Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

Sub-asymptotic motivation for new conditional multivariate extremes models

Thomas Lugrin[1] | Jonathan A. Tawn[2] | Anthony C. Davison[3]

[1]Department of Defence, Swiss Federal Administration, Bern, Switzerland
[2]Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
[3] Institute of Mathematics, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Correspondence
A. C. Davison, EPFL-FSB-MATH-STAT, Station 8, 1015 Lausanne, Switzerland
Email: Anthony.Davison@epfl.ch

Summary

Statistical models for extreme values are generally derived from non-degenerate probabilistic limits that can be used to approximate the distribution of events that exceed a selected high threshold. If convergence to the limit distribution is slow, then the approximation may describe observed extremes poorly, and bias can only be reduced by choosing a very high threshold, at the cost of unacceptably large variance in any subsequent tail inference. An alternative is to use sub-asymptotic extremal models, which introduce more parameters but can provide better fits for lower thresholds. We consider this problem in the context of the Heffernan–Tawn conditional tail model for multivariate extremes, which has found wide use due to its flexible handling of dependence in high-dimensional applications. Recent extensions of this model appear to improve joint tail inference. We seek a sub-asymptotic justification for why these extensions work, and show that they can improve convergence rates by an order of magnitude for certain copulas. We also propose a class of extensions of them that may have wider value for statistical inference in multivariate extremes.

KEYWORDS:
asymptotic dependence, asymptotic independence, conditional extremes, Gaussian distribution, logistic model, sub-asymptotic approximation

1 | INTRODUCTION

Catastrophic events can have a major impact on physical infrastructure and on society. Multivariate extreme value models are used to capture the structure of such events, for a single hazard at multiple sites, for multiple hazards at a single site or multiple hazards at multiple sites, and are used to extrapolate measures of risk beyond the available data. Applications include river flooding (Asadi, Davison, & Engelke 2015; Katz, Parlange, & Naveau 2002; Keef, Svensson, & Tawn 2009; Keef, Tawn, & Svensson 2009), extreme rainfall (Coles & Tawn 1996; Huser & Davison 2014; Süveges & Davison 2012), wave height and extreme sea surge (de Haan & de Ronde 1998) and high concentrations of air pollutants (Heffernan & Tawn 2004). Such methods also provide improved assessments of financial risk (Hilal, Poon, & Tawn 2011 2014; Poon, Rockinger, & Tawn 2003).

The first limit theorems for multivariate extremes were for componentwise maxima of independent and identically distributed random vectors $\mathbf{X}_1, \ldots, \mathbf{X}_n$ which, when suitably normalised, have a non-degenerate limiting distribution as $n \to \infty$, where that limit distribution is a member of the class of multivariate extreme value distributions (de Haan & Resnick 1977; Pickands 1981; Sibuya 1960; Tiago de Oliveira 1962/63). However, these limit distributions are independent for a broad range of distributions, including all non-perfectly dependent Gaussian distributions and copulas identified by Heffernan (2000), suggesting that these limits may provide poor approximations for componentwise maxima when n is finite, as then the maxima may still be strongly dependent. The class of distributions with this property is elucidated by one of the standard

measures of extremal dependence (Coles, Heffernan, & Tawn 1999; Joe 1993),

$$\chi = \lim_{p \to 1} \Pr\{F_2(X_2) > p \mid F_1(X_1) > p\}, \tag{1}$$

where $X_i \sim F_i$ for $i = 1, 2$. The values $\chi > 0$ and $\chi = 0$ are respectively termed asymptotic dependence and asymptotic independence; with a dependent/independent bivariate extreme value distribution, the largest values of X_1 and X_2 can/cannot occur together. A related quantity, $\bar{\chi}$, lies in the interval $[-1, 1]$ and is used to distinguish among different degrees of asymptotic independence (Coles et al. 1999). Owing to the breadth of the class of asymptotically independent distributions, there have been numerous studies of sub-asymptotic properties of bivariate maxima. For example, Bofinger and Bofinger (1965) and Bofinger (1970) derived the correlation of componentwise maxima for bivariate Gaussian and certain other copulas for $n \leq 50$; more recent examples are Beranger, Padoan, and Sisson (2017) and Beranger, Padoan, Xu, and Sisson (2019).

There have been parallel developments for multivariate threshold methods. The underpinning limit theory is based on point processes and multivariate regular variation (Coles & Tawn 1991; de Haan 1985; Resnick 2007). Some work focuses on how second-order features influence estimators (Cai, Einmahl, & de Haan 2011), whereas other approaches reframe the problem by converting second-order features into the primary term in the limit theory. Ledford and Tawn (1997) take X_1 and X_2 to have unit Fréchet marginal distributions and consider $\lim_{t\to\infty} t^{1/\eta}\Pr(X_1 > xt, X_2 > yt)$ for fixed $x, y > 0$ and some constant η, with $0 < \eta \leq 1$, yielding a finite limit which gives a first-order limit model that smoothly encompasses perfect dependence, asymptotic dependence, asymptotic independence and complete independence. This limit characterisation was later extended to $\lim_{t\to\infty} t^{\lambda(\gamma)}\Pr(X_1 > xt^{\gamma}, X_2 > yt^{1-\gamma})$ where $\lambda(\gamma)$ is a positive function of γ ($0 \leq \gamma \leq 1$) satisfying a range of conditions described in Wadsworth and Tawn (2013) and de Valk (2016). These results, and other related asymptotically motivated models (Huser & Wadsworth 2019; Wadsworth, Tawn, Davison, & Elton 2017), encompass both asymptotic dependence and asymptotic independence, but they only consider growth rates in the arguments of the joint survivor function, on Fréchet marginals, that are linked through a power; on exponential margins these growth rates are proportional. Furthermore, current results in these cases are for low-dimensional cases only.

In this paper we focus on the conditional extremal model of Heffernan and Tawn (2004), which places no preconditions on the relative growth of the large variables and has been widely used for substantive applications owing to its ability to handle a wide range of joint tail dependencies, its parsimony, its simple computational properties, and its applicability to high dimensions (Tawn, Shooter, Towe, & Lamb 2018). To simplify the notation we deal with the bivariate case, but extension to the general multivariate case, of both existing methods and our developments, is straightforward. This model was originally presented for marginally Gumbel distributed random vectors, but Keef, Papastathopoulos, and Tawn (2013) showed that formulation on the Laplace scale is preferable when positive or negative dependence is possible, so we first transform (X_1, X_2) to random variables (X, Y) with Laplace margins via the probability integral transform

$$X = \operatorname{sign}\{1 - 2F_1(X_1)\} \log\left[1 - \{|1 - 2F_1(X_1)|\}\right],$$

and similarly for Y, preserving the dependence structure through the copula, according to Sklar's (1959) representation theorem.

Under conditions specified by Heffernan and Tawn (2004), which include the joint distribution of (X, Y) being in the standard domain of attraction of the bivariate extreme value distribution, the conditional extremal model presupposes the existence of normalising functions $a(\cdot) : \mathbb{R}_+ \to \mathbb{R}$ and $b(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ such that for $x > 0$,

$$\Pr\left\{Z_u := \frac{Y - a(X)}{b(X)} \leq z,\ X - u > x \mid X > u\right\} \to H(z)\exp(-x), \quad u \to \infty, \tag{2}$$

where $H(\cdot)$ is a non-degenerate distribution function with no mass at infinity. Under mild assumptions on the distribution of (X, Y), results in Heffernan and Resnick (2007) imply that

$$a(x) = x\mathcal{L}_a(x), \quad b(x) = x^{\beta}\mathcal{L}_b(x),\ \beta < 1, \tag{3}$$

with the functions $\mathcal{L}_a(x)$ and $\mathcal{L}_b(x)$ slowly varying: $\mathcal{L}(xt)/\mathcal{L}(x) \to 1$ for any fixed $t > 0$ as $x \to \infty$, where $\mathcal{L}$ is either $\mathcal{L}_a$ or $\mathcal{L}_b$. Two aspects of the limit distribution should be noted. Firstly, the Laplace margins imply that the exponential limit for $X - u$ is exact for any positive u. Secondly, (2) corresponds to Z_u and X becoming independent as $u \to \infty$, so the finite-u distribution H_u of Z_u depends less and less on X as the limit is approached, i.e., $H_u \to H$.

In order to construct a statistical model, Heffernan and Tawn (2004) and Keef et al. (2013) assume that the limit on the right-hand side of (2) holds exactly above some finite u, i.e., $H_u = H$, and they adopt parametric families for $a(\cdot)$ and $b(\cdot)$ that satisfy (3), yield a parsimonious model and encompass a broad range of asymptotic dependence and asymptotic independence structures. By considering the forms of $a(\cdot)$ and $b(\cdot)$ in a broad class of copulas, they propose taking canonical parametric forms for a and b, i.e.,

$$a_0(x) = \alpha x, \quad \alpha \in [-1, 1], \qquad\qquad b_0(x) = x^{\beta}, \quad \beta \in (-\infty, 1), \tag{4}$$

which include all the normings they found and correspond to approximating the slowly-varying functions by $\mathcal{L}_a(x) \equiv \alpha$ and $\mathcal{L}_b(x) \equiv 1$ in expression (3). The latter is equivalent to setting $\mathcal{L}_b(x) = b$ for any constant $b > 0$, with the change in norming absorbed into the variance of $H(\cdot)$. If $\alpha = 1$ and $\beta = 0$, then (X, Y) are asymptotically dependent with $\chi = \int_0^{\infty} \bar{H}(-z)e^{-z}dz$, and otherwise they are asymptotically independent with $\chi = 0$ and the value of $\bar{\chi}$ dependent on the upper tail form of H.

Evidence is emerging that the canonical norming functions (4) are not optimal for all theoretical copulas or in statistical practice. Papastathopoulos and Tawn (2016) found examples of the inverted multivariate extreme value copula (i.e., the lower joint tail of the multivariate extreme value copula) for which more general forms of $\mathcal{L}_a(x)$ and $\mathcal{L}_b(x)$ of the form (3) are required. Tendijck, Eastoe, Tawn, Randell, and Jonathan (2020) and Simpson, Opitz, and Wadsworth (2020) have also found improved fits using $a_0(x) = \alpha_0 + \alpha x$ for some constant α_0.

In this paper we study possible theoretical justification for this improved performance by exploring the sub-asymptotic behaviour of the conditional multivariate extreme value limit (2) for some well-studied copulas. We quantify the relative benefits of different characterisations in (3) by determining their respective rates of convergence in (2). We also explore whether relaxing the limiting independence assumption for X and Z_u can further improve rates of convergence. Another motivation for our study is that when simulating data on which to assess the performance of methods to fit the conditional model (e.g., Lugrin 2018), the estimates of $a_0(x)$ and $b_0(x)$ for $x > u$ can misleadingly suggest a poor fit, as it is $a(x)$ and $b(x)$ for $x > u$ that are being estimated; sub-asymptotic forms for $a(x)$ and $b(x)$ are helpful in providing a baseline for comparison.

The sub-asymptotic behaviours that we find suggest novel parsimonious sub-asymptotic parametric forms for $a(x)$ and $b(x)$, which should reduce the sensitivity of inferences to the choice of threshold u and enable a lower threshold to be used in practice. This is important, as small differences in parameter estimates and uncertainty at finite levels can lead to large differences when extrapolating to rarer events.

Section 2 introduces the framework used to study the sub-asymptotic behaviour of the conditional tail model and our rate of convergence metrics. In Section 3, we consider three copulas for which incorporating sub-asymptotic structure can lead to improved convergence; the proofs are in the Supplementary Information. In Section 4, we unify our findings and propose sub-asymptotic parametric models that extend the Heffernan–Tawn class of norming functions.

2 | CONVERGENCE FORMULATIONS

The right-hand side of expression (2) encapsulates the limiting conditional independence of $Z_u = \{Y - a(X)\}/b(X)$ and the excesses $X - u$ for large X. We first consider the marginal limiting behaviour of Z. Under further assumptions, relating to convergence and existence of joint densities, Heffernan and Resnick (2007), Resnick and Zeber (2014) and Wadsworth et al. (2017) show that

$$\lim_{x\to\infty} \Pr\left\{\frac{Y - a(x)}{b(x)} \leq z \mid X = x\right\} = H(z), \tag{5}$$

where $a(\cdot)$, $b(\cdot)$ and $H(\cdot)$ are the same as in (2).

The purpose of our sub-asymptotic analysis is to characterise the behaviour of the remainder terms, defined in the notation of (4) by

$$a(x) - a_0(x) \sim r_a(x), \quad b(x) - b_0(x) \sim r_b(x), \qquad x \to \infty,$$

where $r_a(x)$ and $r_b(x)$ satisfy $r_a(x) = o\{a_0(x)\}$ and $r_b(x) = o\{b_0(x)\}$ as $x \to \infty$, and are to be interpreted as the leading order terms only in the differences $a(x) - a_0(x)$ and $b(x) - b_0(x)$ respectively. Specifically, we consider the second-order normalisation for $a(\cdot)$ and $b(\cdot)$, with

$$a_1(x) = a_0(x) + r_a(x), \qquad b_1(x) = b_0(x) + r_b(x). \tag{6}$$

With these sub-asymptotic forms, we are able to refine the normalisation of Y in (5), yielding the sub-asymptotic conditional distribution

$$\Pr\left\{\frac{Y - a_1(X)}{b_1(X)} \leq z \mid X = x\right\} = H_x(z), \quad x > u, \tag{7}$$

with $H_x(z) \to H(z)$ as $u \to \infty$.

Heffernan and Tawn (2004) gave the rate of convergence of the conditional distribution for various copula models in terms of $r_0(x, z) \to 0$, as $x \to \infty$ for fixed $z \in \mathbb{R}$, where

$$r_0(x, z) = \left|\Pr\left\{\frac{Y - a_0(X)}{b_0(X)} \leq z \mid X = x\right\} - H(z)\right|, \tag{8}$$

with (X, Y) on the Gumbel scale, finding that the rate at which $r_0(x, z) \to 0$ did not depend on z. We shall need similar results with Laplace margins. We consider how much we can improve the convergence rate of $r_0(x, z)$, when using the sub-asymptotic norming, by studying the rate of convergence to zero of

$$r_1(x, z) = \left|\Pr\left\{\frac{Y - a_1(X)}{b_1(X)} \leq z \mid X = x\right\} - H(z)\right|. \tag{9}$$

We also want to quantify the sub-asymptotic remainder, using

$$r_1^{(s)}(x, z) = \left|\Pr\left\{\frac{Y - a_1(x)}{b_1(x)} \leq z \mid X = x\right\} - H_x(z)\right|. \tag{10}$$

We hope to show that $r_1^{(s)}(x, z) = o\{r_1(x, z)\}$ and $r_1(x, z) = o\{r_0(x, z)\}$ as $x \to \infty$ for all z, and that the rates of convergence to zero for the distances r, r_1 and $r_1^{(s)}$ do not depend on z. Section 3 gives two examples where this improved convergence is achieved and one where it is impossible to find better normalisations than $a_0(x)$ and $b_0(x)$. We shall present these rates on a scale that is invariant to the marginal choice, by converting to a return period n, where $\Pr(X > x) = n^{-1}$.

3 | EXAMPLES

3.1 | Gaussian distribution

Let (V, W) have a bivariate standard normal distribution with nonzero correlation ρ and let (X, Y) be its marginal transform to the Laplace scale, given by

$$X = \begin{cases} -\log 2\{1 - \Phi(V)\}, & V > 0, \\ \log 2\Phi(V), & V \leq 0, \end{cases}$$

and similarly for Y as a function of W. The dependence structure of (X, Y) displays asymptotic independence, as $\chi = 0$ when $\rho < 1$ and $\bar{\chi} = \rho$ when $\rho \in [-1, 1]$.

Theorem 1. For (X, Y) with the Gaussian dependence structure just defined and $\rho \neq 0$, the ultimate and sub-asymptotic normings (4) and (6) for Y given that $X = x$, with x positive and large, have

$$\begin{aligned} a_0(x) &= \operatorname{sign}(\rho)\rho^2 x, \quad & a_1(x) &= \operatorname{sign}(\rho)\rho^2 x + \frac{1}{2}(1-\rho^2)\log(\pi x) + \tfrac{1}{2}\log(\rho^2), \\ b_0(x) &= x^{1/2}, & b_1(x) &= x^{1/2}\left\{1 - \frac{\rho^2\log(\pi x) - 2}{4x\rho^2}\right\}. \end{aligned} \tag{11}$$

The limit distribution $H(z)$ in (5) is Gaussian with mean zero and variance $2\rho^2(1-\rho^2)$, and the sub-asymptotic distribution (7) is

$$H_x(z) = \Phi\left[\frac{z}{\{2\rho^2(1-\rho^2)\}^{1/2}} + \frac{z^2}{\{2^5(1-\rho^2)\rho^3 x\}^{1/2}}\right], \quad z > -2(|\rho|\, x)^{1/2}.$$

For all z, $r_0(x, z) = O\{\log(x)x^{-1/2}\}$, $r_1(x, z) = O(x^{-1/2})$ and $r_1^{(s)}(x, z) = O\{\log^2(x)x^{-3/2}\}$.

If we choose x such that $\Pr(X > x) = 1/n$, then the rate of convergence to the limit distribution is $O(\log\log n/\sqrt{\log n})$ using the ultimate norming in (8), which is improved to $O(1/\sqrt{\log n})$ by the sub-asymptotic norming in (9), and the sub-asymptotic remainder (10) is $O\{(\log\log n)^2/(\log n)^{3/2}\}$.

If $b_0(x) = 1 + (\rho^2 x)^{1/2}$ then Theorem 1 still holds for all ρ, including $\rho = 0$, though the variances of H and H_x no longer have the ρ^2 term and the third-order terms of $a_1(x)$ and $b_1(x)$ diverge as $\rho \to 0$. Moreover H_x is not Gaussian for finite x, and has a truncated lower tail, with the truncation diminishing as $x \to \infty$.

When assessing the performance of different methods to fit the Heffernan and Tawn model to simulated data above a finite threshold u, it is tempting to use the limiting norming functions $a_0(x)$ and $b_0(x)$ in (4) as the true values of the location and scale functions in (2). In the case of simulated data from the Gaussian copula, Theorem 1 shows that this can be misleading, as the sub-asymptotic norming $a_1(\cdot)$, $b_1(\cdot)$ gives a better approximation to the location and scale functions above u. By replacing x by the threshold u in the logarithmic terms of (11) and taking u large, we derive second-order approximations for $\alpha = a_1(x)/x$ and $\beta = \log b_1(x)/\log x$ of the forms

$$\tilde{\alpha}_1 = \operatorname{sign}(\rho)\rho^2 + \frac{(1-\rho^2)\log(\pi u)}{2u} + \frac{\log(\rho^2)}{2u} \approx \operatorname{sign}(\rho)\rho^2 + \frac{(1-\rho^2)\log(u)}{2u}, \tag{12}$$

$$\tilde{\beta}_1 \approx \frac{1}{2} - \frac{1}{4u}. \tag{13}$$

Figure 1 illustrates convergence of these approximations when $\rho = 0.5$, and for values of u corresponding to Laplace quantiles from 0.975 to the 0.99998, for both the simplified forms in (12) and (13) and those including the next-order term from (11). Convergence is very slow, so it makes sense to consider second-order approximations when measuring the adequacy of finite-sample estimates. For an idea of the amount of data needed to reach such quantiles, we change the scale of the abscissa to the return period scale, using

$$\frac{1}{1 - F_L(x)} \times \frac{1}{n_Y},$$

with $F_L(\cdot)$ the Laplace distribution function, x any quantile on the Laplace scale and $n_Y = 365.25$ the number of observations per year. Even with the equivalent of more than 100 years of daily data, the parameters differ strikingly from their asymptotic values.

3.2 | Inverted logistic distribution

We consider a bivariate random vector (X, Y) with inverted logistic distribution and Laplace margins (Ledford & Tawn 1997; Papastathopoulos & Tawn 2016). Its joint survivor function is

$$\Pr(X > x, Y > y) = \exp\left[-V\left\{\frac{-1}{\log\left(\frac{1}{2}e^{-x}\right)}, \frac{-1}{\log\left(\frac{1}{2}e^{-y}\right)}\right\}\right], \quad x, y > 0,$$

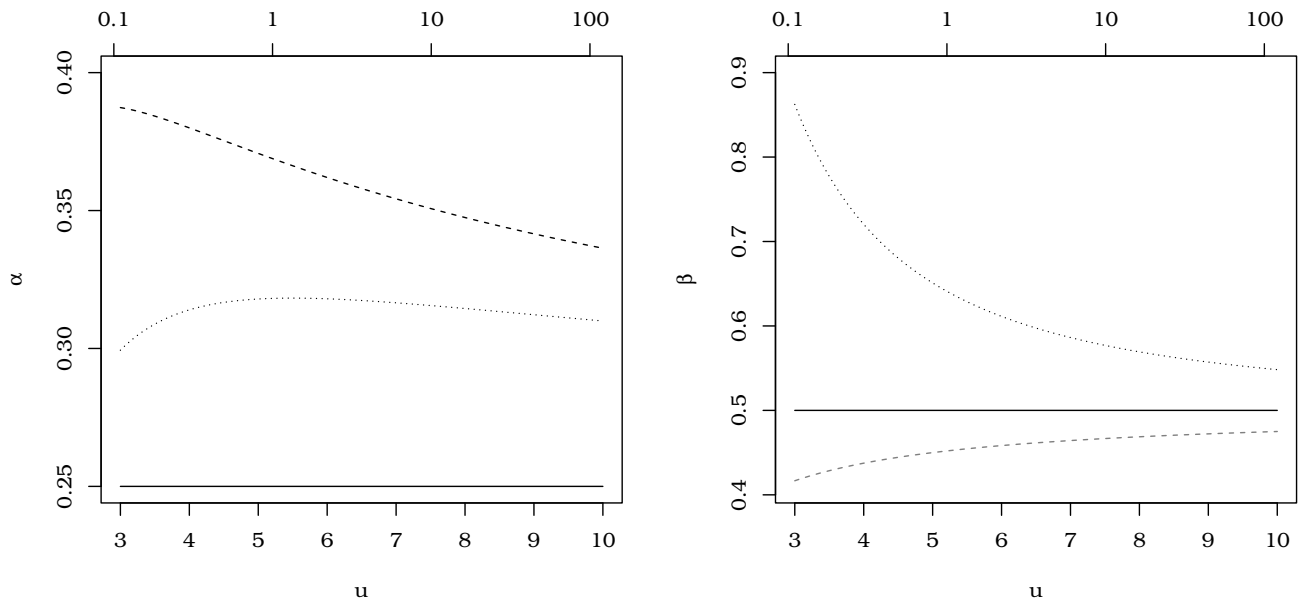

FIGURE 1 Comparison of first-order (black), second-order from (12) and (13) (dashed) and from (11) (dotted) approximations to the Heffernan–Tawn parameters α and β for a Gaussian copula with correlation $\rho = 0.5$. Lower abscissa on Laplace scale; upper abscissa on the return period scale, in years, assuming daily observations.

where

$$V(z,w) = (z^{-1/\gamma} + w^{-1/\gamma})^{\gamma}, \quad 0 < \gamma \leq 1, \tag{14}$$

is the logistic exponent function. Here $\chi = 0$ and $\bar{\chi} = 2^{1-\gamma} - 1$ for $0 < \gamma \leq 1$, with $\gamma \to 0$ corresponding to complete dependence and $\gamma = 1$ corresponding to independence.

Theorem 2. Let (X, Y) have a bivariate inverted logistic distribution with dependence parameter $0 < \gamma \leq 1$ and Laplace margins. Then the asymptotic and sub-asymptotic normings (4) and (6) for Y given that $X = x$, with x large, are

$$a_0(x) = 0, \qquad a_1(x) = -\log 2,$$
$$b_0(x) = x^{1-\gamma}, \quad b_1(x) = x^{1-\gamma},$$

so the sub-asymptotic form for $b_1(\cdot)$ equals $b_0(x)$.

The limit distribution $H(z)$ in (5) is Weibull, specifically $\bar{H}(z) = 1 - H(z) = \exp(-\gamma z^{1/\gamma})$, and the sub-asymptotic distribution $H_x(\cdot)$ in (7) satisfies

$$-\log \bar{H}_x(z) = \gamma z^{1/\gamma} + (1-\gamma)[(1-\log 2)z^{1/\gamma} - \gamma z^{2/\gamma}/2]/x, \tag{15}$$

with bounded support

$$[0, z_x^H] = [0,\ \{x/(1-\gamma) + (1-\log 2)/\gamma\}^{\gamma}] \longrightarrow \mathbb{R}_+, \quad x \to \infty.$$

If $\Pr(X > u) = n^{-1}$, the rate of convergence to the limit distribution is $r_0(x,z) = O\{(\log n)^{-(1-\gamma)}\}$ using the ultimate norming in (8) and $r_1(x,z) = O\{(\log n)^{-1}\}$ using the sub-asymptotic norming in (9). The sub-asymptotic remainder (10) is $r_1^{(s)}(x,z) = O\left\{(\log n)^{-(2-\gamma)}\right\}$.

Figure 2, which illustrates the convergence of H_x to H for $\gamma = 1/3$, 2/3 and 3/4, and x corresponding to 0.8, 0.9, 0.95 and 0.99 quantiles, suggests that the adequacy of the approximation in the upper tail depends heavily on γ.

3.3 | Logistic distribution

Let (X, Y) have a bivariate logistic distribution with Laplace margins,

$$\Pr(X \leq x,\ Y \leq y) = \exp\left[-V\left\{\frac{-1}{\log\left(1 - \frac{1}{2}e^{-x}\right)}, \frac{-1}{\log\left(1 - \frac{1}{2}e^{-y}\right)}\right\}\right], \quad x, y > 0,$$

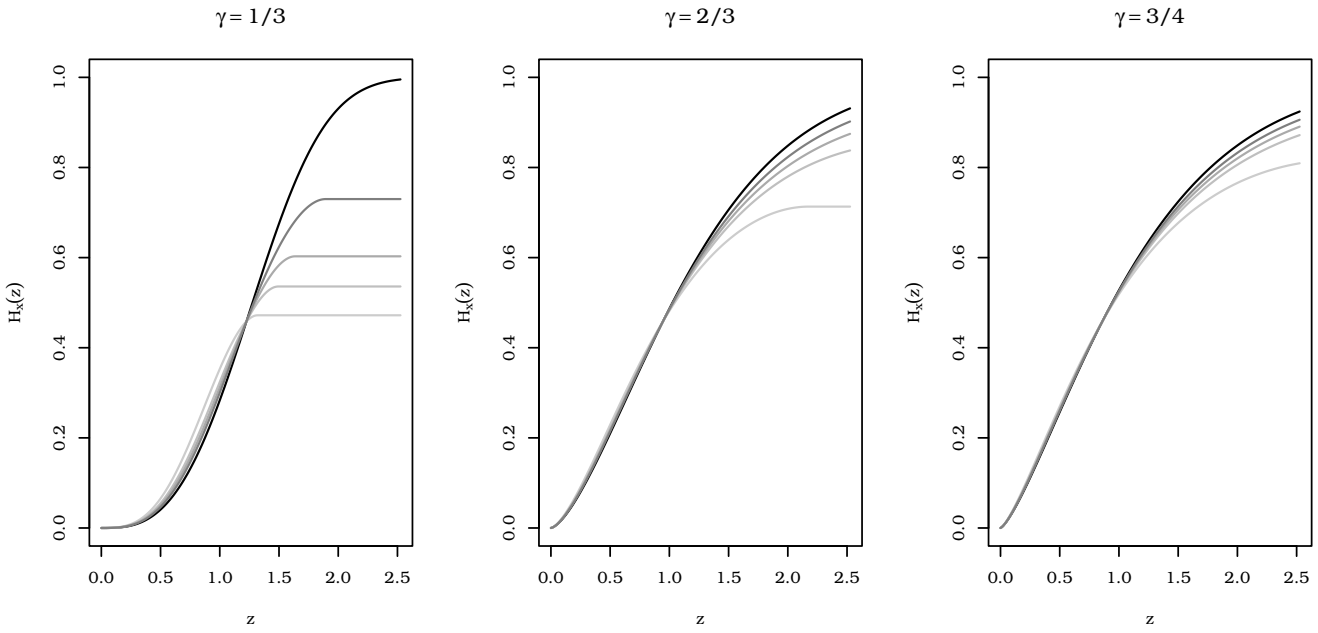

FIGURE 2 Convergence of the sub-asymptotic distribution H_x (grey) towards H (black) for the inverted logistic distribution, with x corresponding to the 0.8, 0.9, 0.95 and 0.99 quantiles (from light to dark). The panels illustrate the convergence for dependence parameter $\gamma = 1/3, 2/3, 3/4$.

with V given at (14). In the following, we do not consider the case $\gamma = 1$ corresponding to complete independence. The degree of asymptotic dependence is $\chi = 2 - 2^{\gamma}$.

Theorem 3. Let (X, Y) have a bivariate inverted logistic distribution with dependence parameter $\gamma \in (0, 1]$ and Laplace margins. Then the ultimate normings (4) for Y given that $X = x$, with x large, are $a_0(x) = x$ and $b_0(x) = 1$, the sub-asymptotic normings (6) are identical to a_0 and b_0, and $r_0(x, z) = O\{r_1(x, z)\} = O\{r_1^{(s)}(x, z)\}$.

4 | SUB-ASYMPTOTIC MODEL

Based on the examples studied in Section 3, in this section we suggest a class of sub-asymptotic models for the Heffernan and Tawn (2004) model that improves convergence rates relative to the limit model and contains the models of Tendijck et al. (2020) and Simpson et al. (2020) and all the terms that improved convergence rates for the three copulas studied above. This model should yield better statistical inferences than the canonical formulation (4). The proposed extension is parsimonious, with just two further parameters in its simplest form,

$$a(x) = \alpha x + \alpha_0 + \frac{\mathcal{L}_a(x)}{x^{\gamma_a}}, \qquad \log b(x) = \left\{\beta + \frac{\mathcal{L}_b(x)}{x^{\gamma_b}}\right\} \log x,$$

where α and β are from the first-order norming functions a_0 and b_0, $\gamma_a > -1$, $\gamma_b \geq 0$, and the functions $\mathcal{L}_a$, $\mathcal{L}_b$ are slowly varying at infinity. For statistical modelling we must specify $\mathcal{L}_a$ and $\mathcal{L}_b$, and as in earlier work (e.g., Ledford & Tawn 1996) we fix them to be constant above some threshold u, i.e., $\mathcal{L}_a(x) = \delta_a$ and $\mathcal{L}_b(x) = \delta_b$ for $x > u$. Second-order effects are hard to estimate, so in practice it may suffice to set $\gamma_a = \gamma_b = 1$.

Author contributions

This work is part of the PhD thesis of TL, jointly supervised by ACD and JAT. The bulk of the theoretical development was done by TL, with help from ACD and, principally, JAT. All authors contributed to the writing of the paper.

Financial disclosure

This research was partially supported by the Swiss National Science Foundation.

Conflict of interest

The authors declare no potential conflict of interests.

Acknowledgements

We thank the reviewers for helpful comments.

References

Asadi, P., Davison, A. C., & Engelke, S. (2015). Extremes on river networks. *The Annals of Applied Statistics*, *9*, 2023–2050.

Beranger, B., Padoan, S. A., & Sisson, S. A. (2017). Models for extremal dependence derived from skew-symmetric families. *Scandinavian Journal of Statistics*, *44*(1), 21-45.

Beranger, B., Padoan, S. A., Xu, Y., & Sisson, S. A. (2019). Extremal properties of the multivariate extended skew-normal distribution. *Statistics and Probability Letters*, *147*, 105-114.

Bofinger, E., & Bofinger, V. J. (1965). The correlation of maxima in samples drawn from a bivariate normal distribution. *Australian Journal of Statistics*, *7*, 57–61.

Bofinger, V. J. (1970). The correlation of maxima in several bivariate non-normal distributions. *Australian Journal of Statistics*, *12*, 1–7.

Cai, J.-J., Einmahl, J. H. J., & de Haan, L. (2011). Estimation of extreme risk regions under multivariate regular variation. *The Annals of Statistics*, *39*, 1803–1826.

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

Coles, S. G., & Tawn, J. A. (1991). Modelling extreme multivariate events. *Journal of the Royal Statistical Society Series B*, *53*, 377–392.

Coles, S. G., & Tawn, J. A. (1996). Modelling extremes of the areal rainfall process. *Journal of the Royal Statistical Society Series B*, *58*, 329–347.

de Haan, L. (1985). Extremes in higher dimensions: the model and some statistics. In *Proceedings of the 45th Session of the International Statistical Institute* (p. 26.3).

de Haan, L., & de Ronde, J. (1998). Sea and wind: multivariate extremes at work. *Extremes*, *1*, 7–45.

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

de Valk, C. (2016). Approximation and estimation of very small probabilities of multivariate extreme events. *Extremes*, *19*, 687–717.

Heffernan, J. E. (2000). A directory of coefficients of tail dependence. *Extremes*, *3*, 279–290.

Heffernan, J. E., & Resnick, S. I. (2007). Limit laws for random vectors with an extreme component. *The Annals of Applied Probability*, *17*, 537–571.

Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society Series B*, *66*, 497–546.

Hilal, S., Poon, S.-H., & Tawn, J. A. (2011). Hedging the Black Swan: conditional heteroskedasticity and tail dependence in S&P500 and VIX. *Journal of Banking and Finance*, *35*, 2374–2387.

Hilal, S., Poon, S.-H., & Tawn, J. A. (2014). Portfolio risk assessment using multivariate extreme value methods. *Extremes*, *17*, 531–556.

Huser, R. G., & Davison, A. C. (2014). Space-time modelling of extreme events. *Journal of the Royal Statistical Society Series B*, *76*, 439–461.

Huser, R. G., & Wadsworth, J. L. (2019). Modeling spatial processes with unknown extremal dependence class. *Journal of the American Statistical Association*, *114*(525), 434–444.

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

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*, 1287–1304.

Keef, C., Papastathopoulos, I., & Tawn, J. A. (2013). Estimation of the conditional distribution of a multivariate variable given that one of its components is large: additional constraints for the Heffernan and Tawn model. *Journal of Multivariate Analysis*, *115*, 396–404.

Keef, C., Svensson, C., & Tawn, J. A. (2009). Spatial dependence in extreme river flows and precipitation for Great Britain. *Journal of Hydrology*, *378*, 240–252.

Keef, C., Tawn, J. A., & Svensson, C. (2009). Spatial risk assessment for extreme river flows. *Journal of the Royal Statistical Society Series C*, *58*, 601–618.

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

Ledford, A. W., & Tawn, J. A. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society Series B*, *59*, 475–499.

Lugrin, T. (2018). *Semiparametric Bayesian Risk Estimation for Complex Extremes* (Unpublished doctoral dissertation). Ecole Polytechnique Fédérale de Lausanne.

Papastathopoulos, I., & Tawn, J. A. (2016). Conditioned limit laws for inverted max-stable processes. *Journal of Multivariate Analysis*, *150*, 214–228.

Pickands, J. (1981). Multivariate extreme value distributions. *Bulletin of the International Statistical Institute*, *49*, 859–878.

Poon, S.-H., Rockinger, M., & Tawn, J. A. (2003). Modelling extreme-value dependence in international stock markets. *Statistica Sinica*, *13*, 929–953.

Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer.

Resnick, S. I., & Zeber, D. (2014). Transition kernels and the conditional extreme value model. *Extremes*, *17*, 263–287.

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

Simpson, E. S., Opitz, T., & Wadsworth, J. L. (2020). *High-dimensional modeling of spatial and spatio-temporal conditional extremes using INLA and the SPDE approach.* arXiv:2011.04486.

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.

Süveges, M., & Davison, A. C. (2012). A case study of a "Dragon-King": the 1999 Venezuelan catastrophe. *The European Physical Journal Special Topics*, *205*, 131–146.

Tawn, J. A., Shooter, R., Towe, R. P., & Lamb, R. (2018). Modelling spatial extreme events with environmental applications. *Spatial Statistics*, *28*, 35–58.

Tendijck, S., Eastoe, E. F., Tawn, J. A., Randell, D., & Jonathan, P. (2020). *Modelling the extremes of bivariate mixture distributions with application to oceanographic data.* Submitted.

Tiago de Oliveira, J. (1962/63). Structure theory of bivariate extremes, extensions. *Estudos de Matemática, Estatística e Econometria*, *7*, 165–195.

Wadsworth, J. L., & Tawn, J. A. (2013). A new representation for multivariate tail probabilities. *Bernoulli*, *19*, 2689–2714.

Wadsworth, J. L., Tawn, J. A., Davison, A. C., & Elton, D. M. (2017). Modelling across extremal dependence classes. *Journal of the Royal Statistical Society Series B*, *79*, 149–175.

How to cite this article: Lugrin, T., Tawn, J. A., and A. C. Davison (2021), Sub-asymptotic motivation for new conditional multivariate extremes models