function $I$ which determines the attenuation rate. For any $a \in [0, \infty)^m$, $I(a)$ and $\kappa(a)$ are identical only if $I(a + x) \ge I(a)$ for all $x \in [0, \infty)^m$. This condition is rather restrictive, as a rate function resembles a density more than it resembles a survival function; see definition (4.3). As an illustration, for
X ∼ \mathcal{N}(0, T) with T a correlation matrix, I(a) = \sum_{j,i \in \{1, \ldots, m\}} (T^{-1})_{ij} \sqrt{a_i a_j} for all a ∈ [0, \infty)^m. In the bivariate case with T_{12} = T_{21} := t and x := (1, 0), therefore, I(x) = (1 - t^2)^{-1}, so I(x) > 1 = \kappa(x) for every t ∈ (0, 1) by (5.10). More generally, in the bivariate case with differentiable \kappa, I(a) > \kappa(a) implies that \kappa_1(a) = 0 or \kappa_2(a) = 0 and therefore, (5.11) yields zero, so (5.12) is not valid.

After this excursion into the topic of short-range tail dependence, we now return to long-range tail dependence described by the LDP (3.9).

5 Changing the standardisation of the marginals

An LDP similar to (3.9) is obtained when we replace the standard exponential marginals of Y by another distribution G which is continuously increasing on its support (0, \infty) and satisfies a Weibull tail limit (cf. (4.6)) with index \tau > 0, so q_G = G^{-1}(1 - e^{-\tau t}) satisfies q_G ∈ RV_{\tau} for every Borel set I. Then for every continuously increasing distribution function G with support in (0, \infty) and Weibull tail limit with index \tau > 0, we have

\text{Theorem 3} Suppose that \{\mu_y; y > 0\} defined by (5.2) is exponentially tight and satisfies the LDP

\begin{align*}
-\inf_{y \to \infty} \mu_y(A) &\leq \lim \inf_{y \to \infty} y^{-1} \log \mu_y(A) \\
&\leq \lim \sup_{y \to \infty} y^{-1} \log \mu_y(A) \\
&\leq -\inf_{\lambda \to \infty} I_G(\lambda)
\end{align*}

for every Borel set A ⊂ \mathbb{R}^m, with I_G := I \circ Id^{/r} a good rate function;

(b) for all x ∈ \mathbb{R}^m, I_G(x) ≥ \max_{j \in \{1, \ldots, m\}} |x_j|^{1/r} and

I_G(0) = 0, \quad \text{and inf}_{x \in \mathbb{R}^m, x_j > 0} I(x) = \lambda^{1/r} for all \lambda > 0 and j \in \{1, \ldots, m\};

(c) a Borel set A ⊂ \mathbb{R}^m is a continuity set of I_G if and only if A^{1/r} := \{x \in \mathbb{R}^m : x^i \in A\} is a continuity set of I;

(d) if all marginals of F satisfy GW tail limits, i.e., (3.14), then

\begin{align*}
-\inf_{y \to \infty} \mu_y(A) &\leq \lim \inf_{y \to \infty} y^{-1} \log \tilde{\mu}_y(A) \\
&\leq \lim \sup_{y \to \infty} y^{-1} \log \tilde{\mu}_y(A) \\
&\leq -\inf_{\lambda \to \infty} I_G(\lambda)
\end{align*}

for all Borel sets A ⊂ \mathbb{R}^m, with

\tilde{\mu}_y(A) := P(X ∈ Q_{G,y}(q_G(y)A))

and for all y > 0, \tilde{Q}_{G,y} := Q_{G,y}^{-1} with Q_{G,y} defined in (5.7).
Theorem 3 above provides some freedom in choosing the standardised marginals in which to express tail dependence.

**Proof** By Dembo & Zeitouni (1998) (Theorem 4.2.13), the LDP \( \text{LDP} \) with good rate function \( I \) for the random vectors \( \{Y/y; y > 0\} \) carries over to any exponentially equivalent family of random vectors and therefore, by Lemma 1(b), to \( \{(Z/q_G(y))/y; y > 0\} \). Therefore, \( \text{LDP} \) holds with \( \mu_G(A) \) replaced by \( P((Z/q_G(y))/y, y > 0) = P(X \in Q_G(q_G(y)A)) \). By the contraction principle (Dembo & Zeitouni 1998, Theorem 4.2.1), this is equivalent to \( \text{LDP} \) with \( I_G \) defined under (a). Moreover, \( \mu_G(y; y > 0) \) is exponentially tight, due to Lemma 1(a). The properties of \( I_G \) under (b) and (c) follow from its definition and \( 5.4, 5.5, 5.6 \). The proof of (d) extends the proof of Theorem 1(a): by exponential equivalence of \( \{Y/y, y > 0\} \) and \( \{\tilde{Y}_y/y, y > 0\} \) and regular variation of \( q_G \), \( \{q(y)/q_G(y); y > 0\} \) and \( \{q_G(Y)/q_G(y), y > 0\} \) are also exponentially equivalent. Therefore, noting that \( \mu_G(y; y > 0) = P(q_G(Y)/q_G(y) \in A) \) and \( \mu_G(y; y > 0) = P(q_G(Y)/q_G(y) \in A) \), \( \mu_G(y; y > 0) \) satisfies the same LDP as \( \mu_G(y; y > 0) \), again by (Dembo & Zeitouni 1998) (Theorem 4.2.13).

We may further modify \( G \) to extend the support of \( \mu_G(y; y > 0) \) and \( \mu_G(y; y > 0) \) from \( [0, \infty)^m \) to \( \mathbb{R}^m \), extending the multivariate tail approximation from only “high” events to all “peripheral” events in \( \mathbb{R}^m \), i.e., to a tail approximation in all directions. In order to do this, take \( G \) symmetric about 0, i.e., \( G = 1 - G(0) \), as well as continuously increasing and satisfying a Weibull tail limit cf. 1.6. Therefore, the marginals of \( Z \) defined by \( 5.1 \) satisfy

\[
P(Z_j > q_G(y)) = P(-Z_j > q_G(y)) = e^{-y}
\]

for all \( y \geq 2 \) and \( j \in \{1, \ldots, m\} \). As before, \( Q_G \) is defined by \( 5.2 \) and \( \mu_G(y; y > 0) \) is defined by \( 5.3 \). One possible choice for \( G \) is the Laplace distribution \( L \), which satisfies \( L(z) = L(-z) = \frac{1}{2} \exp(-z) \) for all \( z \geq 0 \). \( L \) can be regarded as the symmetric variant of the exponential distribution, satisfying \( q_L(y) = y - 2 \) for all \( y \geq 2 \) and \( \lim_{y \to \infty} y^{-1} \log(1 - L(q_L(y)\lambda)) = \lim_{y \to \infty} y^{-1} \log L(-q_L(y)\lambda) = -\lambda \) for all \( \lambda \geq 0 \).

**Theorem 4** For \( X \) a random vector in \( \mathbb{R}^m \) with continuous marginals and \( L \) the Laplace distribution, assume that \( \mu_{L,y; y > 0} \) satisfies

\[
I_L(x) := \inf_{\varepsilon > 0} \sup_{y \to \infty} \inf_{\varepsilon > 0} y^{-1} \log \mu_{L,y}(B_\varepsilon(x)) = \inf_{\varepsilon > 0} \sup_{y \to \infty} y^{-1} \log \mu_{L,y}(B_\varepsilon(x))
\]

for all \( x \in \mathbb{R}^m \). Then for every continuously increasing symmetric distribution function \( G \) satisfying a Weibull tail limit with index \( r > 0 \),

(a) \( \mu_{G,y; y > 0} \) satisfies the LDP for \( \text{LDP} \) with good rate function \( I_G := I_L \circ \text{Id}^{1/r} \);
(b) the assertions under Theorem 3 hold, and in additional,

\[
\inf_{x \in \mathbb{R}^m : x_j < -\lambda} I_G(x) = \lambda^{1/r} \quad \forall \lambda > 0, j \in \{1, \ldots, m\}.
\]

(c) if all marginals of \( F \) satisfy upper and lower GW tail limits (i.e., \( 1.3 \) holds with \( 1.4 \) for \( F = F_j \)), \( q = q_j^+ \) and \( \rho = \rho_j^+ \), and for \( F = 1 - F_j (-\text{Id}) \), \( q = q_j^− \) and \( \rho = \rho_j^− \) for \( j = 1, \ldots, m \), then the LDP \( 5.7 \) holds with \( \tilde{\mu}_{G,y} \) defined by \( 5.7 \), with for all \( y > 0 \) and \( x \in \mathbb{R}^m \), \( \tilde{Q}_G(y,x) = \left( \tilde{Q}_{G,1,y}(x_1), \ldots, \tilde{Q}_{G,m,y}(x_m) \right) \) and

\[
\tilde{Q}_{G,j,y}(x) := \begin{cases} \tilde{q}_{F_j,y}(-\log(1 - G(z))) & \text{if } z \geq 0 \\ \tilde{q}_{1-F_j(-\text{Id}),y}(-\log G(z)) & \text{if } z < 0. \end{cases}
\]

12 Cees de Valk
Proof Apart from readily verifiable identities, the proof involves a minor and straightforward adaptation of the derivation of (3.9) and the proof of Theorem 3. □

Remark 3 Theorem 4 remains valid after replacing $L$ by any continuously increasing symmetric distribution function $G$ satisfying a Weibull tail limit with index $\varrho > 0$, using $I_G := I_G \circ \ell /r$ in (a).

Remark 4 The standard normal distribution $\Phi$ satisfies the conditions on $G$ in Theorem 4 if $\Phi$ seems a particularly attractive choice for $G$: if the components of $X$ are independent, then $I_\Phi = \|\|^2_2$, which is rotation-invariant and convex.

6 Simple estimators for very small probabilities

We are now ready to apply the theory of Sections 5 and 6 to the problem of estimation of probabilities of extreme events $p_n$ satisfying (6.1) from $X^{(1)}, \ldots, X^{(n)}$; with $X^{(1)}, X^{(2)}, \ldots$ a sequence of iid random vectors in $\mathbb{R}^m$ with a distribution function $F$ having continuous marginals $F_1, \ldots, F_m$. Generalising (6.1) to $\tau_2 > \tau_1 > 0$, we will focus on events of the form $B_n := Q_G(q_G(y_n), A)$ for certain $A \subset \mathbb{R}^m$, with $y_n := -\log p_n$ and $G$ some distribution function as in Theorem 4. Under the conditions for Theorem 4 including marginal upper and lower GW tail limits as in Theorem 4(c), the tail LDP (6.6) with (6.7) and (6.10) applies. For every Borel set $A \subset \mathbb{R}^m$ which is a continuity set of $I_G$ satisfying that $\inf I_G(A) \in (0, \infty)$, therefore, $-\log P(X \in B_n) \sim y_n \inf I_G(A)$ (showing that eventually, $P(X \in B_n) \in [n^{-t_2}, n^{-t_1}]$ for some $t_2 > t_1 > 0$) and $-\log P(Q_G(y_n)(Z/\ell) \in B_n) \sim y_n \inf I_G(A)$ for all $\ell > 0$, so by (6.5),

$$
\lim_{n \to \infty} \frac{\log P(X \in B_n)}{\ell^{-1/r} \log P(Q_G(y_n)(Z/\ell) \in B_n)} = 1 \quad \forall \ell > 0.
$$

This limit suggests that the logarithm of $P(X \in B_n)$ could be estimated by modifying the denominator in (6.1) as follows: choose for $G$ a particular distribution function $G$ with Weibull tail index $\varrho > 0$ and satisfying the conditions on $G$ in Theorem 4 and then replace $r$ by $\varrho$ and $P(Q_G(y_n)(Z/\ell) \in B_n)$ by an estimator of $P(Q_G(y_n)(Z/\ell) \in B_n)$ with $Z := Q_G^{-1}(X)$, and $Q_G(y_n)$ an estimator of $Q_G$. The hope is that for each $n$, the value of $\ell$ can be chosen in such a manner that consistency is assured.

Estimation of $Q_G$ (see (6.2)) boils down to a univariate quantile estimation problem, so we will proceed to examine this first. For real scalar iid random variables $X^{(1)}, X^{(2)}, \ldots$ with continuous distribution function $F$ satisfying the upper GW tail limit (6.3), and with $X_{1:n} \leq \ldots \leq X_{n:n}$ the order statistics of the sample $X^{(1)}, \ldots, X^{(n)}$, define a quantile estimator $\hat{q}_{n,X}$ for $q_F$ defined by (1.2) as

$$
\hat{q}_{n,X}(y) := \begin{cases} 
X_{\lceil n(1-e^{-y})\rceil+1:n} & \text{if } y \in (0, y_n] \\
X_{n-k_0(n)+1:n} + \hat{\varrho}(n)h(\varrho(n)/y_n) & \text{if } y > y_n, 
\end{cases}
$$

with $\hat{\varrho}(n)$ and $\hat{\varrho}(n)$ estimators for $\varrho$ and $\varrho$ in (1.3), respectively, $k_0 : \mathbb{N} \to \mathbb{N}$ nondecreasing and such that $k_0(n) \in \{1, \ldots, n\}$ for all $n \in \mathbb{N}$, and

$$
y_n := \log(n/k_0(n)).
$$
The only assumption we will make on the quantile estimator is that the probability-based quantile estimation error $\hat{\nu}_{n,X}$, defined by

$$\hat{\nu}_{n,X}(y) := \frac{q_F^{-1}\hat{q}_{n,X}(y)}{q_F^{-1}q_F(y)} - 1 = y^{-1}q_F^{-1}\hat{q}_{n,X}(y) - 1$$

for $y > 0$, satisfies

$$\lim_{n \to \infty} \sup_{\lambda \in [1,A]} |\hat{\nu}_{n,X}(y,\lambda)| = 0 \ a.s. \ \forall A > 1.$$  

(6.6)

Very simple estimators $\hat{\rho}(n)$ and $\hat{g}(n)$ such that $\hat{q}_{n,X}$ given by (5.3) satisfies this requirement were considered in de Valk (2014) (Theorem 4): taking

$$\hat{\rho}(n) := \log \left( \frac{X_{n-k_1(n)+1:n} - X_{n-k_2(n)+1:n}}{X_{n-k_1(n)+1:n} - X_{n-k_0(n)+1:n}} \right), \ \ \ \hat{g}(n) := \frac{X_{n-k_1(n)+1:n} - X_{n-k_0(n)+1:n}}{h_{\hat{\rho}(n)}(\xi)}$$

with

$$k_i(n) := \left[ (n/k_0(n))^{-\xi} \right] \text{ for } i \in \{1,2\}$$

for some $\xi > 1$, and with $k_0$ satisfying that $\limsup_{n \to \infty} \log k_0(n)/\log n < 1$ and $\lim_{n \to \infty} k_2(n)/\log n = \infty$, (6.6) is ensured.

Returning to the case of $iid$ copies $X^{(1)}, X^{(2)}$, ..., of a random vector $X$ in $\mathbb{R}^m$, let for $j \in \{1,..,m\}$ $X_j,1:n \leq ... \leq X_j,n:n$ denote the marginal order statistics derived from the marginal sample $X_j^{(1)},...,X_j^{(n)}$. Now define the following GW-based estimator $\hat{Q}_{G,j,n}$ for $Q_{G,j}$ (compare (6.10):

$$\hat{Q}_{G,j,n}(x) := (\hat{Q}_{G,1,n}(x_1),...,\hat{Q}_{G,m,n}(x_m))$$

with $\hat{q}_{n,X}$ defined by (6.3), and let $\hat{Q}_{G,n}(x) := (\hat{Q}_{G,1,n}(x_1),...,\hat{Q}_{G,m,n}(x_m)) \in \mathbb{R}^m$ for every $x \in \mathbb{R}^m$. Furthermore, define the following estimator of $2^{(i)}_{j,n}$ defined by (6.2):

$$\hat{z}^{(i)}_{j,n} := \hat{G}^{-1}(\{R^{(i)}_{j,n} - \xi\}/n)$$

with $R^{(i)}_{j,n} := \sum_{i=1}^{n} \mathbf{1}(X^{(i)}_{j} \leq X^{(i)}_{j,n})$ for $i = 1,..,n$ and $j = 1,..,m$. For every sequence of $n$ random vectors $f^{(1)}, f^{(2)},...,f^{(n)}$ in $\mathbb{R}^m$ and every $B \subset \mathbb{R}^m$, define

$$\hat{\rho}_n(f \in B) := n^{-1} \sum_{i=1}^{n} \mathbf{1}(f^{(i)} \in B),$$

and for some chosen $\theta > 0$, determine a value of the analogue of $\ell$ in (6.1) as

$$\ell^+_\theta,n(B) := \sup\{t > 0 : \hat{\rho}_n(\hat{Q}_{G,n}(Z_{n,t}) \in B) \geq (k_0(n)/n)^\theta \}$$

with the convention that $\sup\{0\} = 0$. Now consider the following estimator for $P(X \in B)$:

$$P^\theta_{G,n}(X \in B) := (k_0(n)/n)^\theta (\ell^+_\theta,n(B))^{-1/\theta}$$

(6.13)

Shortly, we will verify that if $B_{\sigma} := Q_G(q_G(\sigma \log n)A)$ is substituted for $B$, with $G$ any continuous symmetric distribution function satisfying a Weibull tail limit, then under mild restrictions on $A$ and $k_0$, this estimator converges in the large deviation sense for all $\sigma > 0$. 


Theorem 5 Let \( X^{(1)}, X^{(2)}, \ldots \) be iid copies of a random vector \( X \) on \( \mathbb{R}^m \) satisfying the conditions of Theorem 4 including continuous marginals with upper and lower GW tail limits as in Theorem 4(c). Let \( G \) be a continuously increasing and symmetric distribution function satisfying a Weibull tail limit with index \( \varrho > 0 \). For \( k_0 : \mathbb{N} \to \mathbb{N} \) satisfying

\[
0 < c' := \liminf_{n \to \infty} \frac{\log k_0(n)}{\log n} \leq \limsup_{n \to \infty} \frac{\log k_0(n)}{\log n} =: c < 1,
\]

consider the estimator (6.18) for \( P(X \in B) \), with the quantile estimator (6.3) satisfying (6.6) and \( \theta \in (0, (1 - c')^{-1}) \). Then for \( B_{n, \tau} := Q_G(\varrho \tau \log n)A \), with \( A \) any continuously increasing and symmetric distribution function satisfying a Weibull tail limit and \( \tau \in \mathbb{R}^m \) any Borel set which is a continuity set of \( I_G \) defined in Theorem 4(a) satisfying inf \( I_G(A) \in (0, \infty) \),

\[
\lim_{n \to \infty} \sup_{\tau \in [T^{-1}, T]} \left| \frac{\log \hat{P}_{G,n}(X \in B_{n, \tau})}{\log P(X \in B_{n, \tau})} - 1 \right| = 0 \quad \forall T > 1.
\]

Proof The proof can be found in Subsection 8.2.

Remark 5 By Theorem 11 \( P(X \in B_{n, \tau}) = n^{-\varrho} \inf I_G(A)(1 + o(1)) \) in (6.15), so the probability range \([1, 1]\) is covered by Theorem 5.

Remark 6 \( G \) in Theorem 5 may be different from \( G \); therefore, the choice of \( G \) for the estimator is of no consequence for its consistency. The case of \( G \) and \( \hat{G} \) having support in \((0, \infty)\) as in Theorem 5 can be handled by small modifications.

Remark 7 For quantile estimator (6.3) with parameter estimator (6.7) and (6.8) for some \( \xi > 1 \), condition (6.6) is satisfied if \( c' > 1 - \xi^{-1} \) in (6.14) (de Valk (2014), Theorem 4).

In practice, computing or approximating (6.12) may not be easy. Therefore, it would be an advantage to replace \( \ell_{G,n}(B) \) in (6.12) by an arbitrary value in some suitable interval. Define for some \( \vartheta \in (0, \varrho] \)

\[
\ell_{G,n}^\vartheta(B) := \sup \{ t > 0 : \hat{p}_n(\hat{Q}_{G,n}(\hat{Z}_n/t) \in B) \geq (k_0(n)/n)^{\vartheta} \}.
\]

Then \( \ell_{G,n}^\vartheta(B) \leq \ell_{G,n}(B) \). Let \( \ell_{G,n}(B) \) be the result of an algorithm designed to satisfy

\[
\ell_{G,n}(B) \in [\ell_{G,n}^\vartheta(B), \ell_{G,n}^{\vartheta'}(B)];
\]

for the present analysis, it is sufficient to assume that \( \ell_{G,n}(B) \) is a random variable adapted to the \( \sigma \)-algebra generated by \( X^{(1)}, \ldots, X^{(n)} \) such that (6.17) holds. The following variation of the estimator (6.13) is based on \( \ell_{G,n}(B) \):

\[
\hat{P}_{G,n}(X \in B) := \left( \hat{p}_n(\hat{Q}_{G,n}(\hat{Z}_n/\ell_{G,n}(B)) \in B) \right)^{(\ell_{G,n}(B))^{-1/\varrho}}.
\]

Theorem 6 Let \( X^{(1)}, X^{(2)}, \ldots \) and \( G \) be as under Theorem 4. For \( k_0 : \mathbb{N} \to \mathbb{N} \) satisfying (6.14), consider the estimator (6.18) for \( P(X \in B) \), with the quantile estimator (6.3) satisfying (6.7) and \( \theta \in (0, (1 - c')^{-1}) \) and \( \varphi \in (0, \varrho] \). Then for \( B_{n, \tau} \) as in Theorem 4

\[
\lim_{n \to \infty} \sup_{\tau \in [T^{-1}, T]} \left| \frac{\log \hat{P}_{G,n}(X \in B_{n, \tau})}{\log P(X \in B_{n, \tau})} - 1 \right| = 0 \quad \forall T > 1.
\]

Proof The proof can be found in Subsection 8.1.
7 Discussion

Like similar methods in the classical multivariate extreme value setting (e.g. de Haan & Sinha (1999), Drees & de Haan (2013), Draisma et al (2014)), the estimators (6.13) and (6.18) exploit homogeneity of a function describing tail dependence; in this case, homogeneity of the rate function $I_G$. This offers the advantage that no explicit estimate of $I_G$ is required. However, in certain situations, there may be reasons to estimate $I_G$, such as if for a given $X$, probabilities need to be estimated for multiple sets in a consistent and reproducible manner. Therefore, estimation of $I_G$ remains a topic deserving elaboration.

The limitation of $A$ to continuity sets of $I_G$ in Theorems 5 and 6 is less restrictive than it may seem, since homogeneity of the rate function (5.5) makes continuity sets relatively common, as noted at the end of Section 3. The other conditions on $A$ are weak. To prove convergence of the estimators under such weak conditions, local uniformity in $d$ of convergence in (6.15) is employed, which is derived from uniformity in $d$ of convergence in (8.13). A variant of (8.13) involving only pointwise convergence is much easier to prove using Hoeffding’s inequality and the Borel-Cantelli lemma. Local uniformity in $\lambda$ of convergence in (8.2), also derived from (8.13), is used to prove local uniformity in $\tau$ of convergence of the estimators in (6.15) and (6.19). In practice, this means that if such an estimator applied to a given dataset produces a fair estimate of $P(X \in B_0)$ for some $B_0 \subset \mathbb{R}^m$, then it may also be applied with confidence to the same dataset to estimate the probability of $B_1 \subset \mathbb{R}^m$ such that $P(X \in B_1) \geq P(X \in B_0)^\tau$ for $\tau > 1$ not too large, e.g. $\tau = 2$. If $P(X \in B_0) \ll 1$, e.g. $P(X \in B_0) = 0.01$, this amounts to extrapolation over several additional orders of magnitude in probability.

Convergence of log-ratios of probabilities as in (6.15) and (6.19) is typical for the probability range (1.1). As observed in de Valk (2014), a stronger notion of convergence might be desirable, but would require restrictive additional assumptions which seem hard to justify. Rather, it is recommended to diagnose bias in estimates and take this into account in estimates of uncertainty. For this reason, modelling of bias and rate of convergence deserves further study. Deriving asymptotic error distributions will require additional assumptions beyond those for Theorems 5 and 6 and is left for a follow-up study as well.

The assumption of marginal GW tail limits is not restrictive, since the more generally applicable log-GW tail limit can be converted to a GW tail limit by a logarithmic transformation (de Valk (2014)). This may also work for lower tail limits, as in the case of the lognormal distribution. In other cases, one may assume that for $j \in \{1, \ldots, m\}$, numbers $c_j$ and $d_j$ in $\{0, 1\}$ exist such that the distribution function of the random variable

$$-\log \max(-d_j X_j, 1) + \min(\max(X_j, -d_j^{-1}), c_j^{-1}) + \log \max(c_j X_j, 1)$$

has upper and lower GW tails, and replace $X_j$ by (7.1). A rule for choosing $c_j$ and $d_j$ based on data of $X_j$ is discussed in de Valk (2015) under assumptions weaker than the existence (log)GW tail limits. In practice, it will often be clear how to choose $c_j$ and $d_j$.

---

4 This remark does not apply to short-range approximation as in (1.9), which is applicable to the intermediate probability range.
Potential applications of approximation based on the tail LDP’s and of estimators like include analysis of financial and economic tail risk, flood hazard analysis, and structural reliability analysis (Ditlevsen & Madsen (2007)). FORM (Hasofer & Lind (1974)), the most widely applied method for approximation of failure probabilities in structural reliability analysis, can be regarded as a large-deviations method, although it is rarely presented as such. Compared to FORM, which requires transformation of $X$ to a $U \sim N(0, I)$ (Hasofer & Lind (1974); Ditlevsen & Madsen (2007)), the tail LDP may offer simplification, as it involves only marginal transformations. In line with FORM, the standard normal distribution $\Phi$ could be chosen for $G$ in Section 5 and for $G$ in Section 6. Furthermore, approximation of failure probabilities and extreme value analysis, normally treated as two entirely separate topics in structural reliability analysis, can be combined within the same large deviation framework, as demonstrated by the estimators in Section 6.

Short-range approximation based on (4.9) may have merits in applications focused on the intermediate probability range, such as in the analysis of market risk of financial institutions; (4.9) can be applied together with classical univariate extreme value methods for the marginals.

Without much effort, the main results of this article can be generalised from a random vector in $\mathbb{R}^m$ to a random element of $C_b(K)$, the continuous functions on a compact metric space $K$. Classical multivariate extreme value theory has been generalised to this settings earlier; see e.g. de Haan & Lin (2001), Part III of de Haan & Ferreira (2006) and Ferreira & de Haan (2014). For the theory presented here, the main difference between the $\mathbb{R}^m$ setting and the $C_b(K)$ setting is that in the latter, exponential tightness of $\{\mu_{G,y}, y > 0\}$ no longer follows from the standardised marginals but is an independent assumption. In loose terms, it entails that all but an exponentially small probability mass is concentrated on equicontinuous sets of functions in $C_b(K)$ (see e.g. Dembo & Zeitouni (1998)).

8 Proofs and lemmas

8.1 Proof of Theorem 6

For convenience, the following definitions will be used, suppressing dependence on $G$ and $G$ in the notation:

$$\hat{\mu}_{n,l}(A) := \hat{p}_n(Q_{G,y}n(\hat{Z}_n/l) \in Q_{G,y}n(q_G(y_n)A)) \quad (8.1)$$

and furthermore, $l_n(A) := \ell_{G,y}n(Q_{G,y}n(q_G(y_n)A))$, $l_n^+(A) := \ell_{G,y}^+n(Q_{G,y}n(q_G(y_n)A))$ and $l_n^-(A) := \ell_{G,y}^-n(Q_{G,y}n(q_G(y_n)A))$.

Proof By Theorem 3 $\{\mu_{G,y}, y > 0\}$ defined by and satisfies the LDP with good rate function $I_G$. Let $r$ be the Weibull index of $G$. As $\theta \in (0, (1 - c')^{-1})$ with $c'$ as in 6.14, there exists a $\Delta > 0$ satisfying

$$(1 - c')^{-1} > \Delta^{1/r} \inf I_G(A) > \theta, \quad (8.2)$$
so condition (8.3) in Lemma 5 is satisfied. Fixing an arbitrary $A > 1$, then by Lemma 5 for every $\delta \in (0, \Delta)$ (see (6.11)),
\[
\lim_{n \to \infty} \sup_{\lambda \in [A^{-1}, A], d \in [\delta, \Delta]} \left| n^{-1} \log \bar{\mu}_{n, (d/\lambda)^{1/r}}(A\lambda) + d^{1/r} \inf I_G(A) \right| = 0 \quad \text{a.s.} \quad (8.3)
\]
and
\[
\limsup_{n \to \infty} \sup_{\lambda \in [A^{-1}, A], d > \Delta} \left| n^{-1} \log \bar{\mu}_{n, (d/\lambda)^{1/r}}(A\lambda) \leq -\Delta^{1/r} \inf I_G(A) < -\theta \quad \text{a.s.} \quad (8.4)
\]
Choosing $\delta < (\theta / \inf I_G(A))^{r}$, since $\Delta > (\theta / \inf I_G(A))^{r}$ by (8.2), we observe that
\[
d^{1/r} \inf I_G(A) \begin{cases} \leq \theta & \text{if } d \in [\delta, (\theta / \inf I_G(A))^{r}] \subset [\delta, \Delta] \\ > \theta & \text{if } d \in ((\theta / \inf I_G(A))^{r}, \Delta] \subset [\delta, \Delta] \end{cases}
\]
in (8.3). Therefore, with (8.3), using (6.12),
\[
\lim_{n \to \infty} \sup_{\lambda \in [A^{-1}, A]} \left| \lambda^{\theta/r} I_{K^{+}}(A\lambda) - (\theta / \inf I_G(A))^{\theta} \right| = 0 \quad \text{a.s.} \quad (8.5)
\]
and similarly, using (6.10), we find that
\[
\lim_{n \to \infty} \sup_{\lambda \in [A^{-1}, A]} \left| \lambda^{\theta/r} I_{K^{-}}(A\lambda) - (\theta / \inf I_G(A))^{\theta} \right| = 0 \quad \text{a.s.} \quad (8.6)
\]
By (8.3), (8.6), (6.17) and (8.3),
\[
\lim_{n \to \infty} \sup_{\lambda \in [A^{-1}, A]} \left| n^{-1} \log \bar{P}_{Q,n}(X \in Q_G(y_n)A\lambda) - \lambda^{1/r} \inf I_G(A) \right| = 0 \quad \text{a.s.} \quad (8.7)
\]
or equivalently, by (6.15),
\[
\lim_{n \to \infty} \sup_{\lambda \in [A^{-1}, A]} \left| n^{-1} \log \bar{P}_{Q,n}^{H}(X \in Q_G(q_G(y_n)A\lambda)) + \lambda^{1/r} \inf I_G(A) \right| = 0 \quad \text{a.s.} \quad (8.8)
\]
By Theorem 4 (6.4) holds with $\inf I_G(A^\circ) = \inf I_G(A)$, so since $\inf I_G(A) > 0$, by (8.8),
\[
\lim_{n \to \infty} \sup_{\lambda \in [A^{-1}, A]} \left| \frac{\log \bar{P}_{Q,n}^{H}(X \in Q_G(q_G(y_n)A\lambda))}{\log P(X \in Q_G(q_G(y_n)A\lambda))} - 1 \right| = 0 \quad \text{a.s.,}
\]
and (6.19) follows from (6.14) and regular variation of $q_G$, since $A > 1$ is arbitrary.  \[ \square \]

8.2 Proof of Theorem 6

Proof Following the proof of Theorem 6 in Subsection 8.1 (8.5) and (6.13) yield
\[
\lim_{n \to \infty} \sup_{\lambda \in [A^{-1}, A]} \left| n^{-1} \log \bar{P}_{Q,n}^{H}(X \in Q_G(q_G(y_n)A\lambda) + \lambda^{1/r} \inf I_G(A) \right| = 0
\]
and the result (6.15) follows as in the proof of Theorem 6.  \[ \square \]
8.3 Lemmas

**Lemma 1** (a) For a random vector $Z$ on $\mathbb{R}^m$ with all its marginals equal to $G$, a continuously increasing distribution function with support in $(0, \infty)$ and satisfying a Weibull tail limit (cf. \((7.7)\)) with index $r > 0$,

$$
\lim_{y \to \infty} y^{-1} \log P(\|Z\|_\infty / q_G(y) > \lambda) = -\lambda^{1/r} \quad \forall \lambda \geq 0.
$$

Moreover, for $Y := -\log(1 - G(Z))$, $\{Y/y; y > 0\}$ and $\{(Z/q_G(y))^{1/r}; y > 0\}$ are exponentially equivalent.

(b) If instead, $G$ is symmetric and continuously increasing on $\mathbb{R}$ and satisfies a Weibull tail limit, then \((8.4)\) holds as well.

**Proof** For every $y > 0$,

$$
1 - G(\lambda q_G(y)) \leq P(\|Z\|_\infty / q_G(y) > \lambda) \leq m(1 - G(\lambda q_G(y))) \quad \forall \lambda \geq 0. \quad (8.10)
$$

Taking logarithms and dividing by $y$, \((8.9)\) follows from the Weibull tail limit. From the definitions of $q_G$ and $Y$,

$$
\max_{j \in \{1, \ldots, m\}} |(q_G(Y_j)/q_G(y))^{1/r} - Y_j/y| > \delta. \quad (8.11)
$$

By the Weibull tail limit, $q_G \in RV(1/r)$, so for every $\Lambda > 1$ and $\delta > 0$, there exists a $y_{\Lambda, \delta} > 0$ such that \((\text{Bingham et al. (1987), Theorem 1.5.2})\),

$$
\sup_{\lambda \in (0, \Lambda]} \left| (q_G(y\lambda)/q_G(y))^{1/r} - \lambda \right| \leq \delta \quad \forall y \geq y_{\Lambda, \delta}.
$$

Therefore, \((8.11)\) for $y > y_{\Lambda, \delta}$ implies $\max_{j \in \{1, \ldots, m\}} |Y_j/y| > \Lambda$, so

$$
\lim_{y \to \infty} y^{-1} \log P \left( \|Z/q_G(y)\|_\infty^{1/r} - Y/y \right| > \delta \right) \leq \lim_{y \to \infty} \sup_{y > \Lambda} \left| \log P(\|Z\|_\infty > Ay) \right|
$$

which equals $-\Lambda$, as seen by applying \((8.9)\) with the standard exponential distribution for $G$. Since $\Lambda > 1$ is arbitrary, we obtain exponential equivalence. For (b), the upper bound in \((8.10)\) needs to be increased to $2m(1 - G(\lambda q_G(y)))$, resulting in \((8.9)\). \(\square\)

**Lemma 2** Let $G$ be a symmetric and continuously increasing distribution function satisfying a Weibull tail limit with index $r > 0$. Let $Z$ be a random vector in $\mathbb{R}^m$ with all its marginals equal to $G$ and $\{\mu_G, y; y > 0\}$ (see \((7.3)\)) satisfying the LDP \((7.4)\) with good rate function $I_G$. Let the Borel set $A \subset \mathbb{R}^m$ be a continuity set of $I_G$ satisfying $\inf I_G(A) \in (0, \infty)$. If $\{y_n > 0\}$ satisfies $\lim_{n \to \infty} y_n = \infty$ and $\Delta > 0$ satisfies

$$
\limsup_{n \to \infty} \frac{y_n}{\log n} < \frac{\Delta^{-1/r}}{\inf I_G(A)}, \quad (8.12)
$$

then for $Z^{(1)}, Z^{(2)}, \ldots$ a sequence of iid copies of $Z$, with $\hat{p}_n$ defined by \((6.11)\),

$$
\lim_{n \to \infty} \sup_{d \in (0, \Delta]} y_n^{-1} \log \hat{p}_n(Z \in dAq_G(y_n)) + d^{1/r} \inf I_G(A) = 0 \quad \text{a.s.} \quad (8.13)
$$

and

$$
\lim_{n \to \infty} \sup_{d > \Delta} y_n^{-1} \log \hat{p}_n(Z \in dAq_G(y_n)) \leq -\Delta^{1/r} \inf I_G(A) \quad \text{a.s.} \quad (8.14)
$$
Proof Let \( A := \cup_{\lambda \geq 1} \lambda A \); by (5.5), \( A \) is a continuity set of \( I_G \) satisfying \( \inf I_G(A) = \inf I_G(A) < \infty \). Define the random variable

\[
\omega := \inf \{ w > 0 : Z \in A \}
\]  

(8.15)

with \( \inf \{ \emptyset \} := \infty \), and let \( \mathcal{F} \) be its distribution function. Since \( \cup_{\lambda \geq 1} \lambda A \subset A \), \( Z \in A \) if and only if \( \inf I_G(A) < \infty \), so by (5.3) and Theorem (1.4),

\[
\lim_{y \to \infty} y^{-1} \log \mathcal{F}(w/q_G(y)) = -w^{-1/r} \inf I_G(A) \quad \forall w > 0.
\]  

(8.16)

Therefore, since \( \inf I_G(A) \in (0, \infty) \), \(- \log \mathcal{F}(1/Id) \in RV_{1/\nu} \) with \( \nu > 0 \), so by Bingham et al (1987) (Theorem 1.5.2) and (8.16) again, for every \( A \) with \( \inf A = \infty \),

\[
\lim_{y \to \infty} \sup_{w \geq a} \left| y^{-1} \log \mathcal{F}(w/q_G(y)) + w^{-1/r} \inf I_G(A) \right| = 0.
\]  

(8.17)

By (8.16), there is for every \( \varepsilon > 0 \) an \( n_\varepsilon \in \mathbb{N} \) such that for all \( n \geq n_\varepsilon \),

\[
n \mathcal{F}(a/q_G(y_n)) \geq e^{-\log n - (\varepsilon + a^{-1/r}) \inf I_G(A)/y_n}.
\]  

(8.18)

Taking \( a = 1/\Delta \), then by (8.12), \( \varepsilon > 0 \) can be chosen small enough that the exponent in (8.18) eventually exceeds \( \varepsilon \log n \). Therefore,

\[
\lim_{n \to \infty} n \mathcal{F}(a/q_G(y_n))/\log n = \infty.
\]  

(8.19)

With \( \mathcal{F}^{-1} \) the left-continuous inverse of \( \mathcal{F} \), almost surely \( \omega^{(i)} = F^{-1}(U^{(i)}) \) for all \( i \in \mathbb{N} \), with \( U^{(1)}, U^{(2)}, \ldots \) independent and uniformly distributed on \( (0, 1) \), so almost surely, \( \hat{p}_n(\omega \leq w/q_G(y_n)) = \mathcal{F}(w/q_G(y_n)) \) for all \( n \in \mathbb{N} \) and all \( w \geq a \). Therefore, by Wellner (1978) (Corollary 1) and (8.19),

\[
\lim_{n \to \infty} \sup_{w \geq a} | \log \hat{p}_n(\omega \leq w/q_G(y_n)) - \log \mathcal{F}(w/q_G(y_n)) | = 0 \quad \text{a.s.}
\]  

(8.20)

and since \( \omega \leq w/q_G(y_n) \Rightarrow Z \in A \) if and only if \( \omega \leq w/l/q_G(y_n) \) for all \( l > 1 \) and \( w > 0 \), using (8.17) and (5.5), as \( a = 1/\Delta \),

\[
\lim_{n \to \infty} \sup_{d \in (0, \Delta]} | y^{-1} \log \hat{p}_n(Z \in dA \cap \mathbb{N} ) + d^{1/r} \inf I_G(A) | = 0 \quad \text{a.s.}
\]  

(8.21)

Therefore, as \( A \subset A \) and \( \inf I_G(A) = \inf I_G(A) \),

\[
\limsup_{n \to \infty} \sup_{d \in (0, \Delta]} y^{-1} \log \hat{p}_n(Z \in dA \cap \mathbb{N} ) + d^{1/r} \inf I_G(A) \leq 0 \quad \text{a.s.}
\]  

(8.22)

\( A \) is a continuity set of \( I_G \) and \( I_G \) satisfies (5.5), so there is a point \( x \in \partial A^\circ \) and a sequence \( (x_n) \subset A^\circ \) such that \( I(x) = \inf I(A) \), \( \| x_n - x \| \downarrow 0 \) and \( I(x_n) \downarrow I(x) \) as \( n \to \infty \). For each \( i \in \mathbb{N} \), we can construct an open \( B_i \subset \mathbb{R}^m \) and \( \eta_i > 1 \) satisfying

\[
\cup_{\lambda \geq 1} \lambda B_i \subset B_i, \quad x_i \in B_i \setminus B_i \setminus \eta_i \subset A^\circ, \quad \text{and}
\]

\[
\inf I_G(B_i^\circ) = \inf I_G(B_i) \in ( \inf I(A), I(x_i) ]
\]  

(8.23)
as follows. First, we find an open cone \( C'_i \) inside \( \cup_{\lambda > 0} \lambda A^0 \) together with a \( \beta_i > 1 \) and a \( \eta_i > 1 \) such that \( x_i \in B'_i := \cup_{\lambda > \beta_i} \lambda A^0 \cap C'_i \) and \( B'_i \setminus B'_i \eta_i \subseteq A^0 \) (these exist). If \( C'_i \) is a continuity set of \( I_G \), then set \( C_i = C'_i \), and

\[
B_i := \cup_{\lambda \geq \beta_i} \lambda A^0 \cap C_i \quad (8.24)
\]

satisfies (8.23). Else, consider the continuous transformation \( f_1 : \mathbb{R}^m \times [0,1] \rightarrow \mathbb{R}^m \) defined by \( f_1(y,a) := ay + (1-a)(\|y\|_\infty / \|x_i\|_\infty)x_i \). It satisfies \( f_1(C'_i, 1) = C'_i \), \( f_1(C'_i, 0) = \cup_{\lambda > 0} \lambda x_i \) and \( f_1(C'_i, a) \subset f_1(C'_i, a') \) if \( a \leq a' \). Therefore, \( a \rightarrow \inf I_G(\cup_{\lambda \geq \beta_i} \lambda A^0 \cap f_1(C'_i, a)) \) is nonincreasing, so with \( a_i \) any of its continuity points in \((0,1)\) and \( C_i = f_1(C'_i, a_i) \). (8.24) is a continuity set of \( I_G \) and satisfies (8.23). By (8.23),

\[
\hat{p}_n(Z \in dA_{QG}(y_n)) \geq \hat{p}_n(Z \in dB_{QG}(y_n)) - \hat{p}_n(Z \in d\eta_{B,QG}(y_n))
\]

and furthermore, (8.24) continues to hold after substituting \( A \) by \( B_i \). Therefore, by (8.23), almost surely, the right-hand side of (8.25) is \( \hat{p}_n(Z \in dB_{QG}(y_n))(1 + o(1)) \) uniformly in \( d \in (0,\Delta) \) and furthermore, using (8.23),

\[
\liminf_{n \to \infty} \inf_{d \in (0,\Delta)} \frac{y_n^{-1} \log \hat{p}_n(Z \in dA_{QG}(y_n))}{d^{1/r} I_G(x_i)} \geq 0 \quad \text{a.s.} \quad (8.26)
\]

Now (8.14) follows from (8.23) and (8.26), because \( \lim_{n \to \infty} I_G(x_i) = \inf I_G(A) \). Finally, by (8.21), as \( \cup_{\lambda \geq 1} A_\lambda \subseteq A \),

\[
\limsup_{n \to \infty} \sup_{d > \Delta} \frac{y_n^{-1} \log \hat{p}_n(Z \in dA_{QG}(y_n))}{d^{1/r} I_G(A)} \leq -\Delta^{1/r} \inf I_G(A) \quad \text{a.s.} \quad (8.27)
\]

and because \( A \subseteq A \) and \( \inf I_G(A) = \inf I_G(A) \), (8.11) follows. \( \Box \)

**Lemma 3** Let \( Z \) be a random vector on \( \mathbb{R}^m \) with all marginals equal to \( G \), a symmetric and continuously increasing distribution function satisfying a Weibull tail limit with index \( r > 0 \). For \( Z(1), Z(2), \ldots \) a sequence of iid copies of \( Z \), let \( \tilde{Z}^{(i)} := (\tilde{Z}^{(i)}_{1,n}, \ldots, \tilde{Z}^{(i)}_{m,n}) \) for \( i = 1, \ldots, n \) with

\[
\tilde{Z}^{(i)}_{j,n} := G^{-1}(\langle R^{(i)}_{j,n} - \frac{1}{n} \rangle).
\]

and \( R^{(i)}_{j,n} := \sum_{i=1}^{n} 1(Z(i) \leq Z^{(i)}_{j}) \). For \( (y_n > 0) \) satisfying \( \liminf_{n \to \infty} y_n / \log n > 0 \),

\[
\sup_{\varepsilon > 0} \limsup_{n \to \infty} y_n^{-1} \log \hat{p}_n \left( \| \tilde{Z} - Z \|_\infty > q_G(y_n)\varepsilon \right) < -\liminf_{n \to \infty} y_n^{-1} \log n \quad \text{a.s.} \quad (8.29)
\]

**Proof** Fix an arbitrary \( \varepsilon > 0 \). Since

\[
\hat{p}_n \left( \| \tilde{Z} - Z \|_\infty > q_G(y_n)\varepsilon \right) \leq \sum_{j=1}^{m} \hat{p}_n \left( \| \tilde{Z}_{j,n} - Z_j \| > q_G(y_n)\varepsilon \right).
\]

it is sufficient to prove (8.23) for the univariate case. Hence, representing \( Z = G^{-1}(U) \) with \( U \) uniformly distributed on \((0,1)\), we can write \( Z = Z^+ + Z^- \) with

\[
Z^+ := \max(Z,0) = q_G(\log \min(1-U,\frac{1}{n}));
\]

\[
Z^- := \min(Z,0) = -q_G(\log \min(U,\frac{1}{n})); \quad (8.30)
\]
and similarly, \( \hat{Z}_n = \hat{Z}_n^+ + \hat{Z}_n^- \) with
\[
\hat{Z}_n^+ = q_G(-\log \min(1 - \frac{n-1}{n}, \frac{i}{n})); \quad \hat{Z}_n^- = -q_G(-\log \min(\frac{n-1}{n}, \frac{i}{n})). \tag{8.31}
\]

By [Wellner 1978] (Corollary 4), the order statistics \( U_{1:n} \leq \ldots \leq U_{n:n} \) derived from \( U^{(1)}, \ldots, U^{(n)} \) satisfy
\[
\lim_{n \to \infty} \sup_{i \in \{\lfloor \log n \rfloor, \ldots, n\}} \left| -\log(U_{i:n}) + \log(n^{-1}(i - \frac{1}{2}), \frac{1}{2}) \right| = 0 \quad \text{a.s.},
\]
so as \( \inf_{n \to \infty} y_n / \log n > 0 \) and \( q_G \in RV(r) \) with \( r > 0 \) (see Bingham et al. 1987, Theorem 1.5.2), using (8.30) and (8.31),
\[
\lim_{n \to \infty} \sup_{i \in \{\lfloor \log n \rfloor, \ldots, n\}} \left| Z_{i:n}^- - \hat{Z}_{i:n}^- \right| / q_G(y_n) = 0 \quad \text{a.s.}
\]
and therefore,
\[
\lim_{n \to \infty} \sup_{i \in \{\lfloor \log n \rfloor, \ldots, n\}} \left| Z_i - \hat{Z}_n \right| / q_G(y_n) = 0 \quad \forall \varepsilon > 0 \quad \text{a.s.} \tag{8.32}
\]

Combined with the analogous result for \( Z^+ - \hat{Z}_n^+ \), we obtain
\[
\lim_{n \to \infty} \sup_{i \geq \delta} \log n^{-1} \hat{p}_n \left( Z_i - \hat{Z}_n > q_G(y_n) \varepsilon \right) \leq 1 \quad \forall \varepsilon > 0 \quad \text{a.s.} \tag{8.33}
\]

which implies the univariate case of (8.29) to be proven. \( \square \)

**Lemma 4** Let the random vector \( X \) on \( \mathbb{R}^n \) have continuous marginals satisfying upper and lower GW tail limits with indices \( \rho \) and \( \tau > 0 \), respectively. Let \( X^{(1)}, X^{(2)}, \ldots \) be a sequence of iid copies of \( X \). With \( Q_G \) and \( \hat{Z}_n \) defined by (6.3) and (6.10), let \( \hat{G} := G^{-1}(\hat{Z}_n) \), and \( Q_{G,n} \) defined by (6.3), let \( k_0 : \mathbb{N} \to \mathbb{N} \) satisfy (6.14) and let \( \hat{q}_{n,X} \) given by (6.14) satisfy (6.4). Then for every \( \delta > 0 \) and \( \varepsilon > 0 \),
\[
\lim_{n \to \infty} \sup_{i \geq \delta} \log n^{-1} \hat{p}_n \left( \left\| Q_{G}^{-1} Q_{G,n}(\hat{Z}_n^+ \varepsilon^{1/\rho}) - \hat{Z}_n^+ \right\|_\infty > q_G(y_n) \varepsilon \right) = -\infty \quad \text{a.s.} \tag{8.34}
\]

**Proof** By the Weibull tail limits of \( G \) and \( q_G \in RV(r) \) and \( q_G \in RV(\rho) \). Fix \( \varepsilon > 0 \) and \( \delta > 0 \). As in Lemma 3, we only need to prove (8.33) for the univariate case, so we proceed with this. As either \( \hat{Z}_n = \hat{Z}_n^+ := \max(\hat{Z}_n, 0) \) and \( \hat{Z}_n = \hat{Z}_n^- := \max(\hat{Z}_n, 0) \) or \( \hat{Z}_n = \hat{Z}_n^- := \min(\hat{Z}_n, 0) \) and \( \hat{Z}_n = \hat{Z}_n^+ := \min(\hat{Z}_n, 0) \),
\[
\hat{p}_n \left( \left\| Q_{G}^{-1} Q_{G,n}(\hat{Z}_n^+ \varepsilon^{1/\rho}) - \hat{Z}_n^+ \right\|_\infty > q_G(y_n) \varepsilon \right) \\
\leq \hat{p}_n \left( \left\| Q_{G}^{-1} Q_{G,n}(\hat{Z}_n^+ \varepsilon^{1/\rho}) - \hat{Z}_n^+ \right\|_\infty > q_G(y_n) \varepsilon \right) \\
+ \hat{p}_n \left( \left\| Q_{G}^{-1} Q_{G,n}(\hat{Z}_n^- \varepsilon^{1/\rho}) - \hat{Z}_n^- \right\|_\infty > q_G(y_n) \varepsilon \right), \tag{8.34}
\]
so by symmetry, we only need to show that almost surely,
\[
\lim_{n \to \infty} \sup_{i \geq \delta} \log n^{-1} \hat{p}_n \left( \left\| Q_{G}^{-1} Q_{G,n}(\hat{Z}_n^- \varepsilon^{1/\rho}) - \hat{Z}_n^- \right\|_\infty > q_G(y_n) \varepsilon \right) = -\infty. \tag{8.35}
\]
Noting that $\hat{Z}^+_n = q_G q^{-1}_G (\hat{Z}^+_n)$, if
\[
\sup_{l \geq \delta} \sup_{j \in \{1, \ldots, n\}} |q_G^{-1} q_G (\hat{Z}^+_n) - l^{-1} q_G q^{-1}_G (\hat{Z}^+_n)| \leq q_G (y_n) \varepsilon \quad \forall n \geq n_{\delta, \varepsilon}.
\]
(8.36)

Fixing $A > \max(1, \delta^{-1/r})/(1 - c')$ with $c'$ as in (6.11), there exists $N_{\delta} \in \mathbb{N}$ such that for all $n \geq N_{\delta}$, $l \geq \delta$ and all $j \in \{1, \ldots, n\}$,
\[
q_G (\log 2) = 0 = \hat{Z}^+_n \min(1, l^{-\varepsilon/r})
\]
\[
\leq \hat{Z}^+_n \max(1, l^{-\varepsilon/r}) \leq q_G (\log 2) \max(1, \delta^{-\varepsilon/r}) \leq q_G (y_n A)
\]
(8.37)
because $q_G \in RV_{(\phi)}$. In the univariate case, $Q^{-1} G_n = q_G \circ q^{-1}_G \circ \hat{q}_{n,X} \circ q^{-1}_G$ on $[0, \infty)$ with $q_F$ defined by (6.2) and $\hat{q}_{n,X}$ by (6.3). Therefore, by (8.37), (8.39) holds if both
\[
\sup_{y \in [\log 2, y_n, A]} |q_G (q^{-1}_G \hat{q}_{n,X} (y)) - q_G (y)| \leq q_G (y_n) \varepsilon/2 \quad \forall n \geq n_{\delta, \varepsilon}
\]
(8.38)
and
\[
\sup_{y \in [\log 2, y_n, A]} \sup_{l \geq \delta} |q_G q^{-1}_G (l^{-\varepsilon/r} q_G (y)) - q_G (y) l^{-1} | \leq q_G (y_n) \varepsilon/2 \quad \forall n \geq n_{\delta, \varepsilon}
\]
(8.39)
hold. Since $q_G \in RV_{(\phi)}$ and $q_G \in RV_{(\phi)}$ with $\phi > 0$ and $r > 0$, (8.38) and (8.39) are both true for some $n_{\delta, \varepsilon}$ if $\hat{q}_{n,X}$ defined by (6.5) satisfies
\[
\lim_{n \to \infty} \sup_{y \in J_n} |\hat{q}_{n,X} (y)| = 0
\]
(8.40)
for $J_n = [y_n, y_n A]$ and for $J_n = [\log 2, y_n]$. By (6.6), (8.40) holds almost surely for $J_n = [y_n, y_n A]$. Furthermore, for all $y \in (0, y_n)$, $y \hat{q}_{n,X} (y) = q_F (X_{n(1 - \nu(y) + 1)}) - y$, which equals $-\log U_{n(1 - \nu(y) + 1)} - y$ for all $y > 0$ and $n \in \mathbb{N}$ for $U^{(i)} := 1 - F (X^{(i)})$ uniformly distributed on $(0, 1)$ for all $i \in \mathbb{N}$. Therefore, by (6.8), Corollary 4 in [Wellner (1978)] implies that $\limsup_{n \to \infty} \sup_{y \in [0, y_n]} |\hat{y}_{n,X} (y)| = 0$ with probability 1, and (8.40) has probability 1 for $J_n = [\log 2, y_n]$ as well. This completes the proof.

\[\square\]

**Lemma 5** Let $G$ and $G$ be symmetric, continuously increasing distribution functions satisfying Weibull tail limits with indices $\phi > 0$ and $r > 0$, respectively. Let the random vector $X$ on $\mathbb{R}^n$ satisfy upper and lower marginal GW tail limits as in Theorem 4, and let $\{\mu_G, y > 0\}$ (see (6.3)) satisfy the LDP (5.3) with good rate function $I_G$. Let the Borel set $A \subset \mathbb{R}^n$ be a continuity set of $I_G$ satisfying inf $I_G (A) \in (0, \infty)$. Let $k_0 : \mathbb{N} \to \mathbb{N}$ satisfy (6.14) and let the quantile estimator $\hat{q}_{n,X}$ given by (6.3) satisfy (6.8). For $c'$ defined in (6.11), let
\[
\Delta \in \left(0, ((1 - c') \inf I_G (A))^{-r}\right).
\]
(8.41)

Then for a sequence $X^{(1)}, X^{(2)}, \ldots$ of iid copies of $X$, $\hat{\mu}$ defined by (8.11) satisfies for every $A > 1$ and $\delta \in (0, \Delta)$,
\[
\lim_{n \to \infty} \sup_{\lambda \in A^{-1} A, d \in [0, \Delta]} \left| y_n^{-1} \log \hat{\mu}_{G, n, (d/\lambda)^{\phi/r}} (\lambda A) + d^{1/r} \inf I_G (A) \right| = 0 \quad \text{a.s.}
\]
(8.42)
Lemma 2 implies for $I$ and it follows that as in (6.14),

$$
\lim sup_{n \to \infty} \sup_{\lambda \in [A^{-1},A], d > \Delta} \frac{y_n^{-1} \log \hat{\mu}_{G,n}(d/\lambda)^{\nu/r}(\lambda A) - \Delta^{1/r}}{\inf I_G(\lambda) \text{ a.s.}} \leq -\frac{1}{\Delta^{1/r}} \inf I_G(A) \text{ a.s.} \quad (8.43)
$$

**Proof.** With $\{\text{prop}\}$ denoting the subset of the underlying probability space satisfying the proposition $\text{prop}$, consider the events

$$
C_{a,b,n} := \{ \sup_{i \geq n} \left| Q_G^{-1} \hat{Q}_{G,n} \left( Z_{\Delta}(i)^{i-\epsilon} - Z_{\Delta}(i)^{-1} \right) \right| \leq q_G(y_n)b \}. \quad (8.44)
$$

for $i \in \{1, \ldots, n\}$. Similar to (6.11), we can define empirical probabilities of such events, e.g., $\hat{p}(C_{a,b,n}) := n^{-1} \sum_{i=1}^{n} 1(C_{a,b,n})$. Combining Lemmas 2 and 4 for $c'$ as in (6.14),

$$
\lim sup_{n \to \infty} y_n^{-1} \log \hat{p}(C_{a,b,n}) < -(1-c')^{-1} \text{ a.s.} \quad \forall a, b > 0. \quad (8.45)
$$

For every $S \subset \mathbb{R}$ and $\epsilon > 0$, let $S' := \{ x \in \mathbb{R} : \inf_{x' \in S} \| x - x' \| \leq \epsilon \} \text{ (closed)}$, and $S := \{ x \in \mathbb{R} : \inf_{x' \in S} \| x - x' \| > \epsilon \} \text{ (open)}$. Set $S_0 := S$. Since $I_G$ is a good rate function, Lemma 4.1.6 of [Dembo & Zeitouni (1998)] implies

$$
\liminf_{i \downarrow 0} I_G(A^i) = \inf_{i \downarrow 0} I_G(A^i) = \inf_{i \downarrow 0} I_G(\cup_{i > 0} A^{-i}) = \liminf_{i \downarrow 0} I_G(A^{-i}). \quad (8.46)
$$

so the nonincreasing function $i \mapsto \inf I_G(A^i)$ is continuous in $(-i_0, i_0)$ for some $i_0 > 0$, and it follows that $A^i$ is a continuity set of $I_G$ for every $i \in (-i_0, i_0)$. Therefore, for $E_{d,i,n}^{(i)} := \{ Z^{(i)} \in d \hat{q}_G(y_n) A^i \}$,

$$
\lim_{n \to \infty} \sup_{d \in (0, \Delta]} \left| y_n^{-1} \log \hat{p}(E_{d,i,n}) + d^{1/r} \inf I_G(A^i) \right| = 0 \quad \forall \epsilon \in (-i_0, i_0). \quad (8.48)
$$

By (8.46) and (8.47), there exist $\epsilon > 0$ and $i_1 \in (0, i_0)$ such that $\inf I_G(A^{-i})d^{1/r} \leq \inf I_G(A^i) \Delta^{1/r} + \epsilon \leq (1-c')^{-1} - \epsilon$ for all $i \in [0, i_1]$ and $d \in (0, \Delta]$. Therefore, by (8.48),

$$
\lim_{n \to \infty} \inf_{d \in (0, \Delta]} \left| y_n^{-1} \log \hat{p}(E_{d,-i,n}) \right| \geq -(1-c')^{-1} + \epsilon \quad \forall \epsilon \in [0, i_1] \quad \text{a.s.} \quad (8.49)
$$

Let

$$
D_{d,A,i,n}^{(i)} := \{ \hat{Q}_{G,n}(Z^{(i)}(\lambda/d)^{\nu/r}) \in \hat{q}_G(y_n) A^i \}. \quad (8.50)
$$

By (8.47), (8.48) and (8.50), for all $d \geq \delta, \lambda \in [A^{-1}, A]$ and $\epsilon > 0$, $E_{d,-i,n}^{(i)} \cap C_{d/\lambda, i, A, n}^{(i)} \subset D_{d,A,i,n}^{(i)}$, so $E_{d,-i,n}^{(i)} \cap C_{d/\lambda, i, A, n}^{(i)} \subset D_{d,A,i,n}^{(i)}$ and therefore, $\hat{p}(D_{d,A,i,n}^{(i)}) \geq \hat{p}(E_{d,-i,n}) - \hat{p}(C_{d/\lambda, i, A, n}^{(i)})$. Therefore,

$$
\inf_{d \in [\delta, \Delta], \lambda \in [A^{-1}, A]} y_n^{-1} \log \hat{p}(D_{\lambda,A,i,n}^{(i)}) - y_n^{-1} \log \hat{p}(E_{d,-i,n}) \geq y_n^{-1} \log (1 - e^{\log \hat{p}(C_{d/\lambda, i, A, n}^{(i)}) - \inf_{d \in [\delta, \Delta]} \log \hat{p}(E_{d,-i,n})}),
$$
so by (8.45) and (8.49), for all $t \in (0, t_1]$,
\[
\lim_{n \to \infty} \inf_{d \in [\delta, \Delta], \lambda \in [A^{-1}, A]} \frac{1}{y_n} \log \hat{p}_n(D_{\lambda,d,n}) - \frac{1}{y_n} \log \hat{p}_n(E_{d,-i,n}) \geq 0 \quad \text{a.s.} \tag{8.51}
\]
and therefore, by (8.45) and (8.46),
\[
\lim_{n \to \infty} \inf_{d \in [\delta, \Delta], \lambda \in [A^{-1}, A]} \frac{1}{y_n} \log \hat{p}_n(D_{\lambda,d,n}) + d^{1/r} \inf I_G(A) \geq 0 \quad \text{a.s.} \tag{8.52}
\]
Similarly, noting that $D^{(i)}_{\lambda,d,n} \cap C^{(i)}_{\delta/A,\lambda/n} \subset E^{(i)}_{d,-i,n}$ for all $i > 0$, $d \geq \delta$ and $\lambda \in [A^{-1}, A]$,
\[
\inf_{d \in [\delta, \Delta], \lambda \in [A^{-1}, A]} \frac{1}{y_n} \log \hat{p}_n(E_{d,-i,n}) - \frac{1}{y_n} \log \hat{p}_n(D_{\lambda,d,n}) \geq \frac{1}{y_n} \log (1 - d^{1/r} \inf \hat{p}_n(C^{(i)}_{\delta/A,\lambda/n} - \inf_{d \in [\delta, \Delta], \lambda \in [A^{-1}, A]} \log \hat{p}_n(D_{\lambda,d,n}))
\]
which by (8.45), (8.52) and (8.41) implies
\[
\lim_{n \to \infty} \inf_{d \in [\delta, \Delta], \lambda \in [A^{-1}, A]} \frac{1}{y_n} \log \hat{p}_n(D_{\lambda,d,n}) - \frac{1}{y_n} \log \hat{p}_n(D_{\lambda,d,n}) \geq 0 \quad \forall t > 0 \quad \text{a.s.}
\]
Therefore, by (8.45) and (8.46),
\[
\lim_{n \to \infty} \sup_{d \in [\delta, \Delta], \lambda \in [A^{-1}, A]} \frac{1}{y_n} \log \hat{p}_n(D_{\lambda,d,n}) + d^{1/r} \inf I_G(A) \leq 0 \quad \text{a.s.,} \tag{8.53}
\]
so with (8.52) and (8.51), (8.54) is obtained. By Lemma 2,
\[
\lim_{n \to \infty} \sup_{d > \Delta} \frac{1}{y_n} \log \hat{p}_n(E_{d,-i,n}) \leq -\Delta^{1/r} \inf I_G(A') \quad \text{a.s.} \quad \forall t \in (-\varepsilon_0, \varepsilon_0).
\tag{8.54}
\]
For all $d \geq \delta$, $\lambda \in [A^{-1}, A]$ and $i > 0$, $D^{(i)}_{\lambda,d,n} \subset D^{(i)}_{\lambda,d,n} \cup C^{(i)}_{\delta/A,\lambda/n}$, so $\hat{p}_n(D^{(i)}_{\lambda,d,n}) \leq 2 \max(\hat{p}_n((C^{(i)}_{\delta/A,\lambda/n})^c), \hat{p}_n(E^{(i)}_{d,-i,n}))$. Therefore, by (8.45), (8.51), (8.50) and (8.41), (8.53) is obtained. □

Acknowledgements The author would like to thank John Einmahl for his helpful comments and suggestions.

References

Bingham, N.H., C.M. Goldie, J.L. Teugels, Regular variation. Cambridge Univ. Press, (1987)
Bruun, J. T. and J.A. Tawn, “Comparison of approaches for estimating the probability of coastal flooding”. J. Roy. Statist. Soc. Ser. C (Applied Statistics) 47, 405-423, (1998)
Coles, S.G. and J.A. Tawn, “Statistical Methods for Multivariate Extremes: An Application to Structural Design”. J. Roy. Statist. Soc. Ser. C (Applied Statistics) 43 (1), 1-48, (1994)
Dembo, A. and O. Zeitouni, Large deviations techniques and applications. Springer, New York, (1998)
Ditlevsen, O. and H. O. Madsen, *Structural Reliability Methods*. Technical University of Denmark, (2007)

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

Drees, H. and L. de Haan, “Estimating Failure Probabilities”. *To appear in Bernoulli*, (2013)

Embrechts, P. and G. Puccetti, “Aggregating risk across matrix structured loss data: the case of operational risk”. *Journal of Operational Risk* 3(2), 29-44, (2007)

Ferreira, A. and L. de Haan, “The generalised Pareto process; with a view towards application and simulation”. *Bernoulli* 20(4), 171701737, (2014)

de Haan, L. and A.K. Sinha, “Estimating the probability of a rare event”. *Annals of Statistics* 27 (2), 732-759, (1999)

de Haan, L. and T. Lin, “On convergence toward an extreme value distribution in c(0, 1]”. *Ann. Probab.*, 29(1):467–483, (2001)

de Haan, L. and A. Ferreira, *Extreme value theory - An introduction*. Springer, (2006)

Hasofer, A.M. and N.C. Lind, “Exact and invariant second moment code format”. *J. Eng. Mech. Div.* 100, 111-121, (1974)

ISO, “Petroleum and natural gas industries - Specific requirements for offshore structures Part 1: Metocean design and operating considerations”. *ISO/FDIS 19901-1:2005(E)*, (2005)

Klüppelberg, C., “On the asymptotic normality of parameter estimates for heavy Weibull-like tails”. *Preprint*, (1991)

Ledford, A.W. and J.A. Tawn, “Statistics for near independence in multivariate extreme values”. *Biometrika* 83 (1): 169-187, (1996)

Ledford, A.W. and J.A. Tawn, “Modelling dependence within joint tail regions”. *J. Royal Statist. Soc. Ser. B* 59: 475-499, (1997)

Ledford, A.W. and J.A. Tawn, “Concomitant tail behaviour for extremes”. *Adv. Appl. Prob.* 30: 179-215, (1998)

Resnick, S. I., *Extreme values, regular variation, and point processes*. Springer Verlag, (1987)

de Valk, C.F., “Approximation of high quantiles from intermediate quantiles”. *Submitted (see http://arxiv.org/abs/1307.6591)*, (2014)

de Valk, C.F., “Tail approximation without tail limits, with an application to high quantile estimation” (*manuscript in preparation*), (2015)

Vrouwenvelder, A.W.C.M. and P. Struik, “Safety philosophy for dike design in the Netherlands”. *Proceedings of the 22nd Conference on Coastal Engineering*, ASCE, (1990)

Wadsworth, J.L. and J.A. Tawn, “A new representation for multivariate tail probabilities”. *Bernoulli* 19(5B), 2689-2714, (2013)

Wellner, J.A., “Limit theorems for the ratio of the empirical distribution function to the true distribution function”. *Z. Wahrscheinlichkeitsrechnung verw. Gebiete* 45, 73-88, (1978)