Limiting Behavior of Maxima under Dependence

Klaus Herrmann[1], Marius Hofert[2], Johanna G. Nešlehová[3]

2024-05-07

Weak convergence of maxima of dependent sequences of identically distributed continuous random variables is studied under normalizing sequences arising as subsequences of the normalizing sequences from an associated iid sequence. This general framework allows one to derive several generalizations of the well-known Fisher–Tippett–Gnedenko theorem under conditions on the univariate marginal distribution and the dependence structure of the sequence. The limiting distributions are shown to be compositions of a generalized extreme value distribution and a distortion function which reflects the limiting behavior of the diagonal of the underlying copula. Uniform convergence rates for the weak convergence to the limiting distribution are also derived. Examples covering well-known dependence structures are provided. Several existing results, e.g. for exchangeable sequences or stationary time series, are embedded in the proposed framework.

Keywords
copula, dependent sequences, distortion function, extremes, maxima, time series, uniform convergence rate

MSC2010
60G70, 62H05

1 Introduction

Consider a sequence $(X_i, i \in \mathbb{N})$ (in short: (X_i)) of random variables with common distribution function F, i.e. $X_1, X_2, \dots \sim F$. The purpose of this article is to study the limiting behavior of the maximum of the first n variables of this sequence under suitable normalization. That is, we seek to identify conditions under which there exists a nondegenerate distribution function G and normalizing sequences (c_n), $c_n > 0$, and (d_n) of constants such that $M_n = \max(X_1, \dots, X_n)$ satisfies the convergence in distribution

$$\frac{M_n - d_n}{c_n} \xrightarrow{\mathrm{d}} G. \tag{1}$$

[1]Département de Mathématiques, Université de Sherbrooke, klaus.herrmann@usherbrooke.ca
[2]Department of Statistics and Actuarial Science, The University of Hong Kong, mhofert@hku.hk
[3]Department of Mathematics and Statistics, McGill University, johanna.neslehova@mcgill.ca

If (1) holds, we say that F is in the maximum domain of attraction of G, denoted by $F \in \mathrm{MDA}(G)$. Our further interest lies in identifying the limit G. Answering these questions is essential for applications that require the extrapolation into the tail of the unknown underlying distribution function F; this is a cornerstone of extreme value analysis as described in standard textbooks such as Beirlant et al. (2004), Haan and Ferreira (2006), Embrechts et al. (1997), and Resnick (1987).

When the random variables X_i, $i \in \mathbb{N}$, are iid the problem has been solved in Fisher and Tippett (1928) and Gnedenko (1943). The so-called Fisher–Tippett–Gnedenko theorem, given, e.g. in Resnick (1987, Proposition 0.3), states that if $F \in \mathrm{MDA}(G)$, then G necessarily belongs to the class of generalized extreme value distributions, denoted by $G \in \mathrm{GEV}$. Distribution functions in the GEV class have parameters $\xi, \mu \in \mathbb{R}$ and $\sigma > 0$, and they are given by

$$H_{\xi,\mu,\sigma}(x) = \begin{cases} \exp(-(1+\xi(x-\mu)/\sigma)^{-1/\xi}), & \text{if } \xi \neq 0, \\ \exp(-\mathrm{e}^{-(x-\mu)/\sigma}), & \text{if } \xi = 0, \end{cases}$$

for all $x \in \mathbb{R}$ with $1+\xi(x-\mu)/\sigma > 0$ and by the limiting values $H_{\xi,\mu,\sigma}(x) = 0$ or $H_{\xi,\mu,\sigma}(x) = 1$ otherwise. Results about the construction of the normalizing sequences in the iid case are also available, see, e.g. Resnick (1987), Embrechts et al. (1997), and Haan and Ferreira (2006).

Our main contribution is to identify the limiting behavior of appropriately centered and scaled maxima of $X_1, X_2, \ldots \sim F$ under dependence. We derive conditions on F and on the dependence structure of (X_i) under which the limiting distribution G in (1) is nondegenerate and are able to identify its form, thereby generalizing the Fisher–Tippett–Gnedenko theorem to the case of dependent sequences. In particular, we will show that G is a composition of a generalized extreme value distribution and a univariate function driven by the dependence structure of the sequence (X_i).

Maxima of dependent sequences are of particular interest in actuarial science, where the principal idea of insurable exposure is that a large number of similar risks are pooled into homogeneous portfolios; see, e.g. Mehr and Cammack (1980, Chapter 2). The behavior of the largest loss in such a portfolio then plays a key role for risk assessment. In such situations and with a large portfolio approximation in mind, one often assumes that risks are identically distributed. Also, such risks are typically dependent due to common factors affecting all risks.

The classical univariate Fisher–Tippett–Gnedenko theorem was extended early on to incorporate specific types of dependence between the random variables X_i, $i \in \mathbb{N}$. Berman Berman (1962b) derived general limiting results for maxima of exchangeable sequences. The special case where the dependence between the random variables of the exchangeable sequence is given by an Archimedean copula was considered in Ballerini (1994) and Wüthrich (2004); extremal properties under Archimedean and related dependence structures were crucial in modeling collateralized debt obligations, e.g. in Schönbucher and Schubert (2001) and Hofert and Scherer (2011). Almost sure limit theorems for maxima under Archimedean dependence were recently studied in Dudziński and Furmańczyk (2017), while Huang et al. (2017) proposed adaptations of the block maxima and peaks-over-threshold methods for exchangeable sequences.

Furthermore, an extensive body of literature exists when (X_i) is a stationary time series. The early work of Watson (1954) under m-dependence, and Loynes (1965) under α-mixing was substantially extended and refined in Leadbetter (1974) and Leadbetter (1983), culminating in Leadbetter et al. (1983). Numerous statistical inference procedures for tail extrapolation in the context of stationary time series have also been developed, see e.g. Beirlant et al. (2004) and Fawcett and Walshaw (2012); a recent review with extensions is provided by Buriticá et al. (2021). Alternatively, Ferreira and Ferreira (2018) consider a local dependence condition for stationary time series, while Haan et al. (2016) and Chavez-Demoulin and Guillou (2018) work under a β-mixing framework. Finally, Russell and Huang (2021) use a bivariate Gumbel copula to model the dependence between consecutive block maxima in a first order Markov process.

Contrary to the existing literature, we study the limiting behavior of M_n in general, i.e. without assuming exchangeability or stationarity. This context is akin to the order statistics literature, which studies the distribution of maxima for various multivariate distributions in finite samples, see e.g. Balakrishnan et al. (1992), Rychlik (1994), and Arellano-Valle and Genton (2008). We only require throughout the classical assumption that the common distribution function F be continuous. The dependence between the random variables $(X_i, i \in \mathbb{N})$ can then be uniquely described using copulas following Sklar's theorem Sklar (1959). Specifically, for each $n \in \mathbb{N}$, there exists a unique copula C_n, i.e. a distribution function on $[0,1]^n$ with standard uniform univariate margins, such that $\mathbb{P}(X_1 \le x_1, \ldots, X_n \le x_n) = C_n(F(x_1), \ldots, F(x_n))$, $(x_1, \ldots, x_n) \in \mathbb{R}^n$. With this representation, we obtain $\mathbb{P}(M_n \le x) = \delta_n(F(x))$ where $\delta_n(u) = C_n(u, \ldots, u)$, $u \in [0,1]$, is the *diagonal* of C_n. As we show in Section 2, the limiting behavior of M_n under dependence is determined by the tail behavior of F and its interplay with the properties of the copula diagonal δ_n. Investigating this relationship allows us to derive a rigorous theoretical framework for the study of the limiting behavior of maxima under dependence.

Our first main result (Theorem 2.2) provides conditions under which maxima of identically distributed dependent random variables (X_i) can be normalized using the same sequences (c_n), $c_n > 0$, and (d_n) of normalizing constants as in the iid case. Our second main result (Theorem 2.10) is then a generalization of the Fisher–Tippett–Gnedenko theorem under dependence. In both cases the limiting distribution of appropriately stabilized maxima is of the form $D \circ H$ for $H \in \mathrm{GEV}$ and D being a *distortion function*, i.e. a distribution function satisfying $D(0) = 0$ and $D(1) = 1$. Uniform convergence rates for the weak convergence of normalized maxima to the limit $D \circ H$ are derived in Section 5(Theorem 5.1).

The results from Section 2 are illustrated through various examples. In Section 3, we first consider so-called power diagonals which still lead to limiting distributions of maxima that are generalized extreme value. We also establish connections between our findings and well-known results about extremes of stationary time series with restricted short-range dependence at extreme levels. Complementing these results, we also provide examples that illustrate how our methodology can be used for time series with long-range dependence. Section 4 then presents examples where the limiting distribution of maxima is no longer generalized extreme value. Notably, we explore the case where C_n is Archimedean or Archimax and provide a generalization of multiplicative frailty models linked to the Fréchet domain of attraction. Conclusions are given in Section 6. Proofs of all results are provided in the Supplementary Material Herrmann et al. (2024).

2 Weak convergence of maxima of dependent sequences

In this section, we derive a generalization of the Fisher–Tippett–Gnedenko theorem when $X_1, X_2, \ldots \sim F$ are identically distributed but dependent. As stated before, we assume that F is continuous and denote by C_n and δ_n the unique copula of $(X_1, \ldots, X_n)$ and its diagonal, respectively.

For a fixed n, general properties of copula diagonals have been studied, e.g., by Jaworski (2009); the connection between the maximum M_n of $X_1, \ldots, X_n$ and δ_n is also made in Hofert et al. (2013) who use it to construct a maximum likelihood estimator for the parameters of Archimedean copula models. In order to exploit the latter connection in an asymptotic context when $n \to \infty$, we first define the so-called diagonal power distortion of δ_n.

Definition 2.1
Let $X_1, X_2, \ldots \sim F$ and C_n be the unique copula of $(X_1, \ldots, X_n)$ with diagonal δ_n. Let the rate $r \colon \mathbb{N} \to (0, \infty)$ be an arbitrary strictly positive function; for simplicity, we write r_n for $r(n)$. The *diagonal power distortion* with respect to the rate r is $D_n^r \colon [0,1] \to [0,1]$ given, for all $u \in [0,1]$, by $D_n^r(u) = \delta_n(u^{1/r_n})$.

Because C_n is componentwise non-decreasing and uniformly continuous Nelsen (2006), D_n^r is a continuous distribution function on $[0,1]$ for any positive rate function r. We begin our investigation towards a generalization of the Fisher–Tippett–Gnedenko theorem for dependent sequences by describing the precise impact of δ_n on the limiting distribution of maxima.

Theorem 2.2
Consider $X_1, X_2, \ldots \sim F$ and let $M_n^* = \max(X_1^*, \ldots, X_n^*)$ denote the maximum corresponding to an iid sequence $X_1^*, X_2^*, \ldots \sim F$. Assume that $F \in \mathrm{MDA}(H)$ for a GEV distribution H with normalizing sequences (d_n^*) and (c_n^*), $c_n^* > 0$, i.e. for any $x \in \mathbb{R}$, $\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n^* - d_n^*}{c_n^*} \leq x\right) = H(x)$.

(i) If there exists a function $r \colon \mathbb{N} \to (0, \infty)$ such that $\lim_{n\to\infty} r_n = \infty$ and D_n^r converges pointwise to a continuous function D on $[0,1]$, then, for all $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - d^*_{\lceil r_n \rceil}}{c^*_{\lceil r_n \rceil}} \leq x\right) = D \circ H(x) \tag{2}$$

and D is a continuous distribution function on $[0,1]$.

(ii) If there exists a function $r\colon \mathbb{N} \to (0,\infty)$ such that $1/r_n = O(1/n)$ and a distribution function D on $[0,1]$ so that, for all continuity points x of $D \circ H$, (2) holds, then $\lim_{n\to\infty} D_n^r(u) = D(u)$ for $u \in \{0,1\}$ and all continuity points $u \in (0,1)$ of D.

Remark 2.3

Although some of the results in Theorem 2.2 can possibly be extended to the case when F is not continuous, we do not consider such a generalization here. First, $F \in \mathrm{MDA}(H)$ does not hold for several well-known discrete distributions even in the iid case Embrechts et al. (1997, Section 3.1) and other approaches, such as maxima of triangular arrays, are typically considered (Anderson et al. (1997)). Second, the copula is no longer unique when F is not continuous and this complicates matters in the present context; see e.g. Feidt et al. (2010).

Remark 2.4

$D_n^r(u)$ can be interpreted as a probability. Indeed, for $(U_1,\dots,U_n) \sim C_n$, $D_n^r(u) = \delta_n(u^{1/r_n}) = \mathbb{P}(U_1 \le u^{1/r_n},\dots,U_n \le u^{1/r_n}) = \mathbb{P}(\{\max(U_1,\dots,U_n)\}^{r_n} \le u)$. Its limiting behavior as $n \to \infty$ can thus be rephrased in terms of weak convergence of the maximum of dependent standard uniform random variables under power normalization to a continuous limit. Theorem 2.2 (i) thus complements and generalizes results of limiting distributions under power normalization in the iid case, pioneered in Pancheva (1985). Due to the positivity of the standard uniform random variables we also have a direct connection to scale transformations via logarithms. If W_i denotes a standard reverse Weibull random variable, i.e. $\mathbb{P}(W_i \le x) = \exp(x)$ for $x \le 0$, taking logarithms leads to $D_n^r(u) = \mathbb{P}(\max(W_1,\dots,W_n) \le \log(u)/r_n)$, where the copula of $(W_1,\dots,W_n)$ is again C_n because of the invariance of copulas with respect to strictly increasing transformations Nelsen (2006, Theorem 2.4.3).

Theorem 2.2 shows that, under suitable conditions on r, the convergence of the diagonal power distortion D_n^r to a limit D is necessary and sufficient for (1) to hold with $G = D \circ H$, provided that the sequences (c_n) and (d_n) are suitable subsequences of (c_n^*) and (d_n^*). In this setup, we necessarily have that $H \in \mathrm{GEV}$ by the classical Fisher–Tippett–Gnedenko theorem. The requirement that $r_n \to \infty$ is central for these statements to be true. To see this, consider a sequence with $X_i = X_1$ for all $i \ge 2$ almost surely. In this case, for every $n \in \mathbb{N}$, C_n is the comonotone copula given, for all $u_1,\dots,u_n \in [0,1]$, by $\min(u_1,\dots,u_n)$ and thus $\delta(u) = u$ for all $u \in [0,1]$. Then $M_n = X_1$ almost surely and (1) holds with $c_n = 1$, $d_n = 0$, $n \in \mathbb{N}$, and $G = F$. This means that the limiting distribution G can be entirely arbitrary here.

We now consider what happens if (instead of assuming $r_n \to \infty$ as in Theorem 2.2) $r_n \to \varrho$ for $n \to \infty$ and $\varrho \in (0,\infty)$. In this case, there is no need to stabilize M_n or to assume that F is in the maximum domain of attraction of a nondegenerate distribution function; however, the limit in (1) is not necessarily generalized extreme value or a distortion thereof anymore.

Proposition 2.5

Let $X_1, X_2, \ldots \sim F$ and $M_n = \max(X_1,\dots,X_n)$, $n \in \mathbb{N}$.

(i) If there exists $r\colon \mathbb{N} \to (0,\infty)$ such that $r_n \to \varrho \in (0,\infty)$ as $n \to \infty$ and D_n^r converges pointwise to a continuous function D on $[0,1]$, then, for all $x \in \mathbb{R}$, $\mathbb{P}(M_n \leq x) \to D(F^\varrho(x))$ as $n \to \infty$.

(ii) Conversely, if $r_n \to \varrho \in (0,\infty)$ such that $|r_n - \varrho| = o(1/n)$, and $\mathbb{P}(M_n \leq x) \to D(F^\varrho(x))$ as $n \to \infty$ for all continuity points of $D \circ F^\varrho$, then $\lim_{n\to\infty} D_n^r(u) = D(u)$ for $u \in \{0,1\}$ and all continuity points $u \in (0,1)$ of D.

The following two examples illustrate Proposition 2.5 for nontrivial dependent sequences.

Example 2.6

Consider again $X_i = X_1$ for all $i \geq 2$ almost surely. Therefore, if we consider any strictly positive function r with the property that $r_n \to \varrho$ as $n \to \infty$ for some $\varrho \in (0,\infty)$, we have that $D_n^r(u) \to u^{1/\varrho}$ and the limit satisfies $D(F^\varrho(x)) = F(x)$ for all $x \in \mathbb{R}$. This example readily generalizes to an eventually almost surely constant sequence of the form $X_1,\dots,X_K,X_K,X_K,\dots$. If δ_K denotes the diagonal of the copula C_K of $(X_1,\dots,X_K)$, then D_n^r converges pointwise to $D(u) = \delta_K(u^{1/\varrho})$ for all $u \in [0,1]$ for any r that satisfies $r_n \to \varrho$ as $n \to \infty$ for some $\varrho \in (0,\infty)$. The limiting distribution of M_n is then $\delta_K \circ F$.

Example 2.7

A special case of the multivariate Cuadras–Augé copula Cuadras and Augé (1981), with parameter $\theta \in (0,1)$, is given, for all $u_1,\dots,u_n \in [0,1]$, by $C_n(u_1,\dots,u_n) = \prod_{i=1}^n u_{(i)}^{(1-\theta)^{i-1}}$, where $u_{(1)} \leq \ldots \leq u_{(n)}$ are the evaluation points sorted in ascending order; see Mai and Scherer (2009). Mai (2018, Example 5) shows that there indeed exists an infinite sequence of identically distributed random variables with the property that for each $n \geq 2$, the copula C_n of $(X_1,\dots,X_n)$ is a Cuadras–Augé copula of this form with parameter $\theta \in (0,1)$. The diagonal of C_n has the form $\delta_n(u) = u^{(1-(1-\theta)^n)/\theta}$ for all $u \in [0,1]$. If $r_n = (1-(1-\theta)^n)/\theta$, then $r_n \to 1/\theta$ as $n \to \infty$. Furthermore, for all $u \in [0,1]$, $D_n^r(u) = D(u) = u$ for all $n \in \mathbb{N}$. By Proposition 2.5 (i), the limiting distribution of M_n is therefore $F^{1/\theta}$.

Proposition 2.5 (i) and the subsequent examples show that any continuous distribution can arise as a limit of suitably normalized maxima of some dependent sequence of identically distributed random variables. In particular, there is no guarantee that the limiting distribution is max-stable, which contrasts the iid case. In general there is thus no hope of obtaining a single nontrivial class of all possible nondegenerate limits of suitably normalized maxima of identically distributed dependent random variables. Nonetheless, two analogues of the Fisher–Tippett–Gnedenko theorem can be obtained under specific assumptions. Assuming $F \in \mathrm{MDA}(H)$ for some $H \in \mathrm{GEV}$, Theorem 2.10 is a simple consequence of Theorem 2.2 (i) and the convergence to types theorem Resnick (1987, Proposition 0.2).

Corollary 2.8

Consider $X_1, X_2, \ldots \sim F$ and let, for $n \in \mathbb{N}$, δ_n be the diagonal of the copula C_n of $(X_1,\dots,X_n)$. Assume that $F \in \mathrm{MDA}(H_{\xi,\mu,\sigma})$ with $H_{\xi,\mu,\sigma} \in \mathrm{GEV}$ and that there exists $r\colon \mathbb{N} \to (0,\infty)$ such that $\lim_{n\to\infty} r_n = \infty$ and D_n^r converges pointwise to a continuous function D on $[0,1]$. If there exist sequences (c_n), $c_n > 0$, and (d_n) such that (1) holds for a nondegenerate G, then there exist $c > 0$ and $d \in \mathbb{R}$ such that $G = D \circ H_{\xi,\tilde{\mu},\tilde{\sigma}}$, where $H_{\xi,\tilde{\mu},\tilde{\sigma}} \in \mathrm{GEV}$ with $\tilde{\mu} = (\mu - d)/c$ and $\tilde{\sigma} = \sigma/c$.

Remark 2.9

While the rate r in Corollary 2.8 is not unique, the speed at which it tends to ∞ is. Suppose that the assumptions of Corollary 2.8 hold and that (c_n^*) and (d_n^*) are the normalizing sequences of the maxima of the associated iid sequence.

Suppose $\tilde{r}\colon \mathbb{N} \to (0,\infty)$ is such that $r_n/\tilde{r}_n \to \theta$ as $n \to \infty$ where $\theta \in (0,\infty)$. Obviously, $\tilde{r}_n \to \infty$ as $n \to \infty$. Because D is continuous, $D_n^r \to D$ uniformly on $[0,1]$ by Lemma A.1 and hence $D_n^{\tilde{r}}(u) \to D(u^\theta) =: \widetilde{D}(u)$ for all $u \in [0,1]$. From Theorem 2.2 (i) we obtain that (1) holds with $c_n = c^*_{\lceil \tilde{r}_n \rceil}$, $d_n = d^*_{\lceil \tilde{r}_n \rceil}$ and $G = \widetilde{D} \circ H_{\xi,\mu,\sigma} = D \circ H^\theta_{\xi,\mu,\sigma}$. By direct calculation, $H^\theta_{\xi,\mu,\sigma} = H_{\xi,\tilde{\mu},\tilde{\sigma}}$, where $\tilde{\mu} = \mu + \sigma \log(\theta)$ if $\xi = 0$ and $\tilde{\mu} = \mu$ otherwise, while $\tilde{\sigma} = \sigma$ if $\xi = 0$ and $\tilde{\sigma} = \sigma\theta^{1/|\xi|}$ otherwise. Applying the convergence to types theorem as in the proof of Corollary 2.8 yields $c^*_{\lceil \tilde{r}_n \rceil}/c^*_{\lceil r_n \rceil} \to \sigma/\tilde{\sigma}$ and $(d^*_{\lceil \tilde{r}_n \rceil} - d^*_{\lceil r_n \rceil})/c^*_{\lceil r_n \rceil} \to \mu - \tilde{\mu}\sigma/\tilde{\sigma}$.

Now suppose that $\tilde{r}\colon \mathbb{N} \to (0,\infty)$ is such that $r_n/\tilde{r}_n \to 0$. In this case, $\tilde{r}_n \to \infty$, but this time the uniform convergence of D_n^r to D implies that $D_n^{\tilde{r}}(u) \to 1$ as $n \to \infty$ if $u \in (0,1]$, while $D_n^{\tilde{r}}(0) = 0$ for all $n \in \mathbb{N}$. Consequently, for any $x \in \mathbb{R}$ such that $H_{\xi,\mu,\sigma}(x) > 0$, $\mathbb{P}(M_n \le c^*_{\lceil \tilde{r}_n \rceil} x + d^*_{\lceil \tilde{r}_n \rceil}) \to 1$ as $n \to \infty$. Thus if $\xi \in (-\infty, 0]$, the limiting behavior of $(M_n - d^*_{\lceil \tilde{r}_n \rceil})/c^*_{\lceil \tilde{r}_n \rceil}$ is degenerate. The same is true when $\xi > 0$ and D is such that $D(u) < 1$ for all $u \in [0,1)$, as can be argued from the convergence to types theorem.

The case of $\tilde{r}\colon \mathbb{N} \to (0,\infty)$ with $r_n/\tilde{r}_n \to \infty$ and $\tilde{r}_n \to \infty$ as $n \to \infty$ is similar. We obtain that whenever $H_{\xi,\mu,\sigma}(x) < 1$, $\mathbb{P}(M_n \le c^*_{\lceil \tilde{r}_n \rceil} x + d^*_{\lceil \tilde{r}_n \rceil}) \to 0$ as $n \to \infty$. If $\xi \in [0,\infty)$, this implies that the weak limit of $(M_n - d^*_{\lceil \tilde{r}_n \rceil})/c^*_{\lceil \tilde{r}_n \rceil}$ is degenerate. For $\xi < 0$, the same follows from the convergence to types theorem if additionally $D(u) > 0$ for all $u \in (0,1]$.

The second generalization of the Fisher–Tippett–Gnedenko theorem can be derived without assuming that $F \in \mathrm{MDA}(H)$ for some $H \in \mathrm{GEV}$, at the cost of stronger assumptions on the underlying dependence structure than those in Corollary 2.8.

Theorem 2.10

Let $X_1, X_2, \ldots \sim F$. Suppose that the diagonal δ_n of the copula C_n of $(X_1, \ldots, X_n)$ is strictly increasing for each n and that there exists a function $r\colon \mathbb{N} \to (0,\infty)$, and a bijection $\lambda\colon (0,\infty) \to (0,\infty)$ such that the following conditions hold:

(i) $\lim_{n\to\infty} r_n = \infty$ and $\lim_{n\to\infty} r_{\lceil tn \rceil}/r_n = \lambda(t)$, $t > 0$;

(ii) the diagonal power distortion D_n^r with respect to r converges pointwise to a continuous and strictly increasing bijection $D\colon [0,1] \to [0,1]$.

If there exist sequences (c_n), $c_n > 0$ and (d_n) such that (1) holds for a nondegenerate G, then $G = D \circ H$ where $H \in \mathrm{GEV}$.

Remark 2.11

An interesting special case of Theorem 2.10 arises when r_n satisfies $n/r_n \to \alpha$ as $n \to \infty$

for some $\alpha > 0$. Condition (i) of Theorem 2.10 then holds with $\lambda(t) = t$ for all $t > 0$. From Corollary A.4 we have that $F^n(cx + d) \to H^\alpha(x)$ for all $x \in \mathbb{R}$, i.e. $F \in \mathrm{MDA}(H^\alpha)$, where H^α is of the same type as H given that H is max-stable.

We close this section with an example where all calculations can be done explicitly.

Example 2.12
Consider the following case of a moving maximum process of Newell (1964) and Deheuvels (1983), see Beirlant et al. (2004, Chapter 10) for an overview. For a fixed $k \geq 0$, let $(Z_i)_{i=-k+1}^{\infty}$ be a sequence of iid standard Fréchet random variables and set $Y_i = (1/(k+1)) \max_{0\leq j\leq k} Z_{i-j}$. Then Y_i is also standard Fréchet. For $n \geq k+1$ the joint distribution of $\boldsymbol{Y}_n = (Y_1, \dots, Y_n)$ can be calculated to be $F_{\boldsymbol{Y}_n}(y_1, \dots, y_n) = \prod_{i=-k+1}^{0} \mathbb{P}(Z_i \leq (k+1) \min_{1\leq \ell \leq i+k} y_\ell) \times \prod_{i=1}^{n-k} \mathbb{P}(Z_i \leq (k+1) \min_{0\leq\ell\leq k} y_{i+\ell}) \times \prod_{i=n-k+1}^{n} \mathbb{P}(Z_i \leq (k+1) \min_{i\leq\ell\leq n} y_\ell)$. Taking into account the marginal distribution of Y_i and using that $\min_{j\in J}(-1/\log(u_j))$ equals $-1/\log(\min_{j\in J} u_j)$ for any index set J, the copula of $\boldsymbol{Y}_n$ is given by

$$C_n(u_1, \dots, u_n) = \prod_{i=-k+1}^{0} \min_{1\leq\ell\leq i+k} u_\ell^{1/(k+1)} \prod_{i=1}^{n-k} \min_{0\leq\ell\leq k} u_{i+\ell}^{1/(k+1)} \prod_{i=n-k+1}^{n} \min_{i\leq\ell\leq n} u_\ell^{1/(k+1)}.$$

Comparing C_n to Mulinacci (2015, Equation (5.3)) (see also Li (2008)), the copula can be identified as Marshall–Olkin copula, i.e. the survival copula of a multivariate Marshall–Olkin distribution with appropriate choices of within-group intensities to keep only certain terms in the product over all non-empty subsets of $\{1, \dots, n\}$. We now have $\delta_n(u) = u^{(n+k)/(k+1)}$, leading to $D_n^n(u) = \delta_n(u^{1/n}) = u^{(n+k)/(n(k+1))} \to u^{1/(k+1)} = D(u)$ as $n \to \infty$.

Now pick an arbitrary continuous distribution function $F \in \mathrm{MDA}(H_{\xi,\mu,\sigma})$ and set $X_i = F^{-1}(F_{Y_i}(Y_i))$ to obtain a sequence (X_i) with univariate margins F and the Marshall–Olkin copula C_n as the copula of $(X_1, \dots, X_n)$. If (c_n^*) and (d_n^*) denote the normalizing sequences from the iid case and if the rate function is given by $r_n = n$, we recover from Theorem 2.2 (i) that $\mathbb{P}\{(M_n - d_n^*)/c_n^* \leq x\} \to H_{\xi,\mu,\sigma}^\theta(x)$, $x \in \mathbb{R}$, where $\theta = 1/(k+1) \in (0, 1]$ and $M_n = \max(X_1, \dots, X_n)$. The fact that the distortion $D(u) = u^\theta$ is a power function has the interesting effect that the weak limit of the maximum under dependence is again GEV, as already addressed in Remark 2.9. This result about the moving maximum process is well-known and $\theta = 1/(k+1)$ is indeed its extremal index Beirlant et al. (2004, Example 10.5).

Finally, notice that any rate function $r_n = n/\alpha$ for $\alpha > 0$ could have been used to obtain $\delta_n(u^{1/(n/\alpha)}) \to u^{\alpha/(k+1)}$. Because $\lim_{n\to\infty} r_{\lceil nt\rceil}/r_n = t$ for any $\alpha > 0$, Theorem 2.10 implies that changing the rate function will only lead to a location-scale transform of the limiting distribution but will not change its functional form otherwise. This observation is in line with the fact that the changed rate $r_n = n/\alpha$ also influences the utilized normalizing constants, leading to the location-scale transform of the limit.

Remark 2.13
Although it is a limit of rescaled copula diagonals, D in Theorem 2.10 is not necessarily

a copula diagonal itself. Indeed, by the Fréchet–Hoeffding inequality Nelsen (2006, Theorem 2.10.12), any copula diagonal must satisfy that for all $u \in (0,1)$, $\delta_n(u) \leq u$. However, the limiting D of the moving maximum process in Example 2.12 satisfies $D(u) = u^\theta > u$ for $u \in (0,1)$ whenever $\theta < 1$.

3 Sequences with asymptotic power diagonals

An appealing feature of the moving maximum process in Example 2.12 is that the limiting distribution of normalized maxima is still generalized extreme value. In this section, we explore the consequences of the results in Section 2 for a broad class of sequences which behave similarly in the sense that the weak limit of suitably normalized maxima is GEV because the diagonal power distortion converges to a power. We first formalize this property in a definition.

Definition 3.1

A sequence of identically distributed random variables $X_1, X_2, \ldots \sim F$ with continuous marginal distribution function F has a *power diagonal* if, for all $n \in \mathbb{N}$ and $u \in (0,1)$, $\delta_n(u) = u^{\eta_n}$ for some sequence (η_n) such that $\eta_n \to \infty$ as $n \to \infty$. And $X_1, X_2, \ldots \sim F$ has an *asymptotic power diagonal with index* $\theta > 0$ if there exists a rate function $r : \mathbb{N} \to (0,\infty)$ with $r_n \to \infty$ as $n \to \infty$ and such that the diagonal power distortion satisfies $D_n^r(u) \to u^\theta$ for all $u \in [0,1]$.

Before proceeding, note that a sequence with a power diagonal necessarily has an asymptotic power diagonal; it suffices to set $r_n = \eta_n$ as then $D_n^{r_n}(u) = u$ for all $u \in [0,1]$. Furthermore, $\eta_n \to \infty$ and $r_n \to \infty$ as $n \to \infty$ are required in view of Proposition 2.5, so that all possible limiting distributions of normalized maxima can be characterized. Also, any copula is bounded above by the comonotone copula Nelsen (2006, Theorem 2.10.12), so that $D_n^r(u) \leq u^{1/r_n}$ for all $u \in (0,1)$. Given that $u^{1/r_n} \to 1$ as $n \to \infty$ if $r_n \to \infty$, any θ such that $D_n^r(u) \to u^\theta$, $u \in (0,1)$, must satisfy $\theta \geq 0$. We excluded the case $\theta = 0$ in Definition 3.1 because the limiting distortion D would then be degenerate.

If $X_1, X_2, \ldots \sim F$ has an asymptotic power diagonal and its marginal distribution function satisfies $F \in \mathrm{MDA}(H)$, Theorem 2.2 (i) guarantees that $(M_n - d^*_{\lceil r_n \rceil})/c^*_{\lceil r_n \rceil} \xrightarrow{\mathrm{d}} H^\theta$, where (d_n^*) and (c_n^*) are the normalizing sequences from the iid case, i.e. such that $F^n(c_n^* x + d_n^*) \to H(x)$ as $n \to \infty$ for any $x \in \mathbb{R}$. Because $H \in \mathrm{GEV}$, H^θ is generalized extreme value with the same shape parameter as H, as detailed in Remark 2.9. The latter also implies that if $r_n/n \to 0$ or $r_n/n \to \infty$, the weak limit of $(M_n - d_n^*)/c_n^*$ is degenerate.

If $X_1, X_2, \ldots \sim F$ has an asymptotic power diagonal but not necessarily $F \in \mathrm{MDA}(H)$, any non-degenerate G in (1) is still GEV by Theorem 2.10 as long as δ_n is strictly increasing for each n and $r_{\lceil tn \rceil}/r_n \to \lambda(t)$ for all $t > 0$ and some bijection $\lambda \colon (0,\infty) \to (0,\infty)$.

We discuss sequences with power diagonals in Section 3.1 and relate to sequences that satisfy the so-called distributional mixing condition in Section 3.2.

3.1 Sequences with power diagonals

Apart from an iid sequence where C_n is the independence copula $\Pi_n(u_1, \ldots, u_n) = \prod_{i=1}^n u_i$ with $\delta_n(u) = u^n$, $u \in (0, 1)$, an example of a sequence with a power diagonal is the moving maximum process in Example 2.12. We can generalize it if we notice that the copula C_n in the latter example is in fact extreme value. From, e.g. Huang (1992), Beirlant et al. (2004), and Gudendorf and Segers (2010), this means that C_n is of the form $C_n(u_1, \ldots, u_n) = \exp[-\ell_n\{-\log(u_1), \ldots, -\log(u_n)\}]$ for all $u_1, \ldots, u_n \in (0, 1)$, where $\ell_n : [0, \infty)^n \to [0, \infty)$ is a *stable tail dependence function (stdf)*, i.e. a map which is homogeneous of order one and has further analytical properties identified in Ressel (2013). This motivates the following definition.

Definition 3.2
We call any sequence $(X_i, i \in \mathbb{N})$ of identically distributed random variables with common continuous marginal distribution F a *meta-extreme sequence*, if the copula C_n of $(X_1, \ldots, X_n)$ is an extreme value copula for any $n \in \mathbb{N}$.

Because any stdf ℓ_n is homogeneous of order one, the diagonal of an extreme value copula satisfies $\delta_n(u) = \exp\{(\log u)\ell_n(1, \ldots, 1)\} = u^{\eta_n}$, where $\eta_n = \ell_n(1, \ldots, 1)$ is the extremal coefficient of Smith Smith (1990); see also Falk (2019, Chapter 2.3). This means that meta-extreme sequences have a power diagonal provided that $\eta_n \to \infty$ as $n \to \infty$. Example 2.7 shows that this latter condition is indeed a restriction; the Cuadras–Augé copula appearing therein is extreme value but such that η_n has a finite limit as $n \to \infty$.

By the characterization of multivariate max-stable distributions in terms of D-norms described in Falk (2019, Section 2.3, Theorem 2.3.3.) and the Takahashi characterization (Falk (2019, Corollary 1.3.2)), we have $\eta_n = n$ for all $n \in \mathbb{N}$ if and only if the elements of the meta-extreme sequence (X_i) are iid. Next, we discuss two specific examples of meta-extreme sequences.

Example 3.3
Exchangeable meta-extreme sequences are fully characterized in Mai (2019), where it is shown on p. 167 that the naïve, bottom-up construction approach of selecting an exchangeable STDF $\tilde{\ell}$ in a fixed dimension n in the hope that there exists an exchangeable meta-extreme sequence with $\ell_n = \tilde{\ell}$ fails in general. Generalizing de Haan's spectral representation Haan (1984), it is shown in Mai (2019) that (X_i) is an exchangeable meta-extreme sequence if and only if there exists $b \in [0, 1]$ and an exchangeable sequence (W_i) of non-negative random variables with unit mean so that for each $n \geq 2$ and $t_1, \ldots, t_n \geq 0$, $\ell_n(t_1, \ldots, t_n) = b \sum_{j=1}^n t_j + (1 - b)\mathbb{E}[\max_{1 \leq i \leq n}(t_i W_i)]$. Clearly, $b = 1$ corresponds to an iid sequence (X_i). When $b > 0$, the sequence (X_i) has a power diagonal, while when $b \in [0, 1)$, the behavior of the extremal coefficient $\eta_n = \ell_n(1, \ldots, 1)$ depends on (W_i). Also when $b = 0$, $\ell_n(t_1, \ldots, t_n) = \|(t_1, \ldots, t_n)\|_D$, where $\| \cdot \|_D$ is the so-called D-norm generated by $(W_1, \ldots, W_n)$ Falk (2019, Lemma 1.1.3).

The special case when $b = 0$ and (W_i) is an iid sequence is treated in Mai (2018), where it is also shown how to construct meta-extreme sequences starting from the distribution function F_W of W_1, provided that $F_W(0) < 1$. This construction first defines a stochastic process H_t, $t \geq 0$, by $H_t = -\ln[\prod_{k=1}^{\infty} F_W\{(\varepsilon_1 + \cdots + \varepsilon_k)/t-\}]$, where $F_W(w-)$ is the left limit of F_W at w and (ε_i) is an iid sequence of unit exponentials. The meta-extreme sequence (X_i) with univariate margin F is then obtained by setting $X_i = F^{-1}(e^{-Y_i})$, where $Y_i = \inf\{t > 0 : H_t > \xi_i\}$ and (ξ_i) is an independent copy of (ε_i). From Mai (2018, Lemma 2 and 3), (X_i) has a power diagonal if and only if $\inf\{t : F_W(t) = 1\} = \infty$.

Specific choices for F_W lead to meta-extreme sequences with well-known stdfs. From Mai (2018, Example 1 and Example 2) and Belzile and Nešlehová (2017, Section 6), we can take W to be scaled gamma with parameters $\alpha > 0$, $\rho > -\alpha$ and density $f_W(w) = \{|1/\rho|/\Gamma(\alpha)\}a^{-\alpha/\rho}x^{\alpha/\rho-1}e^{-(w/a)^{1/\rho}}$, $w > 0$, where $\Gamma(\cdot)$ is Euler's gamma function and $a = \Gamma(\alpha)/\Gamma(\rho+\alpha)$. Setting $\rho = 1$ yields the (symmetric) Coles–Tawn extremal Dirichlet sequence, while when $\alpha = 1$, then $\rho > 0$ and $\rho \in (-1, 0)$ lead to the negative and positive logistic sequences, respectively. The positive logistic (or Gumbel–Hougaard) stdf is usually parametrized in terms of $\theta = -1/\rho$ and is given, for for all $t_1, \ldots, t_n \geq 0$, by $\ell_{n,\mathrm{Gu}}(t_1, \ldots, t_n) = (t_1^\theta + \cdots + t_n^\theta)^{1/\theta}$.

Example 3.4

Another approach to construct meta-extreme sequences is through a more general class of simple max-stable processes on $[0, 1]$, where we follow Haan and Ferreira (2006, Chapter 9). Let $C[0, 1]$ denote the space of continuous functions on $[0, 1]$ equipped with the supremum norm and $C^+[0, 1]$ its subspace of strictly positive functions. The process S on $C^+[0, 1]$ is *simple max-stable* if for all $t \in [0, 1]$, $\mathbb{P}(S(t) \leq z) = e^{-1/z}$ for all $z \geq 0$ and if for all $k \in \mathbb{N}$, $(1/k) \vee_{i=1}^{k} S_i \stackrel{\mathrm{d}}{=} S$, where $S_1, S_2, \ldots$ are iid copies of S, $\vee$ is the pointwise maximum operator, and $\stackrel{\mathrm{d}}{=}$ denotes equality in distribution.

Starting with a simple max-stable process and a strictly increasing sequence (t_i) in $[0, 1]$, e.g. $t_i = 1 - 1/(i + 1)$ for $i \geq 1$, we can define (X_i) by $X_i = F^{-1}(e^{-1/S(t_i)})$. Because $S(t_i)$ is standard Fréchet, $X_i \sim F$. The invariance principle Schweizer and Sklar (1983, Theorem 6.5.6) implies that for each $n \in \mathbb{N}$, the copula C_n of $(X_1, \ldots, X_n)$ is the same as that of $(S(t_1), \ldots, S(t_n))$ and consequently extreme value, so that (X_i) is a meta-extreme sequence. Following Haan (1984), S admits the stochastic representation $S \stackrel{\mathrm{d}}{=} \vee_{k\geq 1}(\xi_k W_k)$, where (ξ_k) is an enumeration of points of a Poisson point process ξ on $[0, 1]$ with intensity $t^{-2}\,\mathrm{d}t$ and which is independent of the iid copies $W_1, W_2, \ldots$ of a stochastic process W on $C^+[0, 1]$ such that $\mathbb{E}[W(t)] = 1$ for all $t \in [0, 1]$ and $\mathbb{E}\big[\sup_{t\in[0,1]} W(t)\big] < \infty$ Haan and Ferreira (2006, Corollary 9.4.5). This implies that the stdf ℓ_n of C_n satisfies $\ell_n(x_1, \ldots, x_n) = \mathbb{E}[\max_{1\leq i\leq n} x_i W(t_i)]$ for all $x_1, \ldots, x_n \geq 0$, so that $\ell_n(x_1, \ldots, x_n)$ is the D-norm generated by $(W(t_1), \ldots, W(t_n))$ Falk (2019, Lemma 1.1.3). Note however that the meta-extreme sequence (X_i) constructed this way cannot have a power diagonal because $\eta_n \not\to \infty$. Indeed, for each $n \in \mathbb{N}$, $\eta_n = \ell_n(1, \ldots, 1) = \mathbb{E}\{\max_{1\leq i\leq n} W(x_i)\} \leq \mathbb{E}\{\sup_{t\in[0,1]} W(t)\} < \infty$. Since $\eta_n \geq 1$ and the sequence (η_n) is non-decreasing, there exists $\varrho > 0$ so that $\eta_n \to \varrho$. Proposition 2.5 (i) thus implies that $M_n = \max(X_1, \ldots, X_n)$ converges weakly to F^ϱ.

The next example shows that power diagonals arise not only from extreme value copulas, demonstrating that the previous discussion is not limited to meta-extreme sequences.

Example 3.5
Consider a sequence of iid bivariate random vectors (U_i, V_i), $i \in \{0, 1, \dots\}$ such that $(U_i, V_i) \sim C_{\mathrm{FN}}$ where $C_{\mathrm{FN}}(u, v) = \min\{u, v, (u^2 + v^2)/2\}$, $u, v \in [0, 1]$. This copula is a special case of the so-called bivariate diagonal copulas introduced and studied in Fredricks and Nelsen (1997). For each $i \in \mathbb{N}$, set $Y_i = \max(V_{i-1}^2, U_i^2)$. This construction is similar to the (finite-dimensional) product type copula construction of Liebscher (2008) and to the copulas discussed in Mazo et al. (2015). Given that V_{i-1} and U_i are independent, the random variables Y_i are standard uniform and the joint distribution function of $(Y_1, \dots, Y_n)$ is a copula given by $C_n(u_1, \dots, u_n) = \sqrt{u_1} \times \prod_{j=1}^{n-1} C_{\mathrm{FN}}(\sqrt{u_j}, \sqrt{u_{j+1}}) \times \sqrt{u_n}$, $u_1, \dots, u_n \in [0, 1]$. Clearly, $\delta_n(u) = u^n$ for all $u \in [0, 1]$, i.e. that C_n has the same power diagonal as Π_n. This means that if we take any continuous distribution function F, the maxima of the sequence $X_i = F^{-1}(Y_i)$, $i \in \mathbb{N}$, behave in the same way as the maxima of the associated iid sequence. Yet, C_{FN} and hence also C_n is not an extreme value copula, so that (X_i) is not meta-extreme.

3.2 Stationary sequences with short-range extremal dependence

We now show that asymptotic power diagonals are inherent to strictly stationary sequences with a certain form of asymptotic independence in the tail. Conditions that formalize the latter property are regularly used to study maxima of strictly stationary time series, see e.g. Leadbetter et al. (1983, Chapter 3), or Beirlant et al. (2004, Chapter 10) and Embrechts et al. (1997, Chapter 4).

Definition 3.6
A strictly stationary sequence $X_1, X_2, \dots \sim F$ is said to satisfy:

(i) The *distibutional mixing condition* $\mathcal{D}(u_n)$ if for any integers p, q, n and indices $1 \leq i_1 < \dots < i_p < j_1 < \dots < j_q \leq n$ such that $j_1 - i_p \geq s$, $|\mathbb{P}(\max_{i \in A \cup B} X_i \leq u_n) - \mathbb{P}(\max_{i \in A} X_i \leq u_n)\mathbb{P}(\max_{i \in B} X_i \leq u_n)| \leq \alpha(n, s)$, where $A = \{i_1, \dots, i_p\}$, $B = \{j_1, \dots, j_q\}$, and $\alpha(n, s_n) \to 0$ as $n \to \infty$ for some positive integer sequence $s_n = o(n)$.

(ii) The *anticlustering condition* $\mathcal{D}'(u_n)$ if $\lim_{k \to \infty} \limsup_{n \to \infty} n \sum_{j=2}^{\lfloor n/k \rfloor} \mathbb{P}(X_1 > u_n, X_j > u_n) = 0$.

The following result, relating $\mathcal{D}(u_n)$ and $\mathcal{D}'(u_n)$ to the behavior of δ_n, is a consequence of Theorem 2.2 (ii) and limit theorems for maxima of stationary series in Leadbetter (1974), Leadbetter (1983), and Leadbetter et al. (1983).

Corollary 3.7
Let $X_1, X_2, \dots \sim F$ be a strictly stationary sequence. Suppose that F is continuous and satisfies $F \in \mathrm{MDA}(H)$ with normalizing sequences (c_n), $c_n > 0$ and (d_n).

(i) If the condition $\mathcal{D}(u_n)$ is satisfied with $u_n = c_n x + d_n$ for each x such that $H(x) > 0$ and there exists a $u \in (0, 1)$ such that $D_n^r(u) \to \gamma$ where $\gamma \in (0, 1)$ and $r_n = n$ for each

$n \in \mathbb{N}$, then the sequence $X_1, X_2, \ldots \sim F$ has an asymptotic power diagonal with index $\theta \in (0,1]$ and θ is the extremal index of (X_i), meaning that $(M_n - d_n)/c_n \xrightarrow{\mathrm{d}} H^\theta$.

(ii) If the conditions $\mathcal{D}(u_n)$ and $\mathcal{D}'(u_n)$ hold with $u_n = c_n x + d_n$ for each $x \in \mathbb{R}$, then the sequence $X_1, X_2, \ldots \sim F$ has an asymptotic power diagonal with index $\theta = 1$ and the rate may be chosen as $r_n = n$ for each $n \in \mathbb{N}$.

Remark 3.8

When the rate is $r(n) = n$, $n \in \mathbb{N}$, Lipschitz continuity of copulas implies that $|D_n^r(v_n) - D_n^r(u)| \to 0$ as $n \to \infty$ whenever $v_n \to u$ where $u \in (0,1)$, as argued in the proof of Theorem 2.2 (ii). Consequently, provided that $F \in \mathrm{MDA}(H)$ with normalizing sequences (c_n), $c_n > 0$ and (d_n), the condition that $D_n^r(u)$ converges for some $u \in (0,1)$ is equivalent to $\mathbb{P}(M_n \le c_n x + d_n)$ converging for some x such that $H(x) \in (0,1)$.

Finally, if $F \in \mathrm{MDA}(H)$ and (X_i) has extremal index $\theta \in (0,1]$, meaning that $(M_n - d_n)/c_n$ converges weakly to H^θ with the same normalizing sequences (c_n) and (d_n) that stabilize the maxima of the associated iid sequence $X_1^*, X_2^*, \ldots \sim F$, then Theorem 2.2 (ii) implies that (X_i) has an asymptotic power diagonal with index θ and rate $r_n = n$, $n \in \mathbb{N}$. Thus, for large enough n, $D_n^r(u) = \delta_n(u^{1/n}) \approx D(u) = u^\theta$, so that $\delta_n(u) \approx u^{\theta n}$, where $u^{\theta n}$ may be interpreted as the copula diagonal of θn independent variables.

The moving maximum process in Example 2.12 is a case in point where Corollary 3.7 (i) applies. The $\mathcal{D}(u_n)$ condition holds and the extremal index equals $\theta = 1/(k+1)$, $k \ge 0$ Beirlant et al. (2004, Example 10.5). The fact that the moving maximum process has an asymptotic power diagonal could thus have been alternatively obtained from Corollary 3.7 (i).

As we illustrate in the next example, Corollary 3.7 reveals facts about the limiting behavior of copula diagonals of stationary sequences whose copulas may be intractable or not even explicit. This is the case for most classical time series models, such as ARMA or GARCH processes (although we note that copulas have also been used explicitly to construct time series models, examples of the latter are Markov processes Darsow et al. (1992), models that use neural networks to capture cross-sectional dependence Hofert et al. (2022), or vine copula models Nagler et al. (2022)).

Example 3.9

Consider a stationary Gaussian sequence (Y_i), meaning that all finite-dimensional distributions are multivariate normal. From Sklar's theorem, the copula C_n of $(Y_1, \ldots, Y_n)$ is a Gaussian copula. Although the univariate normal distribution is in the maximum domain of attraction of the Gumbel extreme value distribution Λ, the tail properties of δ_n may not be easy to investigate. For example, take the Gaussian AR(1) process given by $Y_n = \phi Y_{n-1} + Z_n$ for $n \in \mathbb{Z}$, where $\phi \in (-1,1)$ and (Z_n) is an iid sequence with $Z_n \sim \mathrm{N}(0, \sigma^2)$. It is easily shown that $(Y_1, \ldots, Y_n) \sim \mathrm{N}(\mathbf{0}, \Sigma_n)$ with $(\Sigma_n)_{ij} = (\sigma^2 \phi^{|i-j|})/(1-\phi^2)$. Denote by Φ and Φ_{Σ_n} the distribution function of the $\mathrm{N}(0,1)$ and $\mathrm{N}(\mathbf{0}, \Sigma_n)$, respectively. The diagonal of C_n is then $\delta_n(u) = \Phi_{\Sigma_n}\big(\sigma\Phi^{-1}(u)/\sqrt{1-\phi^2}, \ldots, \sigma\Phi^{-1}(u)/\sqrt{1-\phi^2}\big)$, $u \in (0,1)$.

It is not easy to investigate the limiting behavior of $\delta_n(u^{1/r_n})$ for some suitable rate r.

However, for a stationary Gaussian sequence (V_i), the so-called Berman condition

$$\lim_{n\to\infty} \operatorname{cov}(V_1, V_n)\ln(n) = 0 \tag{3}$$

ensures that the $\mathcal{D}(u_n)$ and $\mathcal{D}'(u_n)$ conditions hold for any $u_n = c_n^* x + d_n^*$ and $x \in \mathbb{R}$, where (c_n^*) and (d_n^*) are the normalizing sequences of the associated iid sequence Embrechts et al. (1997, Lemma 4.4.7). Corollary 3.7 (ii) thus ensures that (V_i) has an asymptotic power diagonal with index $\theta = 1$ and rate $r_n = n$. This is the case for the above AR(1) process, because $\operatorname{cov}(Y_1, Y_n) = \sigma^2\phi^n/(1-\phi^2)$ Embrechts et al. (1997, Example 4.4.9 and Example 7.1.1).

Our framework now allows us to derive the limiting behavior of M_n for any sequence (X_i) of the form $X_i = F^{-1}[\Phi\{(Y_i - \mu)/\sigma\}]$ where F is a continuous distribution function and (Y_i) is a stationary Gaussian sequence with $Y_i \sim \mathrm{N}(\mu, \sigma^2)$ that satisfies (3). Indeed, because the copulas of $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ are identical, $D_n^r(u) \to u$ as $n \to \infty$ with rate $r_n = n$. If $F \in \mathrm{MDA}(H)$, Theorem 2.2 (i) implies that (1) holds with $G = H$ and the normalizing sequences of the associated iid sequence.

Example 3.9 can be generalized to a number of stochastic processes for which the extremal index is known to be non-zero. This includes the uniform autoregressive process or the GARCH(1,1) process; see McNeil et al. (2015, p. 142) or Ferreira (2018, Section 3) for a list of suitable examples.

While the distributional mixing condition rules out long-range dependence in the tail, this is not the case for Theorem 2.2 and Theorem 2.10. The final result in this section characterizes sequences with asymptotic power diagonals for which the results in Section 2 imply that the limit in (1) is generalized extreme value and yet the $\mathcal{D}(u_n)$ condition is violated.

Proposition 3.10
Let $X_1, X_2, \ldots \sim F$ be an exchangeable sequence with an asymptotic power diagonal with index $\theta > 0$ and rate $r : \mathbb{N} \to (0,\infty)$, $\lim_{n\to\infty} r(n) = \infty$. Assume that there exists a bijection $\lambda : (0,\infty) \to (0,\infty)$ such that $r_{\lceil nt \rceil}/r_n \to \lambda(t)$ for all $t > 0$ and suppose that λ restricted to $(0,1)$ is not linear, i.e., not of the form $\lambda(t) = \alpha t$, $t \in (0,1)$, $\alpha > 0$. Suppose further that δ_n is strictly increasing for all $n \in \mathbb{N}$ and that (1) holds for some normalizing sequences (c_n), $c_n > 0$, and (d_n) and non-degenerate distribution G. Then $G \in \mathrm{GEV}$ while $\mathcal{D}(u_n)$ is violated for all thresholds of the form $u_n = c_n x + d_n$ with x such that $G(x) \in (0,1)$.

Prime examples of sequences which meet the conditions of Proposition 3.10 are exchangeable sequences with power diagonals, where the power η_n grows slower than n. For example, we can consider the meta-extreme sequence (X_i) from Example 3.3 with continuous margin $F \in \mathrm{MDA}(H)$ for some $H \in \mathrm{GEV}$ and the logistic stdf $\ell_{n,\mathrm{Gu}}$ with $\theta > 1$. As explained in Section 3.1, the Gumbel–Hougaard copula C_n has a power diagonal with $\eta_n = \ell_{n,\mathrm{Gu}}(1, \ldots, 1) = n^{1/\theta}$. We then have $r_n = \eta_n$ and $\lambda(t) = t^{1/\theta}$, which is not linear.

4 Sequences with non-GEV limits

In this section, we study sequences for which the limiting distribution in (1) is no longer generalized extreme value. We first investigate meta-Archimax sequences in Section 4.1 and two generalizations in Section 4.2, one to arbitrary exchangeable sequences and one to mixtures of not necessarily exchangeable sequences.

4.1 Meta-Archimax sequences

In this section, we treat sequences defined as follows.

Definition 4.1
A continuous function $\psi : [0, \infty) \to [0, 1]$ which satisfies $\psi(0) = 1$, $\psi(t) \to 0$ as $t \to \infty$ and which is strictly decreasing on $[0, \inf\{x : \psi(x) = 0\}]$ is called an *(Archimedean) generator*. A sequence (X_i) of identically distributed random variables is called a *meta-Archimax sequence* with Archimedean generator ψ and a sequence (ℓ_n) of n-variate stdfs, if, for each $n \geq 2$, the copula C_n of $(X_1, \ldots, X_n)$ is an Archimax copula with generator ψ and stdf ℓ_n, i.e.

$$C_n(u_1, \ldots, u_n) = \psi[\ell_n\{\psi^{-1}(u_1), \ldots, \psi^{-1}(u_n)\}], \tag{4}$$

for all $u_1, \ldots, u_n \in [0, 1]$. If, for all $n \geq 1$ and $x_1, \ldots, x_n \in [0, \infty)$, $\ell_n(x_1, \ldots, x_n) = x_1 + \ldots + x_n$, the sequence is called *meta-Archimedean*.

We emphasize that the generator of a meta-Archimax sequence does not depend on n; its inverse ψ^{-1} is well-defined on $(0, 1]$ and $\psi^{-1}(0) := \inf\{x : \psi(x) = 0\}$ by convention. Table 1 provides several well-known parametric families of generators. For additional examples, see e.g. Nelsen (2006, Chapter 4.6) and Hofert (2011); these generators can be further transformed to obtain richer classes of models as in Charpentier and Segers (2009, Table 2) or Hofert and Scherer (2011). When we say that (X_i) is a meta-Archimax sequence with generator ψ and stdfs (ℓ_n), we implicitly assume that C_n in (4) is a bona-fide copula for each n and that (ℓ_n) satisfies $\ell_{n+1}(x_1, \ldots, x_n, 0) = \ell_n(x_1, \ldots, x_n)$ for all $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in [0, \infty)$, where $\ell_1(x) = x$ by convention. The question when a given ψ and (ℓ_n) give rise to a sequence of Archimax copulas can only be answered in special cases; we elaborate on this in Example 4.2 and Example 4.3.

Archimax and notably Archimedean copulas have been studied extensively, viz. Capéraà et al. (2000), Charpentier et al. (2014), Chatelain et al. (2020), McNeil and Nešlehová (2009), and Nelsen (2006). It is easily seen that when $\psi(t) = e^{-t}$ for all $t \geq 0$, the meta-Archimax sequence with generator ψ reduces to a meta-extreme sequence with stdfs (ℓ_n) that we investigated in Section 3.1. Example 4.2 and Example 4.3 show explicit constructions of meta-Archimedean and meta-Archimax sequences for arbitrary generators, relating them to scale mixtures of certain iid or meta-extreme sequences.

Example 4.2
From Kimberling (1974a) it is well-known that ψ is an Archimedean generator of a meta-

Copula	Generator $\psi(t)$	Inverse $\psi^{-1}(u)$	ρ	$-\psi'(0)$
Independence	$\exp(-t)$	$-\log(u)$	1	1
Ali-Mikhail-Haq ($\theta \in (0,1)$)	$\frac{1-\theta}{\exp(t)-\theta}$	$\log\left(\frac{1-\theta(1-u)}{u}\right)$	1	$\frac{1}{1-\theta}$
Clayton ($\theta > 0$)	$(1+t)^{-1/\theta}$	$u^{-\theta}-1$	1	$1/\theta$
Frank ($\theta > 0$)	$-\frac{1}{\theta}\log\left(1+\exp(-t)\left(\mathrm{e}^{-\theta}-1\right)\right)$	$-\log\left(\frac{\exp(-\theta u)-1}{\exp(-\theta)-1}\right)$	1	$\frac{\mathrm{e}^{\theta}-1}{\theta}$
Example 4.11	$1/\left(t(1+1/t)^{1+t}\right)$	(no closed form)	1	∞
Gumbel-Hougaard ($\theta > 1$)	$\exp\left(-t^{1/\theta}\right)$	$(-\log(u))^{\theta}$	$1/\theta$	∞
Joe ($\theta > 1$)	$1-(1-\exp(-t))^{1/\theta}$	$-\log\left(1-(1-u)^{\theta}\right)$	$1/\theta$	∞

Table 1 Completely monotone generators, their inverses, coefficients of regular variation ρ such that $1-\psi(1/x) \in \mathrm{RV}_{-\rho}$ and negative right-hand side generator derivatives at 0 for selected Archimedean copulas. Note that $-\psi'(0) < \infty$ implies $\rho = 1$ as discussed in Example 4.9.

Archimedean sequence if and only if it is completely monotone, i.e. differentiable on $(0,\infty)$ of all orders with k-th derivative satisfying $(-1)^k\psi^{(k)}(\cdot) \geq 0$. This result allows us to construct an arbitrary meta-Archimedean sequence $X_1, X_2, \ldots \sim F$ explicitly as follows. The Bernstein–Widder theorem Feller (1971, p. 439) implies that a completely montone ψ must be a Laplace–Stieltjes transform of a positive random variable V (also called *frailty*), i.e. $\psi(t) = \mathbb{E}[e^{-tV}]$, $t \geq 0$. Let (E_i) be a sequence of iid unit exponential random variables independent of V. As observed in Marshall and Olkin (1988), the survival copula of the multiplicative hazard (or frailty) model $(E_1/V, \ldots, E_n/V)$ is an Archimedean copula with generator ψ, viz. $C_n(u_1,\ldots,u_n) = \psi\{\psi^{-1}(u_1)+\cdots+\psi^{-1}(u_n)\}$, $u_1,\ldots,u_n \in (0,1)$. Consequently, the sequence (Y_i) given by $Y_i = V/E_i$ is then meta-Archimedean with generator ψ and a continuous univariate margin given by $\psi(1/x)$ for $x > 0$ and by 0 otherwise. To obtain a meta-Archimedean sequence with generator ψ and an arbitrary univariate margin F, it suffices to set $X_i = F^{-1}\{\psi(E_i/V)\}$ for all $i \in \mathbb{N}$.

Example 4.3

When ψ is a completely monotone Archimedean generator, Charpentier et al. (2014) show that (4) is a bona-fide copula for any $n \geq 2$ and any n-variate stdf ℓ_n. Complete monotonicity of ψ is not necessary for certain fixed sequences of n-variate stdfs McNeil and Nešlehová (2009) and Charpentier et al. (2014), but when it holds, it again allows us to use the Bernstein–Widder theorem to construct a meta-Archimax sequence explicitly, as follows.

As in Example 4.2, let V be a positive random variable with Laplace–Stieltjes transform ψ. Let also (Z_i) be a meta-extreme sequence independent of V, with unit Fréchet margins and stdfs ℓ_n, as in Definition 3.2. Define the sequence (Y_i) via $Y_i = VZ_i$ for all $i \geq 1$. Charpentier et al. (2014, Remark 3.2) implies that (Y_i) is meta-Archimax with generator ψ and stdfs (ℓ_n). Its univariate margin is again given by $\psi(1/x)$ for $x > 0$ and by 0 otherwise. To generalize this construction to a meta-Archimax sequence with an arbitrary margin F, we set $X_i = F^{-1}\{\psi(E_i/V)\}$ for all $i \in \mathbb{N}$, as in Example 4.2.

We now describe the limiting behavior of maxima of meta-Archimax sequences using the theory we developed in Section 2. To this end, observe that the diagonal of an Archimax copula is of the form $\delta_n(u) = \psi\{\psi^{-1}(u)\eta_n\}$, where $\eta_n = \ell_n(1, \dots, 1)$ is Smith's extremal coefficient as in Section 3.1. Therefore, the asymptotic properties of δ_n will depend on the behavior of η_n and of the Archimedean generator, notably its regular variation. We recall that a measurable function $f > 0$ is regularly varying at ∞ with index $\rho \in \mathbb{R}$, in notation $f \in \mathrm{RV}_\rho$, if it satisfies $\lim_{x\to\infty} f(tx)/f(x) = t^\rho$ for all $t > 0$, see e.g. Bingham et al. (1987, Chapter 2).

Theorem 4.4
Let (X_i) be a meta-Archimax sequence with generator ψ and stdfs (ℓ_n). For each $n \geq 1$, set $\eta_n = \ell_n(1, \dots, 1)$ and define the rate $r : \mathbb{N} \to (0, \infty)$ by $r(n) = r_n = 1/(1 - \psi(1/\eta_n))$, $n \geq 1$. Furthermore assume:

(i) $\eta_n \to \infty$ as $n \to \infty$;

(ii) $1 - \psi(1/\cdot) \in \mathrm{RV}_{-\rho}$ for $\rho \in (0, 1]$.

Then, for all $u \in [0, 1]$, the diagonal power distortion $D^r_n(u) = \psi\{\eta_n\psi^{-1}(u^{1/r_n})\}$ converges to $D(u) = \psi\{(-\log u)^{1/\rho}\}$ as $n \to \infty$. If also the univariate marginal distribution F of (X_i) satisfies $F \in \mathrm{MDA}(H)$ for some $H \in \mathrm{GEV}$ and normalizing sequences (c^*_n), $c^*_n > 0$, and (d^*_n), i.e. $F^n(c^*_n x + d^*_n) \to H(x)$, $x \in \mathbb{R}$, then (1) holds with $G = D \circ H$ and normalizing sequences given by $c_n = c^*_{\lceil r_n \rceil}$ and $d_n = d^*_{\lceil r_n \rceil}$.

The assumptions in Theorem 4.4 are not restrictive. We already discussed assumption (i) in Section 3: If η_n had a finite limit as in Example 2.7, we would be in the scope of Proposition 2.5 and the possible limits in (1) would be too broad. Assumption (ii) is satisfied by nearly all known Archimedean generators, including all listed in Table 1, see Charpentier and Segers (2009) and Larsson and G. Nešlehová (2011).

Remark 4.5
The limiting distribution G obtained in Theorem 4.4 can also be expressed in a different way. To see this, first write $H = H_{\xi,\mu,\sigma}$ for some $\xi, \mu \in \mathbb{R}$ and $\sigma > 0$. Using the fact that for all $x \in \mathbb{R}$, $\{-\log H_{\xi,\mu,\sigma}(x)\}^{1/\rho} = -\log H_{\rho\xi,\mu,\rho\sigma}(x)$, we also have, for all $x \in \mathbb{R}$,

$$G(x) = \psi\big\{-\log H_{\rho\xi,\mu,\rho\sigma}(x)\big\}. \tag{5}$$

Viewing $\psi(t^{1/\rho})$, $t \in [0, \infty)$, as an outer power transformation of the Archimedean generator ψ when $\rho < 1$, the probabilistic interpretation in Hofert (2011) does not apply here, because $\psi(t^{1/\rho})$ is a valid Archimedean generator only for $\rho \geq 1$ Nelsen (2006, Theorem 4.5.1).

Remark 4.6
The alternative expression (5) also allows us to connect Theorem 4.4 with the results of Wüthrich Wüthrich (2004), who investigated maxima of meta-Archimedean sequences. The assumptions of Proposition 5.6 in the latter paper are the same as in Theorem 4.4, and the limit is precisely as in (5), although the normalizing constants are different. Rather than

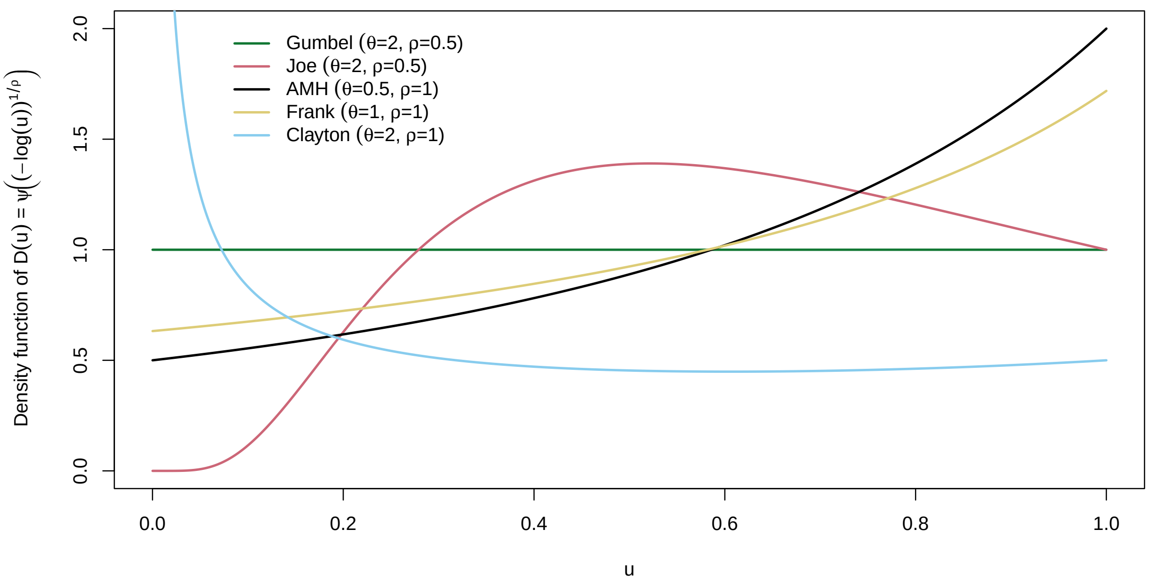

Figure 1 Densities of $D(u) = \psi\{(-\log u)^{1/\rho}\}$ for different Archimedean generators from Table 1.

investigating copula diagonals as done here, the approach in Wüthrich (2004) relates M_n to the maximum $\widetilde{M}_n$ of iid observations drawn from the univariate distribution function $\exp\{-\psi^{-1} \circ F\}$. While the normalizing sequences used in Theorem 4.4 are subsequences of (c_n^*) and (d_n^*) used to stabilize the iid maximum M_n^*, Wüthrich (2004) uses the normalizing constants of $\widetilde{M}_n$. Clearly, the two sets of sequences must be related through the convergence to types theorem Resnick (1987, Proposition 0.2). When $F \in \mathrm{MDA}(H_{\xi,\mu,\sigma})$ with $\xi > 0$, it is also possible to see this directly. In Wüthrich (2004), $c_n = F^{-1}[\psi\{\log n/(n-1)\}]$ and $d_n = 0$, while the approach taken here gives $c_n = c^*_{\lceil r_n \rceil} = F^{-1}\{(\lceil r_n \rceil - 1)/\lceil r_n \rceil\}$ and $d_n = d^*_{\lceil r_n \rceil} = 0$. Assumption (ii) of Theorem 4.4 implies that $1 - \psi\{\log n/(n-1)\} \approx 1 - \psi(1/n)$ so that if we set n^* to be an integer such that $(n^* - 1)/n^* \approx \psi\{\log n/(n-1)\}$, we obtain that $n^* \approx \lceil 1/\{1 - \psi(1/n)\}\rceil$. As $\eta_n = n$ for a meta-Archimedean sequence, $n^* \approx \lceil r_n \rceil$.

As expected from Theorem 2.2, the distortion function D identified in Theorem 4.4 is a proper distribution function on $[0, 1]$. We formally state this result in the following corollary, while Figure 1 illustrates the density of D for various choices of ψ. As can be seen, the possible density shapes range from being constant to strictly increasing, unimodal or unbounded and thus cover a wide range of possible distortions.

Corollary 4.7

The function $D(u) = \psi\{(-\log u)^{1/\rho}\}$ is a distribution function on $[0, 1]$, with corresponding quantile function $D^{-1}(u) = \mathrm{e}^{-\{\psi^{-1}(u)\}^\rho}$, $u \in [0,1]$, and density

$$d(u) = \frac{-\psi'\{(-\log u)^{1/\rho}\}(-\log u)^{(1-\rho)/\rho}}{\rho u}, \quad u \in (0,1).$$

The values of d at the boundary are given by $d(0) = \lim_{y\to\infty} -\psi'(y)y^{1-\rho}\mathrm{e}^{(y^\rho)}/\rho$ and by $d(1) = -\psi'(0)$ if $\rho = 1$ and $d(1) = \lim_{y\to 0} -\psi'(y)y^{1-\rho}/\rho$ if $\rho \in (0,1)$.

We now draw conclusions from Theorem 4.4 and illustrate that the behavior of η_n controls the speed of convergence while the generator ψ determines the functional form of the limit. We focus on archetypal meta-Archimax sequences with generators that induce different degrees of dependence in the upper tail of the associated Archimedean copula.

Example 4.8

By Chatelain et al. (2020, Lemma 2.4), if the generator ψ satisfies assumption (ii) of Theorem 4.4, the meta-Archimax sequence (X_i) with generator ψ and stdfs (ℓ_n) is meta-extreme if and only if ψ is the Gumbel–Hougaard generator $\psi(t) = \exp(-t^{1/\theta})$ with parameter $\theta \geq 1$; the constant c in Chatelain et al. (2020, Lemma 2.4) can be set to $c = 1$ without loss of generality since the associated Archimax copula remains the same. When $\theta = 1$, ψ is the independence generator $\psi(t) = e^{-t}$ and (X_i) is a meta-extreme sequence with the same stdfs (ℓ_n). When $\theta > 1$, expressing (X_i) as a meta-extreme sequence requires changing the stdfs to $(\ell_n^\star)$, where for all $n \geq 2$ and $x_1, \ldots, x_n \in [0,\infty)$, $\ell_n^\star(x_1,\ldots,x_n) = \ell_n^{1/\theta}(x_1^\theta,\ldots,x_n^\theta)$. When (X_i) is meta-Archimedean, $\ell_n(x_1,\ldots,x_n) = x_1 + \cdots + x_n$ so that $\ell_n^\star$ is the logistic stdf $\ell_{n,\mathrm{Gu}}$ seen at the end of Example 3.3.

Assuming that $\eta_n = \ell_n(1,\ldots,1) \to \infty$ as $n \to \infty$, we can compare Theorem 4.4 to the results in Section 3. To this end, suppose that the univariate margin F of (X_i) satisfies $F \in \mathrm{MDA}(H_{\xi,\mu,\sigma})$ and note that for $\theta \geq 1$, ψ satisfies assumption (ii) of Theorem 4.4 with $\rho = 1/\theta$, viz. Table 1. Set $r_n^\star = \eta_n^\star = \eta_n^{1/\theta}$ with $\eta_n^\star = \ell_n^\star(1,\ldots,1)$ so that $\lim_{n\to\infty} r_n^\star = \infty$. From $\lim_{t\to\infty} t(1-e^{-1/t}) = 1$, the rate r_n from Theorem 4.4 linked to $\eta_n^\star$ satisfies $\lim_{n\to\infty} r_n/r_n^\star = \lim_{n\to\infty}[\eta_n^\star\{1-\psi(1/\eta_n^\star)\}]^{-1} = \lim_{n\to\infty}\{\eta_n^{1/\theta}(1-e^{-1/\eta_n^{1/\theta}})\}^{-1} = 1$. Remark 2.9 shows that the rate in Theorem 4.4 can thus be replaced by $r_n^\star = \eta_n^\star$ used in Section 3, leading to the same limit as in (1). The latter is $\psi\{(-\log H_{\xi,\mu,\sigma})^\theta\} = \exp[-\{(-\log H_{\xi,\mu,\sigma})^\theta\}^{1/\theta}] = H_{\xi,\mu,\sigma}$, which is indeed what we obtained in Section 3.1.

Example 4.9

Let ψ be an Archimedean generator with $-\psi'(0) \in (0,\infty)$; see Table 1 for examples. From Charpentier and Segers (2009, Section 4.3), the Archimedean copula generated by ψ is upper tail independent and assumption (ii) of Theorem 4.4 holds with $\rho = 1$; see also Larsson and G. Nešlehová (2011, Proposition 1). When ψ is also completely monotone, the frailty V in Example 4.2 has a finite mean.

Let (X_i) be a meta-Archimax sequence with generator ψ such that $-\psi'(0) \in (0,\infty)$ and a sequence (ℓ_n) of stdfs with $\eta_n = \ell_n(1,\ldots,1) \to \infty$ as $n \to \infty$. Suppose that the univariate margin F of (X_i) satisfies $F \in \mathrm{MDA}(H_{\xi,\mu,\sigma})$. If we set $\tilde{r}_n = \eta_n$ for all $n \in \mathbb{N}$, we obtain via the formulation of r_n provided in Theorem 4.4 that $\lim_{n\to\infty} r_n/\tilde{r}_n =$

$\lim_{n\to\infty}[\eta_n\{1-\psi(1/\eta_n)\}]^{-1} = \lim_{n\to\infty}[\eta_n\{\psi(0)-\psi(1/\eta_n)\}]^{-1} = -1/\psi'(0)$. Remark 2.9 thus implies that the rate $\tilde{r}_n$ could have been used instead of r_n and that (1) would then hold with $(d^*_{\lceil\eta_n\rceil})$ and $(c^*_{\lceil\eta_n\rceil})$ and a limiting distribution of the form $G(x) = \psi\{-\log H_{\xi,\mu,\sigma}^{-1/\psi'(0)}(x)\} = \psi(-\log H_{\xi,\tilde{\mu},\tilde{\sigma}})$, $x \in \mathbb{R}$, with altered location $\tilde{\mu}$ and scale $\tilde{\sigma}$ for which the precises formulas are provided in Remark 2.9; as usual, (d^*_n) and (c^*_n) are the normalizing sequences corresponding to the iid sequence with margin F. Interestingly, when $\ell_n(x_1,\ldots,x_n) = x_1+\cdots+x_n$ for all $n \geq 2$ so that (X_i) is meta-Archimedean, $\eta_n = n$ and (d^*_n) and (c^*_n) can be used to stabilize M_n.

Example 4.10
Let ψ be an Archimedean generator that satisfies assumption (ii) of Theorem 4.4 and is such that $-\psi'(0) = \infty$. In Charpentier and Segers (2009), this case is referred to as *asymptotic dependence in the upper tail* when $\rho < 1$ and *near asymptotic dependence* when $\rho = 1$. Examples are again given in Table 1. As in Example 4.9, let (X_i) be a meta-Archimax sequence with such generator ψ, a sequence (ℓ_n) of stdfs with $\eta_n = \ell_n(1,\ldots,1) \to \infty$ as $n \to \infty$, and univariate margin $F \in \mathrm{MDA}(H_{\xi,\mu,\sigma})$. Let also (d^*_n) and (c^*_n) be the normalizing sequences corresponding to the iid sequence with margin F.

If we set $\tilde{r}_n = \eta_n$ for all $n \in \mathbb{N}$ as in Example 4.9, we can easily see that $r_n/\tilde{r}_n \to 0$, where r_n is as in Theorem 4.4. Because any Archimedean generator is such that $\psi(t) < 1$ for all $t > 0$, Remark 2.9 and Theorem 4.4 imply that the choice $c_n = c^*_{\lceil\eta_n\rceil}$ and $d_n = d^*_{\lceil\eta_n\rceil}$ leads to a degenerate limit in (1). In the special case of a meta-Archimedean sequence, this means that the simplified choice of normalizing constants (c^*_n) and (d^*_n) from the iid case cannot be used to stabilize M_n. Theorem 4.4 however still applies.

Example 4.11
Ballerini (1994) explored maxima of meta-Archimedean sequences whose generator satisfies the so-called polynomial growth condition, viz.

$$\lim_{t\to\infty} t^\rho\{1-\psi(1/t)\} = c \tag{6}$$

for some $\rho \in (0,1]$ and $c \in (0,\infty)$. It is easily seen that ψ then satisfies assumption (ii) of Theorem 4.4 with the same ρ, because $\{1-\psi(1/t)\} = t^{-\rho}L(t)$ for $L(t) \to \theta$ as $t \to \infty$. The latter property immediately renders L slowly varying, but also shows that the polynomial growth condition is more restrictive than assumption (ii), because slowly varying functions need not tend to a positive constant at ∞. This is also apparent from Example 4.9 and Example 4.10; in the special case $\rho = 1$, the polynomial growth condition implies that $-\psi'(0) = 1/c \in (0,\infty)$. One counterexample of an Archimedean generator for which assumption (ii) holds with $\rho = 1$ and yet $-\psi'(0) = \infty$ is provided by Family 23 of Charpentier and Segers (2009, Table 1), another is given in Table 1. We provide justification for the latter in the Supplementary Material Herrmann et al. (2024) and also explain therein that the family in Ballerini (1994, Example 3) is actually not a counterexample.

Let (X_i) be a meta-Archimax sequence with a generator that satisfies (6), a sequence (ℓ_n) of stdfs such that $\eta_n = \ell_n(1,\ldots,1) \to \infty$ as $n \to \infty$, and univariate margin $F \in$

$\mathrm{MDA}(H_{\xi,\mu,\sigma})$. Now set $\tilde{r}_n = \eta_n^\rho$, $n \in \mathbb{N}$. Because of (6), r_n in Theorem 4.4 satisfies $r_n/\tilde{r}_n \to 1/c \in (0,\infty)$. Remark 2.9 thus implies that M_n could have been alternatively normalized using $c_n = c^*_{\lceil \eta_n^\rho \rceil}$ and $d_n = d^*_{\lceil \eta_n^\rho \rceil}$, leading to the limit $G(x) = D\{H^{1/c}_{\xi,\mu,\sigma}(x)\} = \psi[c^{-\rho}\{-\log H_{\xi,\mu,\sigma}(x)\}^\rho]$, $x \in \mathbb{R}$, in (1). In the special case of a meta-Archimedean sequence, these are precisely the normalizing constants and the limit derived in Ballerini (1994, Theorem 2).

For meta-Archimax sequences it is possible to verify the conditions of Theorem 2.10 provided that the stdf associated with the extreme value part is sufficiently regular. As shown next, this guarantees the uniqueness of the limiting distributions in the sense of Theorem 2.10.

Proposition 4.12
Let (X_i) be a meta-Archimax sequence with stdfs (ℓ_n) and generator ψ which is strict, i.e. $\psi^{-1}(0) = \infty$. Set $\eta_n = \ell_n(1,\ldots,1)$ and define the rate $r : \mathbb{N} \to (0,\infty)$ by $r(n) = r_n = 1/(1-\psi(1/\eta_n))$, $n \geq 1$. Suppose that assumptions (i) and (ii) of Theorem 4.4 hold and that $\lim_{n\to\infty} \eta_{\lceil nt \rceil}/\eta_n = \kappa(t)$ for any $t > 0$, where $\kappa\colon (0,\infty) \to (0,\infty)$ is a bijection. Then the rate function r satisfies $\lim_{n\to\infty} r_{\lceil nt \rceil}/r_n = (\kappa(t))^\rho$, $t > 0$. Furthermore, whenever $(M_n - d_n)/c_n$ converges to a non-degenerate limit G, it holds that $G = D \circ H$, where $D(u) = \psi\{(-\log u)^{1/\rho}\}$ and $H \in \mathrm{GEV}$.

An example of stdfs that satisfy $\lim_{n\to\infty} \eta_{\lceil nt \rceil}/\eta_n = \kappa(t)$ for all $t > 0$ and some bijection $\kappa : (0,\infty) \to (0,\infty)$ is the logistic family $(\ell_{n,\mathrm{Gu}})$ from Example 3.3.

Remark 4.13
The assumptions of Proposition 4.12 always hold when the sequence is meta-Archimedean, provided that ψ satisfies assumption (ii) of Theorem 4.4. Indeed, ψ is completely monotone as explained in Example 4.2 and hence strict. Moreover, $\eta_n = n$, so that $\lceil nt \rceil/n \to t$ for any $t > 0$. Hence, the limit is always of the form $\psi\{(-\log H_{\xi,\mu,\sigma})^{1/\rho}\} = \psi(-\log H_{\rho\xi,\mu,\rho\sigma})$ in view of (5). A similar result has been derived in Wüthrich (2004) albeit under stronger assumptions: Wüthrich (2004, Theorem 3.2) shows that if the limit in (1) is of the form $\psi(-\log H)$ with a non-degenerate H, H must be generalized extreme value.

We close this section with a discussion about when the limiting distribution of normalized maxima of a meta-Archimax sequence is in fact generalized extreme value. From Example 4.8, we already know that this is the case when ψ is the Gumbel–Hougaard generator. The following result shows that this is the only possibility. To avoid assuming (ii) of Theorem 4.4, we use the alternative expression (5) of the limiting distribution in (1).

Lemma 4.14
Suppose that ψ is an Archimedean generator. Then $G = \psi(-\log H)$ is generalized extreme value for all $H \in \mathrm{GEV}$ if and only if there exists $c > 0$ and $\theta \geq 1$ such that $\psi(t) = e^{-(ct)^{1/\theta}}$ for all $t \geq 0$.

An interesting consequence of Lemma 4.14 is the following observation. Suppose that

(X_i) is a stationary meta-Archimax sequence so that (1) holds with $G = \psi\{-\log H\}$ for some $H \in \mathrm{GEV}$. At the same time, suppose that (X_i) fulfills the distributional mixing assumption $\mathcal{D}(u_n)$ from Definition 3.6 for any threshold $u_n = c_n x + d_n$ with $x \in \mathbb{R}$ and (c_n) and (d_n) from (1). From Leadbetter et al. (1983, Theorem 3.3.3), we then necessarily have that $G \in \mathrm{GEV}$, and from Lemma 4.14 we know that this only happens when ψ is the Gumbel–Hougaard generator with parameter $\theta \geq 1$. As argued in Example 4.8, (X_i) is then meta-extreme. In the special case when (X_i) is meta-Archimedean and ψ is regularly varying at 0, (X_i) is necessarily stationary and any non-degenerate limit in (1) is of the form $G = \psi\{-\log H\}$ for some $H \in \mathrm{GEV}$, as seen in Remark 4.13. Moreover, Example 4.8 shows that (X_i) with the Gumbel–Hougaard generator is a meta-extreme sequence with the logistic stdf $\ell_{n,\mathrm{Gu}}$. As discussed at the end of Section 3.2, Proposition 3.10 implies that $\theta = 1$, so that (X_i) is in fact iid. As a consequence, there are many examples of stationary series (X_i) that violate the $\mathcal{D}(u_n)$ condition, but still can be analyzed in our framework. We summarize this as follows.

Corollary 4.15
Suppose that (X_i) is a meta-Archimedean sequence with generator ψ that fulfills assumption (ii) of Theorem 4.4, and assume that (1) holds for some non-degenerate G. Then $\mathcal{D}(u_n)$ from Definition 3.6 holds for all thresholds $u_n = c_n x + d_n$ with $x \in \mathbb{R}$ and (c_n) and (d_n) from (1) if and only if (X_i) is an iid sequence.

4.2 Extensions

We now discuss two generalizations of the results for meta-Archimax sequences from Section 4.1: The first are arbitrary exchangeable sequences and the second are mixtures of not necessarily exchangeable sequences.

All meta-Archimedean sequences are exchangeable, as are meta-Archimax sequences whose stdfs have the symmetry property that $\ell_n(x_1, \ldots, x_n) = \ell_n(x_{\pi(1)}, \ldots, x_{\pi(n)})$ for each $n \in \mathbb{N}$ and each permutation π on $\{1, \ldots, n\}$. Maxima of exchangeable sequences have first been investigated in Berman (1962b). From de Finetti's theorem, there exists a real-valued random variable W with the property that $X_1, X_2, \ldots$ are conditionally independent and identically distributed given W. If G_W denotes the conditional distribution function of X_i given W, we obtain that for each $n \in \mathbb{N}$, $H_n(x_1, \ldots, x_n) = \mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \mathbb{E}[\prod_{i=1}^n G_W(x_i)]$. The results in Berman (1962b) can be broadly summarized as follows. Suppose $\widetilde{F}$ is such that $\widetilde{F} \in \mathrm{MDA}(H)$ with normalizing constants (c_n), $c_n > 0$ and (d_n), where H is either the Weibull, Fréchet or Gumbel distribution. Then (1) holds with these normalizing constants if and only if $\log G_W(x)/\log \widetilde{F}(x)$ converges in distribution to some non-degenerate distribution function A concentrated on $[0, \infty)$ as $x \uparrow x_{\widetilde{F}}$, where $x_{\widetilde{F}} = \sup\{x : \widetilde{F}(x) < 1\}$. The limit in (1) is then of the form $\mathcal{LS}_A(-\log H)$, where $\mathcal{LS}_A$ is the Laplace–Stieltjes transform of A.

Example 4.16
When (X_i) is meta-Archimedean, W is the frailty V from Example 4.2 and $G_W(x) =$

$\mathrm{e}^{-W\psi^{-1}\{F(x)\}}$. The approach of Berman (1962b) can thus be adopted with $\widetilde{F} = \mathrm{e}^{-\psi^{-1}\circ F}$; A is then the distribution function of V and $\mathcal{LS}_A = \psi$. Proceeding this way recovers the results in Wüthrich (2004) which we already related to ours in Remark 4.6. As we explained in Remark 4.13, our theory allows us to show that $\psi(-\log H)$, $H \in \mathrm{GEV}$, are the only possible limits in (1).

The difficulty with the approach in Berman (1962b) is that $\widetilde{F}$ is left unspecified, and it may not be clear whether it exists and how to choose it. To instead use the results derived here, we can rely on Durante and Mai (2010) according to which the copula C_n of the first n elements of an exchangeable sequence (X_i) with continuous margin F is given, for all $u_1, \ldots, u_n \in (0,1)$, by $C_n(u_1, \ldots, u_n) = \int_0^1 \prod_{i=1}^n \frac{\partial B(u_i,t)}{\partial t}\,\mathrm{d}t$, where B is the copula of (X_1, W). This way, $D_n^r(u) = \int_0^1 \{\frac{\partial}{\partial t} B(u^{1/r_n}, t)\}^n\,\mathrm{d}t$. Given that $|\frac{\partial}{\partial t} B(v,t)| \le 1$ for all $v \in (0,1)$ and almost all $t \in (0,1)$ Nelsen (2006, Theorem 2.2.7), the dominated convergence theorem allows us to conclude that $D_n^r \to D$ pointwise on $[0,1]$ with $D(u) = \int_0^1 b(u,t)\,\mathrm{d}t$, provided that $\lim_{n\to\infty}\{\frac{\partial}{\partial t} B(u^{1/r_n}, t)\}^n = b(u,t)$. Because $D_n^r(0) = 0$ and $D_n^r(1) = 1$, we also have $D(0) = 0$ and $D(1) = 1$. The limiting behavior of M_n can then be deduced from Theorem 2.2, if its conditions hold.

Example 4.17

For a meta-Archimedean sequence with generator ψ which is a Laplace transform of a frailty V with distribution function F_V, $B(u,t) = \int_0^{F_V^{-1}(t)} \mathrm{e}^{-v\psi^{-1}(u)}\,\mathrm{d}F_V(v)$ and $\frac{\partial}{\partial t} B(u^{1/r_n}, t) = \mathrm{e}^{-F_V^{-1}(t)\{\psi^{-1}(u^{1/r_n})\}}$. When $1 - \psi(1/\cdot) \in \mathrm{RV}_{-\rho}$, we showed in the proof of Theorem 4.4 that $n\psi^{-1}(u^{1/r_n}) \to (-\log u)^{1/\rho}$, so that the limiting distortion is given by $D(u) = \mathbb{E}[\exp\{-(-\log u)^{1/\rho} F_V^{-1}(T)\}]$ where $T \sim \mathrm{U}(0,1)$. This expression readily simplifies to $D(u) = \mathbb{E}[\mathrm{e}^{-V(-\log u)^{1/\rho}}] = \psi\{(-\log u)^{1/\rho}\}$, which is precisely what we obtained in Theorem 4.4.

Example 4.18

From Durante and Mai (2010), exchangeable sequences can also be constructed by choosing an arbitrary B, F, and a distribution of W. If we set B to be the Eyraud–Farlie–Gumbel–Morgenstern copula $B_\theta(u,t) = ut\{1 + \theta(1-u)(1-t)\}$ Nelsen (2006, Example 3.9), then $\frac{\partial}{\partial t} B_\theta(u,t) = u + \theta u(u-1)(2t-1)$, leading to $b_\theta(u,t) = u^{\theta(2t-1)+1}$ upon setting $r_n = n$. For each $u \in (0,1)$, $D_\theta(u) = \int_0^1 u^{\theta(2t-1)+1}\,\mathrm{d}t = \frac{u^{1+\theta} - u^{1-\theta}}{2\theta \log u}$. If $F \in \mathrm{MDA}(H_{\mu,\sigma,\xi})$, Theorem 2.2 (i) shows that (1) holds with the non-degenerate limit $D_\theta \circ H_{\mu,\sigma,\xi}$. The normalizing constants are those that stabilize the maximum of an iid sequence with margin F. The latter fact allows us to use the results in Berman (1962b) and conclude that there must exist some non-degenerate distribution on $[0,\infty)$ with Laplace–Stieltjes transform $D_\theta \circ e^{-t}$. Indeed, we recognize it to be the uniform distribution on $[1 - |\theta|, 1 + |\theta|]$. From Corollary 2.8 we obtain that if (1) holds with a non-degenerate limit, the latter must be of the form $D_\theta \circ H$, where H is of the same type as $H_{\mu,\sigma,\xi}$. Interestingly, $D_\theta = D_{-\theta}$, even though B_θ and $B_{-\theta}$ embody very different types of dependence.

The second extension is to mixtures of not necessarily exchangeable sequences.

Example 4.19
Consider $X_1, X_2, \ldots \sim F$ with the property that for each $d \in \mathbb{N}$ and distinct $t_1, \ldots, t_d \in \mathbb{N}$, the copula $C_{t_1,\ldots,t_d;\theta}$ of $(X_{t_1}, \ldots, X_{t_d})$ depends on a parameter $\theta \in \Theta \subseteq \mathbb{R}^k$. Suppose that the mapping $\theta \mapsto C_{t_1,\ldots,t_d;\theta}(u_1, \ldots, u_d)$ is Borel-measurable for every $u_1, \ldots, u_d \in [0,1]$. These conditions are fulfilled for several sequences constructed in this paper, notably all meta-Archimedean sequences with parametric generators. Let T be some distribution function on Θ. From Durante and Sempi (2016, Corollary 1.4.7, p. 17), $\widetilde{C}_{t_1,\ldots,t_d}$ given by $\widetilde{C}_{t_1,\ldots,t_d}(u_1, \ldots, u_d) = \int_\Theta C_{t_1,\ldots,t_d;\theta}(u_1, \ldots, u_d)\,\mathrm{d}T(\theta)$ is a copula. By the Daniell–Kolmogorov theorem, there exists a sequence (U_i) of standard uniform variables such that for each $n \in \mathbb{N}$ and distinct $t_1, \ldots, t_n \in \mathbb{N}$, $(U_{t_1}, \ldots, U_{t_n}) \sim \widetilde{C}_{t_1,\ldots,t_n}$. A mixture sequence $(\widetilde{X}_i)$ with margin F is then obtained by setting $\widetilde{X}_i = F^{-1}(U_i)$.

To simplify the notation, write $C_{n;\theta} = C_{1,\ldots,n;\theta}$ and $\widetilde{C}_n = \widetilde{C}_{1,\ldots,n}$, and let $\delta_{n;\theta}$ and $\widetilde{\delta}_n$ denote the diagonal of $C_{n;\theta}$ and $\widetilde{C}_n$. Suppose that there exists a rate function r, independent of θ, such that for each $\theta \in \Theta$, the diagonal power distortion $D^r_{n;\theta}$ of $C_{n;\theta}$ converges pointwise to a continuous map D_θ. Then the dominated convergence theorem implies that for all $u \in [0,1]$, $\lim_{n\to\infty} \widetilde{D}^r_n(u) = \lim_{n\to\infty} \widetilde{\delta}_n\{u^{1/r_n}\} = \lim_{n\to\infty} \int_\Theta D^r_{n;\theta}(u)\,\mathrm{d}T(\theta) = \int_\Theta D_\theta(u)\,\mathrm{d}T(\theta) = \widetilde{D}(u)$, i.e. the limiting distortion is a mixture with the same mixing distribution T. Moreover, if for some $H_{\xi,\mu,\sigma} \in \mathrm{GEV}$ and sequences (c^*_n), $c^*_n > 0$, and (d^*_n), $F^n(c^*_n x + d^*_n) \to H_{\xi,\mu,\sigma}(x)$ for all $x \in \mathbb{R}$, Theorem 2.2 (i) shows that (1) holds with $G = \widetilde{D} \circ H_{\xi,\mu,\sigma}$ and normalizing sequences given, for all $n \in \mathbb{N}$, by $c_n = c^*_{\lceil r_n \rceil}$ and $d_n = d^*_{\lceil r_n \rceil}$. The same sequences can be used to stabilize the maximum of (X_i) for each $\theta \in \Theta$.

As a concrete example, let (X_i) be the meta-Archimedean sequence with some margin $F \in \mathrm{MDA}(H)$ and the Ali–Mikhail–Haq generator ψ_θ with parameter $\theta \in (0,1)$, viz. Table 1. Because $-\psi'_\theta(0) < \infty$, Example 4.9 shows that we can choose $r_n = n$; the limiting diagonal power distortion is then given, for all $u \in [0,1]$, by $D_\theta(u) = \psi_\theta(-\log u)$. Various mixtures can now be obtained by choosing a distribution T on $(0,1)$. For example, if T is uniform, $\widetilde{D}(u) = \int_0^1 \frac{1-\theta}{1/u-\theta}\,\mathrm{d}\theta = 1 - \frac{u-1}{u}\log(1-u)$.

5 Uniform convergence rates of maxima under dependence

We now discuss uniform convergence rates of the distribution of normalized maxima to their weak limit, akin to Berry–Esseen type theorems for sums. Numerous such results exist for maxima of iid sequences, a review is provided in Leadbetter and Rootzén (1988, Section 2.8), starting with the early results for normal Hall (1979) and Nair (1981) and exponential Hall and Wellner (1979) random variables. Convergence rates for various classes of distributions have been investigated since, see e.g. Peng et al. (2010); also noteworthy are general results for the Fréchet domain of attraction in Pereira (1983). Convergence rates for maxima of dependent sequences have merely been obtained for specific combinations of dependence patterns and marginal distributions, such as stationary normal sequences Cohen (1982) and Rootzén (1983), or chain-dependent sequences McCormick and Seymour (2001); results for continuous-time stochastic processes can be found in Kratz and Rootzén

(1997).

Here, we derive a general result for dependent sequences by exploiting the interplay between their diagonal power distortions and the extremal properties of the associated iid sequence. This allows us to transfer uniform convergence rates from the iid case to dependent sequences by only slightly strengthening the conditions in Theorem 2.2 (i).

Theorem 5.1 (Uniform convergence rate under dependence)
Consider a sequence $X_1, X_2, \ldots \sim F$ with continuous margin F and assume the following conditions to hold.

(i) There exist $H \in \mathrm{GEV}$ and sequences (c_n^*), $c_n^* > 0$, and (d_n^*) and a map $\beta^* \colon \mathbb{N} \to [0, \infty)$ satisfying $\beta^*(n) \to 0$ as $n \to \infty$ such that $\sup_{x\in\mathbb{R}} |F^n(c_n^* x + d_n^*) - H(x)| \le \beta^*(n)$.

(ii) There exists a rate $r \colon \mathbb{N} \to (0, \infty)$ with $r_n \to \infty$ as $n \to \infty$ and functions $D \colon [0,1] \to [0,1]$ and $s \colon \mathbb{N} \to [0, \infty)$ with $s(n) \to 0$ for $n \to \infty$ such that the diagonal power distortion D_n^r corresponding to r satisfies $\sup_{u\in[0,1]} |D_n^r(u) - D(u)| \le s(n)$.

(iii) The limit D in (ii) is Hölder continuous with constant K and parameter $0 < \kappa \le 1$.

Then

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(M_n \le c^*_{\lceil r_n \rceil} x + d^*_{\lceil r_n \rceil}) - D \circ H(x)| \le K \left\{ \beta^* \left(\lceil r_n \rceil \right) + 3\mathrm{e}^{-1} \mathbb{1}_{\{r_n \notin \mathbb{N}\}} / r_n \right\}^{\kappa} + s(n),$$

where the right-hand side converges to 0 for $n \to \infty$.

Assumption (ii) in Theorem 5.1 is not strong. Indeed, if D_n^r converges pointwise to D, Lemma A.1 implies that this convergence is uniform on $[0,1]$, so assumption (ii) merely introduces the notation for the convergence rate. To apply Theorem 5.1 to any model where the rate in the associated iid case is known, one therefore only needs to explicitly identify the rate function s linked to the dependence structure.

Example 5.2
For the moving maximum process in Example 2.12, set $r_n = n$ to obtain $D_n^r(u) = u^{(n+k)/(n(k+1))}$ and $D(u) = u^{1/(k+1)}$. Because $k \ge 0$, Lemma A.3 in the Supplementary Material Herrmann et al. (2024) implies that $s(n) = \sup_{u\in[0,1]} |D_n^r(u) - D(u)| = \frac{k}{n+k}(1 + \frac{k}{n})^{-n/k}$. With $k/(n+k) \to 0$ and $(1 + k/n)^{-n/k} \to 1/\mathrm{e}$, we indeed see that $s(n) \to 0$ as $n \to \infty$. Furthermore, $D(u) = u^{1/(k+1)}$ is Hölder continuous with constant $K = 1$ and $\kappa = 1/(k+1)$.

Using Theorem 5.1, we can now combine the above convergence rate of D_n^r with uniform convergence rates from the iid case. For example, if F is the standard normal distribution function Φ, Hall (1979) provides sequences of constants (c_n^*) and (d_n^*) such that $\sup_{x\in\mathbb{R}} |\Phi^n(c_n^* x + d_n^*) - \Lambda(x)| \le 3/\log(n)$, where $\Lambda(x) = H_{0,0,1}(x)$ is the (standard) Gumbel distribution. Theorem 5.1 applied to a sequence of standard normal random variables with dependence specified by the moving maximum process thus leads to the convergence rate $\sup_{x\in\mathbb{R}} |\mathbb{P}(M_n \le c_n^* x + d_n^*) - \Lambda(x)^{1/(k+1)}| \le (\frac{3}{\log(n)})^{1/(k+1)} + \frac{k}{n+k}(1 + \frac{k}{n})^{-n/k}$. From the power $1/(k+1)$ of $\beta^*(n) = 3/\log(n)$ we see that the upper bound tends to 0 slower than in the iid case, with higher values of k leading to slower convergence rates.

Applying Theorem 5.1 is particularly easy in case of sequences with asymptotic power diagonals in the sense of Definition 3.1. Indeed, assumption (iii) always holds given that the limiting distortion u^θ is Hölder continuous with $\kappa = \theta$. In the special case of sequences with power diagonals, $r_n = \eta_n$ and $D_n(u) = u$ as discussed in Section 3, so that assumption (ii) is trivially true with $s(n) = 0$. This leads to the following corollary.

Corollary 5.3
Suppose that $X_1, X_2, \ldots \sim F \in \mathrm{MDA}(H)$ is a sequence with a power diagonal and that the continuous margin F satisfies assumption (i) of Theorem 5.1. Then $\sup_{x\in\mathbb{R}} |\mathbb{P}(M_n \leq c^*_{\lceil\eta_n\rceil} x + d^*_{\lceil\eta_n\rceil}) - H(x)| \leq \beta^*(\lceil\eta_n\rceil) + 3\mathrm{e}^{-1}\mathbb{1}_{\{\eta_n\notin\mathbb{N}\}}/\eta_n$. Additionally, if β^* is monotonically decreasing and C_n is positively quadrant dependent (PQD) for each $n \in \mathbb{N}$, i.e. $C_n(\boldsymbol{u}) \geq \Pi_n(\boldsymbol{u})$ for all $\boldsymbol{u} \in [0,1]^d$, then $\beta^*(\lceil\eta_n\rceil) \geq \beta^*(n)$.

Corollary 5.3 implies that under its hypothesis and when C_n is PQD for each n, the upper bound to the uniform convergence rate in (1) is higher compared to the iid case, but the limiting distribution remains unchanged. The following example illustrates this point.

Example 5.4
As discussed in Section 3.1, meta-extreme sequences have power diagonals provided that $\ell_n(1,\ldots,1) \to \infty$ as $n \to \infty$. Because any stdf is bounded above by 1, their copulas are also PQD. Consider e.g. the logistic sequence with standard normal margins, constructed as in Example 3.3. Here, for each $n \in \mathbb{N}$, C_n is the Gumbel–Hougaard copula with parameter $\theta \geq 1$, where $\theta = 1$ corresponds to the iid case, and $\eta_n = r_n = n^{1/\theta}$. Using the same normalizing sequences (c^*_n) and (d^*_n) as in Example 5.2, Corollary 5.3 leads to the uniform bound $\sup_{x\in\mathbb{R}} |\mathbb{P}(M_n \leq c^*_{\lceil\eta_n\rceil} x + d^*_{\lceil\eta_n\rceil}) - \Lambda(x)| \leq \frac{3}{\log(\lceil n^{1/\theta}\rceil)} + 3\mathrm{e}^{-1} n^{-1/\theta} \approx \frac{3\theta}{\log(n)}$. Again, a higher degree of dependence, i.e. a larger value of θ, leads to a higher upper bound of the uniform convergence rate.

While it may be intuitive that dependence slows down convergence rates, the next example shows that this may not always be true. In fact, convergence may even be faster compared to the iid case if the maximum does not need to be normalized. An extreme case in point is a comonotone sequence, where $M_n \sim F$ for each n and the convergence is instantaneous.

Example 5.5
Consider the meta-extreme sequence in Example 2.7, where C_n is the Cuadras-Augé copula with parameter $\theta \in (0,1)$. Because $\eta_n = \ell_n(1,\ldots,1) = (1-(1-\theta)^n)/\theta$ so that $\eta_n \to 1/\theta$, this sequence does not have a power diagonal in the sense of Definition 3.1, so that Corollary 5.3 does not apply. Nonetheless, we can calculate that $\sup_{x\in\mathbb{R}} |\mathbb{P}(M_n \leq x) - F^{1/\theta}(x)| = \sup_{x\in\mathbb{R}} |F^{\eta_n}(x) - F^{1/\theta}(x)| = \sup_{u\in[0,1]} |u^{1/\theta-(1-\theta)^n/\theta} - u^{1/\theta}|$, where equality holds because F is continuous. By Lemma A.3 (i) in the Supplementary Material Herrmann et al. (2024), $\beta(n) := \sup_{u\in[0,1]} |u^{1/\theta-(1-\theta)^n/\theta} - u^{1/\theta}| \leq (1-\theta)^n\{1-(1-\theta)^n\}^{(1-\theta)^{-n}-1}$. For all n large enough we have, by Lemma A.3 (ii), that $\beta(n) \leq 3\mathrm{e}^{-1}(1-\theta)^n$, which shows that not only $\lim_{n\to\infty}\beta(n) = 0$ but also (in contrast to Example 5.2 and Example 5.4) that a higher degree of dependence, i.e. a larger value of θ, leads to a *smaller* upper bound on the

uniform convergence rate. When $\theta = 1$, the sequence is comonotone and $\beta(n) = 0$ for all n.

We conclude this section with a partial converse to Theorem 5.1 which allows us to infer the convergence rate of the dependence structure when the overall convergence rate is known.

Proposition 5.6
Consider $X_1, X_2, \ldots \sim F$ where the continuous margin F satisfies assumption (i) of Theorem 5.1. Suppose also that there exists $r\colon \mathbb{N} \to (0,\infty)$ with $r_n \to \infty$ for $n \to \infty$ such that the following conditions are fulfilled:

(i) There exists a distribution function $D : [0,1] \to [0,1]$ such that $\sup_{x\in\mathbb{R}} |\mathbb{P}(M_n \le c^*_{\lceil r_n \rceil} x + d^*_{\lceil r_n \rceil}) - D \circ H(x)| \le \gamma(n)$, where $\gamma(n) \to 0$ as $n \to \infty$;

(ii) For each $n \in \mathbb{N}$, the diagonal power distortion D^r_n with respect to r is Hölder continuous with constants K_n and κ_n, and there exist $K, \kappa > 0$ such that $\sup_{n\in\mathbb{N}} K_n \le K$ and $\inf_{n\in\mathbb{N}} \kappa_n \ge \kappa$.

Then $\sup_{u\in[0,1]} |D^r_n(u) - D(u)| \le K \left(\beta^*\left(\lceil r_n \rceil\right) + 3\mathrm{e}^{-1}\mathbb{1}_{\{r_n \notin \mathbb{N}\}}/r_n\right)^{\kappa} + \gamma(n)$, where the right-hand side converges to 0 as $n \to \infty$.

Proposition 5.6 can be used to transfer convergence-rate results for time series to convergence rates of the underlying dependence structure represented by D^r_n.

Example 5.7
For the moving maximum process in Example 2.12, the diagonal power distortion is given by $D^r_n(u) = u^{(n+k)/(nk+n)}$ with $r_n = n$. Clearly, D^r_n is Hölder continuous with $\kappa_n = (n+k)/(nk+n) \le 1$ and $K_n = 1$. Because $\kappa_n \to 1/(k+1)$ as $n \to \infty$, assumptions (i) and (ii) in Proposition 5.6 hold.

6 Conclusion

We have investigated weak convergence limits of maxima of dependent sequences of identically distributed random variables after suitable affine normaliziation accomplished through subsequences of the normalizing sequences from the iid case. We have worked under broad conditions on the diagonal of the underlying copula and demonstrated through various examples that these conditions are fulfilled by a large number of sequences. The limiting distribution is a composition of the generalized extreme value limit of the associated iid sequence and a function, termed distortion, which can have various shapes. This characterization could be particularly useful for future development of statistical inference methods.

We close by remarking that there exist dependent sequences whose maxima converge after standardization with sequences that cannot be subsequences of the normalizing sequences from the iid case. Consider the following example from Berman (1962a). Let (Z_i) be a sequence of iid standard normal random variables, and set, for each $i \in \mathbb{N}$, $X_i = \sqrt{\varrho} Z_0 + \sqrt{1-\varrho} Z_i$. It is easily verified that $X_i \sim \mathrm{N}(0,1)$ for each $i \in \mathbb{N}$, and that $\operatorname{cor}(X_i, X_j) = \varrho$

whenever $i \neq j$. As is well-known, there exist sequences (c_n^*), $c_n^* > 0$ and (d_n^*), such that $\{\max(Z_1, \ldots, Z_n) - d_n^*\}/c_n^* \xrightarrow{\mathrm{d}} H_{0,\mu,\sigma}$; from the convergence to types theorem Resnick (1987, Proposition 0.2) and Embrechts et al. (1997, Example 3.3.29), one has $c_n^* = O((2\log n)^{-1/2})$. However, a direct calculation shows that $\{\max(X_1, \ldots, X_n) - d_n\}/c_n \xrightarrow{\mathrm{d}} \mathrm{N}(0, \varrho)$, where for each $n \in \mathbb{N}$, $d_n = \sqrt{1-\varrho} d_n^*$ and $c_n = 1$. If $\tilde{c}_n$ is of the form $c^*_{\lceil r_n \rceil}$ for some rate r with $r_n \to \infty$ as $n \to \infty$, we cannot possibly have that $\tilde{c}_n / c_n \to c \in (0, \infty)$. Even though several existing results are embedded in our framework, a unified theory thus remains to be developed.

A Auxiliary results and their proofs

Lemma A.1

Denote by f_n a sequence of non-decreasing functions that converges pointwise to a continuous limit f on an interval $[a, b]$. Then the convergence of f_n to f is uniform.

Proof. Given that f is continuous on a compact interval, it is uniformly continuous and hence for a given $\varepsilon > 0$ there is a partition $a = t_0 < t_1 < \cdots < t_K = b$ such that $x, y \in [t_i, t_{i+1}] \Rightarrow |f(x) - f(y)| < \varepsilon/2$. For every t_i, $i = 0, \ldots, K$, we also have $f_n(t_i) \to f(t_i)$ and hence there is a n_ε such that for all $n \geq n_\varepsilon$ we have $|f_n(t_i) - f(t_i)| < \varepsilon/2$ simultaneously for all $i = 0, \ldots, K$. This immediately yields that for all $n \geq n_\varepsilon$, $f_n(t_{i+1}) < f(t_{i+1}) + \varepsilon/2$ and $f(t_i) - \varepsilon/2 < f_n(t_i)$. Now fix $\varepsilon > 0$ and an arbitrary $x \in [a, b]$. Then $x \in [t_i, t_{i+1}]$ for some $i \in \{0, \ldots, K\}$. By the monotonicity of f_n, $f_n(t_i) \leq f_n(x) \leq f_n(t_{i+1})$, and hence with the previous two inequalities $f(t_i) - \varepsilon/2 < f_n(t_i) \leq f_n(x) \leq f_n(t_{i+1}) < f(t_{i+1}) + \varepsilon/2$. Considering now $f(x)$ we have by design of the partition that $f(x) - \varepsilon/2 < f(t_i)$ and $f(t_{i+1}) < f(x) + \varepsilon/2$. Therefore, $f(x) - \varepsilon < f_n(x) < f(x) + \varepsilon$, or equivalently $|f_n(x) - f(x)| < \varepsilon$. Since ε and x were arbitrary this concludes the proof. □

Lemma A.2

Denote by (f_n) and (g_n) two sequences of functions $f_n, g_n \colon \mathbb{R} \to \mathbb{R}$, and by $(h_n)_{n \geq 1}$ the composition $h_n = f_n \circ g_n$, i.e. $h_n(x) = f_n(g_n(x))$ for all $x \in \mathbb{R}$. Suppose that f_n converges uniformly to a function $f \colon \mathbb{R} \to \mathbb{R}$ which is continuous on $\mathcal{A} \subseteq \mathbb{R}$ and that g_n converges to pointwise on $\mathcal{A}$ to an arbitrary function $g \colon \mathbb{R} \to \mathbb{R}$, then h_n converges pointwise to $h = f \circ g$ on $\mathcal{A}$.

Proof. Fix an arbitrary $x \in \mathcal{A}$ and set $y_n = g_n(x)$ and $y = g(x)$ for further reference. Then $h_n(x) = f_n(y_n)$ and $h(x) = f(y)$ which leads to $|h_n(x) - h(x)| = |f_n(y_n) - f(y) + f(y_n) - f(y_n)| \leq |f_n(y_n) - f(y_n)| + |f(y_n) - f(y)|$. Due to pointwise convergence of g_n to g on $\mathcal{A}$, $y_n \to y$ and hence the second term tends to zero by continuity of f. For the first term we have $|f_n(y_n) - f(y_n)| \leq \sup_{y \in \mathbb{R}} |f_n(y) - f(y)|$ which converges to zero because $f_n \to f$ uniformly by assumption. □

Lemma A.3

(i) For arbitrary fixed a, b such that $0 < a < b$,

$$\sup_{u\in[0,1]} |u^a - u^b| = (1 - a/b)(b/a)^{a/(a-b)}. \tag{7}$$

(ii) For any fixed $a > 0$ a sequence $b_n \to 0$ we have for n large enough that $\sup_{u\in[0,1]} |u^a - u^{a+b_n}| \leq 3\mathrm{e}^{-1}|b_n|/a$.

(iii) For any fixed $b > 0$ a sequence $a_n \to \infty$ we have for n large enough that $\sup_{u\in[0,1]} |u^{a_n} - u^{a_n+b}| \leq 3\mathrm{e}^{-1}b/a_n$.

Proof.

(i) Fix arbitrary $0 < a < b$ and observe that the function f given, for all $u \in [0,1]$, by $f(u) = u^a - u^b$ is continuous, positive, and satisfies $f(0) = f(1) = 0$. Furthermore, $u^* = (b/a)^{1/(a-b)}$ is easily seen to be the maximizer of f given that $f'(u) = au^{a-1} - bu^{b-1}$. Equation (7) then follows from the fact that $f(u^*) = (1 - a/b)(b/a)^{a/(a-b)}$.

(ii) Fix an arbitrary $a > 0$ and let $n \in \mathbb{N}$ be sufficiently large such that we have $b_n \in (-a\log(2), a\log(2)/(1-\log(2)))$. If $b_n = 0$, then the claim is trivially true. Otherwise, a case distinction needs to be made, depending on the sign of b_n.

In the case when $b_n > 0$, we make use of the inequality $1-(1/x)\log(1+x) \leq x/(x+1)$ valid for all $x > 0$. By setting $x = b_n/a$, and noting that $b_n < a\log(2)/(1-\log(2))$ is equivalent to $b_n/(b_n+a) < \log(2)$, we obtain $1 - \frac{a}{b_n}\log(1+\frac{b_n}{a}) \leq \frac{b_n}{b_n+a} < \log(2)$. This allows us to apply Leadbetter et al. (1983, Lemma 2.4.1 (ii)) to derive that for some $\theta \in (0,1)$,

$$\begin{aligned}(1+b_n/a)^{-a/b_n} - \mathrm{e}^{-1} &= \mathrm{e}^{-a\log(1+b_n/a)/b_n} - \mathrm{e}^{-1} \\ &= \mathrm{e}^{-1}\left[\{1 - a\log(1+b_n/a)/b_n\} + \theta\{1 - a\log(1+b_n/a)/b_n\}^2\right] \\ &\leq \mathrm{e}^{-1}\frac{b_n}{a+b_n}\left(1 + \theta\frac{b_n}{a+b_n}\right) \leq 2\mathrm{e}^{-1}\frac{b_n}{a+b_n},\end{aligned}$$

where the last inequality stems from $b_n/(b_n+a) \leq \log(2) < 1$. Equation (7) from part (i) then leads to

$$\begin{aligned}\sup_{u\in[0,1]} |u^a - u^{a+b_n}| &= (1+b_n/a)^{-a/b_n}\frac{b_n}{a+b_n} = \left\{\mathrm{e}^{-1} + (1+b_n/a)^{-a/b_n} - \mathrm{e}^{-1}\right\}\frac{b_n}{a+b_n} \\ &\leq \left(\mathrm{e}^{-1} + 2\mathrm{e}^{-1}\frac{b_n}{a+b_n}\right)\frac{b_n}{a+b_n} = \mathrm{e}^{-1}\frac{b_n}{a+b_n}\left(1 + \frac{2b_n}{a+b_n}\right).\end{aligned}$$

Using again that $b_n/(b_n+a) < 1$ and further that $\theta \in (0,1)$ and $b_n/(a+b_n) \leq b_n/a$, we finally get $\sup_{u\in[0,1]} |u^a - u^{a+b_n}| \leq 3\mathrm{e}^{-1}b_n/a$.

The case when $b_n < 0$ is similar, but here we make use of the inequality $\log(1+x) \leq x$ valid for all $x > -1$, which leads to $1 - (1+1/x)\log(1+x) \leq -x$ whenever $x \in (-1, 0)$.

Setting $x = b_n/a$, we then have that $1 - (1 + \frac{a}{b_n}) \log(1 + \frac{b_n}{a}) \leq -\frac{b_n}{a} < \log(2)$, where the last inequality holds because n is assumed large enough so that $b_n > -a\log(2)$. Leadbetter et al. (1983, Lemma 2.4.1 (ii)) then implies that for some $\theta \in (0,1)$,

$$\begin{aligned}
&(1 + b_n/a)^{-(1+a/b_n)} - \mathrm{e}^{-1} = \mathrm{e}^{-(1+a/b_n)\log(1+b_n/a)} - \mathrm{e}^{-1} \\
&= \mathrm{e}^{-1}\left[1 - (1 + a/b_n)\log(1 + b_n/a) + \theta\big\{1 - (1 + a/b_n)\log(1 + b_n/a)\big\}^2\right] \\
&\leq \frac{|b_n|}{a}\mathrm{e}^{-1}\left(1 + \theta\frac{|b_n|}{a}\right) \leq 2\mathrm{e}^{-1}\frac{|b_n|}{a},
\end{aligned}$$

where the last inequality derives from $|b_n|/a < \log(2) < 1$. Equation (7) from part (i) then yields

$$\begin{aligned}
\sup_{u\in[0,1]} |u^a - u^{a+b_n}| &= -\frac{b_n}{a}(1 + b_n/a)^{-(1+a/b_n)} = -\frac{b_n}{a}\big\{\mathrm{e}^{-1} + (1 + b_n/a)^{-(1+a/b_n)} - \mathrm{e}^{-1}\big\} \\
&\leq -\frac{b_n}{a}\left(\mathrm{e}^{-1} + 2\mathrm{e}^{-1}\frac{|b_n|}{a}\right) \leq 3\mathrm{e}^{-1}\frac{|b_n|}{a},
\end{aligned}$$

as claimed, where the last step utilizes that $|b_n|/a < 1$ one more time.

(iii) First note that $(1 - \log(2))/\log(2) \approx 0.4427$. Suppose that n is sufficiently large so that $b(1-\log(2))/\log(2) \leq a_n$. We then have $b/(a_n+b) \leq \log(2)$. Calling again on the inequality $1 - \log(1 + x)/x \leq x/(x + 1)$ valid for all $x > 0$, we obtain with $x = b/a_n$ that $1 - (a_n/b)\log(1 + b/a_n) \leq b/(a_n + b) \leq \log(2)$. Applying Leadbetter et al. (1983, Lemma 2.4.1 (ii)) then yields that for some $\theta \in (0,1)$,

$$\begin{aligned}
(1 + b/a_n)^{-a_n/b} - \mathrm{e}^{-1} &= \mathrm{e}^{-a_n\log(1+b/a_n)/b} - \mathrm{e}^{-1} \\
&= \mathrm{e}^{-1}\left[\{1 - a_n\log(1 + b/a_n)/b\} + \theta\{1 - a_n\log(1 + b/a_n)/b\}^2\right] \\
&\leq \mathrm{e}^{-1}\frac{b}{a_n + b}\left(1 + \frac{\theta b}{a_n + b}\right) \leq 2\mathrm{e}^{-1}\frac{b}{a_n + b},
\end{aligned}$$

where the last inequality holds since $b/(b + a_n) \leq \log(2) < 1$. Equation (7) in part (i) then gives

$$\begin{aligned}
\sup_{u\in[0,1]} |u^{a_n} - u^{a_n+b}| &\leq \frac{b}{b + a_n}(1 + b/a_n)^{-a_n/b} = \frac{b}{b + a_n}\left\{\mathrm{e}^{-1} + (1 + b/a_n)^{-a_n/b} - \mathrm{e}^{-1}\right\} \\
&\leq \frac{b}{b + a_n}\left(\mathrm{e}^{-1} + 2\mathrm{e}^{-1}\frac{b}{a_n + b}\right) \leq 3\mathrm{e}^{-1}\frac{b}{b + a_n},
\end{aligned}$$

where $b/(b+a_n) < 1$ is used in the last step. Since $b/(b+a_n) < b/a_n$, the claim follows.

□

Corollary A.4

Denote by (y_n) a sequence of real numbers such that $y_n \to y > 0$ and by (F_n) a sequence of distribution functions on $\mathbb{R}$ that converges weakly to a distribution function F on $\mathbb{R}$. Then $F_n^{y_n}$ converges weakly to F^y.

Proof. Because F^y and F have the same points of continuity, we need to show that $F_n^{y_n}(x) \to F^y(x)$ as $n \to \infty$ at any continuity point x of F. To this end, set $a = y$ and $b_n = y_n - y$ and define the functions g_n and g for all $u \in [0,1]$ by $g_n(u) = u^{y_n}$ and $g(u) = u^y$, respectively. Lemma A.3 (ii) then implies that $\sup_{u\in[0,1]} |g(u) - g_n(u)| = \sup_{u\in[0,1]} |u^a - u^{a+b_n}| \leq \frac{3}{y}\mathrm{e}^{-1}|y_n - y| \to 0$, i.e. g_n converges uniformly to g. Given that g is obviously continuous, and $F_n(x) \to F(x)$ at any continuity point of F we can conclude from Lemma A.2 that the composition $g_n \circ F_n$ converges pointwise to $g \circ F$ at any continuity point of F as claimed. □

Corollary A.5
Denote by (r_n) a sequence of positive real numbers such that $r_n \to \infty$. Then for all $n \in \mathbb{N}$ we have

$$\sup_{u\in[0,1]} |u - u^{r_n/\lceil r_n \rceil}| \leq 3\mathrm{e}^{-1}\frac{1}{\lceil r_n \rceil},$$

where the right-hand side tends to 0 as $n \to \infty$.

Proof. Set $a = 1$ and $b_n = r_n/\lceil r_n \rceil - 1 = (r_n - \lceil r_n \rceil)/\lceil r_n \rceil$. Clearly, $b_n \leq 0$ and $b_n \to 0$ as $n \to \infty$. Part (ii) of Lemma A.3 and its proof then yield the inequality holds as soon as $b_n > -\log(2)$, which happens whenever $r_n > 1$. However, for $r_n \leq 1$ we have $3\mathrm{e}^{-1}/\lceil r_n \rceil = 3\mathrm{e}^{-1} \approx 1.103638$ in which case the statement continues to be correct since we always have $|u - u^{r_n/\lceil r_n \rceil}| \leq 1$. □

Corollary A.6
Denote by (r_n) a sequence of positive real numbers such that $r_n \to \infty$ and by F a distribution function on $\mathbb{R}$. We then have for all $n \in \mathbb{N}$ that

$$\sup_{x\in\mathbb{R}} |F^{\lceil r_n \rceil}(x) - F^{r_n}(x)| \leq 3\mathrm{e}^{-1}\frac{1}{r_n}$$

and the right-hand side tends to 0 as $n \to \infty$.

Proof. First note that $\sup_{x\in\mathbb{R}} |F^{\lceil r_n \rceil}(x) - F^{r_n}(x)| \leq \sup_{x\in\mathbb{R}} |F^{r_n+1}(x) - F^{r_n}(x)| \leq \sup_{u\in[0,1]} |u^{r_n} - u^{r_n+1}|$. Upon setting $b = 1$ and $a_n = r_n$, part (iii) of Lemma A.3 and its proof imply that the right-hand side is bounded above by $3\mathrm{e}^{-1}/r_n$ as soon as $r_n \geq (1 - \log(2))/\log(2) \approx 0.442695$. However, when $r_n < (1 - \log(2))/\log(2)$ we have $3\mathrm{e}^{-1}/r_n > 1$ and so $|F^{\lceil r_n \rceil}(x) - F^{r_n}(x)| \leq 1$ is always true. As such the statement holds for all $n \in \mathbb{N}$, even though it only becomes informative when $3\mathrm{e}^{-1}/r_n < 1$. □

B Proofs of the main results

Proof of Theorem 2.2. For part (i), define f_n and g_n for all $x \in \mathbb{R}$ and $u \in [0,1]$ by $f_n(x) = F^{\lceil r_n \rceil}(c^*_{\lceil r_n \rceil} x + d^*_{\lceil r_n \rceil})$ and $g_n(u) = u^{r_n/\lceil r_n \rceil}$, respectively. The fact that $r_n \to \infty$ as

$n \to \infty$ implies that $f_n \to H$ pointwise as $n \to \infty$. Now consider $h_n = g_n \circ f_n$ so that for all $x \in \mathbb{R}$, $h_n(x) = F^{r_n}(c^*_{\lceil r_n \rceil} x + d^*_{\lceil r_n \rceil})$. Corollary A.5 together with Lemma A.2 implies that $h_n \to H$ pointwise on $\mathbb{R}$ as $n \to \infty$. Furthermore, Lemma A.1 implies that the convergence $D^r_n \to D$ is uniform. Hence, for all $x \in \mathbb{R}$ by Lemma A.2, $\mathbb{P}(M_n \leq c^*_{\lceil r_n \rceil} x + d^*_{\lceil r_n \rceil}) = \delta_n[\{F^{r_n}(c^*_{\lceil r_n \rceil} x + d^*_{\lceil r_n \rceil})\}^{1/r_n}] = (D^r_n \circ h_n)(x) \to D\{H(x)\}$. As a uniform limit of functions with these properties, D is continuous, non-decreasing, $D(0) = 0$, and $D(1) = 1$. Hence D is a continuous distribution function on $[0,1]$.

For part (ii), $D^r_n(u) = \delta_n(u^{1/r_n})$ implies $D^r_n(u) = D(u)$ for $u \in \{0,1\}$ and all $n \in \mathbb{N}$. For a given continuity point $u \in (0,1)$ of D, pick $x \in \mathbb{R}$ such that $u = H(x)$ and set $u_n = F^{r_n}(c^*_{\lceil r_n \rceil} x + d^*_{\lceil r_n \rceil})$. By construction and Corollary A.4, $u_n = (F^{\lceil r_n \rceil}(c^*_{\lceil r_n \rceil} x + d^*_{\lceil r_n \rceil}))^{r_n/\lceil r_n \rceil} \to u$, and, by the triangle inequality,

$$|D^r_n(u) - D(u)| \leq |D^r_n(u) - D^r_n(u_n)| + |D^r_n(u_n) - D(u)|. \tag{8}$$

Now choose an arbitrary $\varepsilon > 0$. By assumption, there exists $n_1 \in \mathbb{N}$ such that $|D^r_n(u_n) - D(u)| = |\mathbb{P}((M_n - d^*_{\lceil r_n \rceil})/c^*_{\lceil r_n \rceil} \leq x) - D \circ H(x)| < \varepsilon/2$ for all $n > n_1$. To handle the first term in (8), Lipschitz continuity of copulas Nelsen (2006, Theorem 2.10.7) implies that $|D^r_n(u) - D^r_n(u_n)| \leq n|u^{1/r_n} - u_n^{1/r_n}|$. Next, let $f_n(y) = y^{1/r_n}$, $y \in [0,1]$, and pick $\delta > 0$ small enough so that $\delta < \min(u, 1-u)$. By assumption and $u_n \to u$, there exists n_2 so that for all $n > n_2$, $1/r_n < 1$ and at the same time $u_n \in (u-\delta, u+\delta)$. The derivative $f'_n(y) = (y^{1/r_n - 1})/r_n$ is then continuous and decreasing in y on $[u-\delta, u+\delta]$ and attains its maximum at $u-\delta$. This means that f_n is Lipschitz on $[u-\delta, u+\delta]$ with constant $L_n = f'_n(u-\delta)$ whenever $n > n_2$. This constant can be bounded above via $L_n = (u-\delta)^{1/r_n - 1}/r_n < 1/((u-\delta)r_n)$. Because $1/r_n = O(1/n)$, there exists $\kappa > 0$ so that for all $n > n_2$, $n/r_n \leq \kappa$. Put together, we obtain that $n|u^{1/r_n} - u_n^{1/r_n}| \leq nL_n|u - u_n| < \frac{n}{(u-\delta)r_n}|u - u_n| \leq \frac{\kappa}{(u-\delta)}|u - u_n|$. Hence, there exists n_3 so that for all $n > n_3$, $|u - u_n| < \varepsilon(u-\delta)/2\kappa$. Thus $|D^r_n(u) - D^r_n(u_n)| < \varepsilon/2$ if $n > \max(n_2, n_3)$ and $|D^r_n(u) - D(u)| < \varepsilon$ if $n > \max(n_1, n_2, n_3)$. □

Proof of Proposition 2.5. Part (i) follows immediately via the same techniques as Theorem 2.2 (i). For part (ii), Lemma A.3 (ii) yields that $\sup_{x \in \mathbb{R}} |F^\varrho(x) - F^{r_n}(x)| = \sup_{u \in [0,1]} |u^\varrho - u^{\varrho + (r_n - \varrho)}| \leq \frac{3}{\varrho \mathrm{e}}|r_n - \varrho|$. The result now follows along the same steps as in the proof of Theorem 2.2 (ii) with the alternative definitions $u = F^\varrho(x)$ and $u_n = F^{r_n}(x)$; these are possible since F is continuous. The key inequality then becomes $n|u^{1/r_n} - u_n^{1/r_n}| < \frac{n}{r_n(u-\delta)}|u - u_n| = \frac{n}{r_n(u-\delta)}|F^\varrho(x) - F^{r_n}(x)| \leq \frac{n}{r_n(u-\delta)}\frac{3}{\varrho \mathrm{e}}|r_n - \varrho| \to 0$ which implies $D^r_n(u) \to D(u)$ for $u \in \{0,1\}$ and all $u \in (0,1)$ where D is continuous. □

Proof of Corollary 2.8. As in Theorem 2.2 (i), let M^*_n denote the maximum of the first n variables of an iid sequence with marginal distribution function F and (c^*_n), $c^*_n > 0$ and (d^*_n) be such that $(M^*_n - d^*_n)/c^*_n \xrightarrow{\mathrm{d}} H_{\xi,\mu,\sigma}$. From (2) and the Convergence to Types Theorem, we have that $c_n/c^*_{\lceil r_n \rceil} \to c$ and $(d_n - d^*_{\lceil r_n \rceil})/c^*_{\lceil r_n \rceil} \to d$ as $n \to \infty$ for some $c > 0$ and $d \in \mathbb{R}$, and that $G(x) = D \circ H_{\xi,\mu,\sigma}(cx + d)$, $x \in \mathbb{R}$. It is easily seen that $H_{\xi,\mu,\sigma}(cx + d) = H_{\xi,\tilde{\mu},\tilde{\sigma}}(x)$ with $\tilde{\mu} = (\mu - d)/c$ and $\tilde{\sigma} = \sigma/c$. □

Proof of Theorem 2.10. Because D is a strictly increasing, continuous bijection, it has a continuous inverse D^{-1} Royden and Fitzpatrick (2010, p. 53). Hence, we can set $H = D^{-1} \circ G$ and note that the continuity points of H and G coincide. We now need to prove that H is max-stable. By assumption, for each continuity point x of G, $D_n^r\{F^{r_n}(c_n x + d_n)\} \to G(x)$ as $n \to \infty$. Again by assumption, D_n^r is strictly increasing, so it is invertible and $(D_n^r)^{-1} \to D^{-1}$ pointwise on $[0,1]$; see Resnick (1987, Proposition 0.1). By Lemma A.1 this convergence is uniform and hence, by Lemma A.2,

$$\lim_{n\to\infty} F^{r_n}(c_n x + d_n) = \lim_{n\to\infty} (D_n^r)^{-1}\big[D_n^r\{F^{r_n}(c_n x + d_n)\}\big] = D^{-1}\{G(x)\} = H(x) \quad (9)$$

for each continuity point x of H. Therefore, for any $t > 0$ and any continuity point x of H,

$$\lim_{n\to\infty} F^{tr_n}(c_n x + d_n) = H^t(x) \quad (10)$$

as $n \to \infty$. At the same time, Corollary A.4 yields

$$\lim_{n\to\infty} F^{tr_n}(c_{\lceil tn\rceil} x + d_{\lceil tn\rceil}) = \lim_{n\to\infty} \big\{F^{r_{\lceil tn\rceil}}(c_{\lceil tn\rceil} x + d_{\lceil tn\rceil})\big\}^{tr_n/r_{\lceil tn\rceil}} = \{H(x)\}^{t/\lambda(t)}. \quad (11)$$

Applying the convergence to types theorem with $U = H^{t/\lambda(t)}$ and $V = H^t$, to (10) and (11), there exist functions $c(t) > 0$ and $d(t) \in \mathbb{R}$ so that $H^{\lambda(t)}(x) = H\{c(t)x + d(t)\}$, $x \in \mathbb{R}$. Because λ is a bijection, it then follows that for any $t > 0$ and $x \in \mathbb{R}$, $H^t(x) = H\big[c\{\lambda^{-1}(t)\}x + d\{\lambda^{-1}(t)\}\big]$. This means that H is max-stable, and hence, by the classical Fisher–Tippett–Gnedenko theorem, $H \in$ GEV. □

Proof of Corollary 3.7. To prove (i), set $x = H^{-1}(u)$ and $v_n = F^n(c_n x + d_n)$ so that $v_n \to H(x) = u$ as $n \to \infty$. For this x and rate $r(n) = n$, $\mathbb{P}(M_n \le c_n x + d_n) = D_n^r(v_n)$. As in the proof of Theorem 2.2 (ii), we can argue that $|D_n^r(v_n) - D_n^r(u)| \to 0$ as $n \to \infty$, so that $\mathbb{P}(M_n \le c_n x + d_n) \to \gamma$. From Leadbetter (1983, Theorem 2.2) we deduce that there exists $\theta \in [0,1]$ so that $\mathbb{P}(M_n \le c_n z + d_n) \to H^\theta(z)$ for any $z \in \mathbb{R}$. Because $\gamma \in (0,1)$, the degenerate case $\theta = 0$ is excluded. The claim now follows from Theorem 2.2 (ii) and Definition 3.1.

Part (ii) is a direct consequence of Leadbetter et al. (1983, Theorem 3.5.2) and Theorem 2.2 (ii). □

Proof of Proposition 3.10. Under the assumptions of Proposition 3.10, the conditions of Theorem 2.10 are satisfied and hence $G = H^\theta$ for some $H \in$ GEV. This implies that $G \in$ GEV. To show that the distributional mixing condition is violated, pick an arbitrary $x \in \mathbb{R}$ with $G(x) \in (0,1)$, which also means that $H(x) \in (0,1)$. Because λ is not of the form $\lambda(t) = \alpha t$, $t \in (0,1)$, $\alpha > 0$, (Aczél (1987, Proposition 1)) implies that there exist $t_1, t_2 \in (0,1)$ such that $t_1 + t_2 \in (0,1)$ while $\lambda(t_1 + t_2) \neq \lambda(t_1) + \lambda(t_2)$. For large enough n so that $\lceil nt_1\rceil + \lceil nt_2\rceil < n$, set $A = \{1, \ldots, \lceil nt_1\rceil\}$ and $B = \{n - \lceil nt_2\rceil + 1, \ldots, n\}$. The index gap between these two sets is $n - \lceil nt_2\rceil - \lceil nt_1\rceil = O(n)$, and is also eventually larger than s_n for any sequence $(s_n) = o(n)$. Next, recall from the proof of Theorem 2.10 that

(9) holds, which, together with Lemma A.1 and Lemma A.2, implies that $\mathbb{P}(\max_{i\in A} X_i \leq u_n) = \delta_{\lceil nt_1\rceil}\{F(c_n x + d_n)\} = \delta_{\lceil nt_1\rceil}([\{F^{r_n}(c_n x + d_n)\}^{r_{\lceil nt_1\rceil}/r_n}]^{1/r_{\lceil nt_1\rceil}})$ tends to $H^{\theta\lambda(t_1)}(x)$ as $n\to\infty$. Similarly, $\mathbb{P}(\max_{i\in B} X_i \leq u_n) \to H^{\theta\lambda(t_2)}(x)$ as $n\to\infty$. Because the sequence is exchangeable, and $\lceil n(t_1+t_2)\rceil \leq \lceil nt_1\rceil + \lceil nt_2\rceil \leq \lceil n(t_1+t_2)\rceil + 1$ the same argument also implies that

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{P}(\max_{i\in A\cup B} X_i \leq u_n) &\leq \lim_{n\to\infty}\mathbb{P}(\max_{i\in\{1,\dots,\lceil n(t_1+t_2)\rceil\}} X_i \leq u_n) = H^{\theta\lambda(t_1+t_2)}, \quad \text{and that}\\
\lim_{n\to\infty}\mathbb{P}(\max_{i\in A\cup B} X_i \leq u_n) &\geq \lim_{n\to\infty}\mathbb{P}(\max_{i\in\{1,\dots,\lceil n(t_1+t_2)\rceil+1\}} X_i \leq u_n)\\
&\geq \lim_{n\to\infty}\{\mathbb{P}(\max_{i\in\{1,\dots,\lceil n(t_1+t_2)\rceil\}} X_i \leq u_n) - \mathbb{P}(X_{\lceil n(t_1+t_2)\rceil+1} > u_n)\}.
\end{aligned}$$

For the last inequality, note that

$$\mathbb{P}(\max_{i\in\{1,\dots,\lceil n(t_1+t_2)\rceil+1\}} X_i \leq u_n) = \mathbb{P}(\max_{i\in\{1,\dots,\lceil n(t_1+t_2)\rceil\}} X_i \leq u_n, X_{\lceil n(t_1+t_2)\rceil+1} \leq u_n)$$

Subtracting $\mathbb{P}(X_{\lceil n(t_1+t_2)\rceil+1} > u_n) - \mathbb{P}(\max_{i\in\{1,\dots,\lceil n(t_1+t_2)\rceil\}} X_i \leq u_n, X_{\lceil n(t_1+t_2)\rceil+1} > u_n) \geq 0$ from the right hand side leads to the result. Now $\mathbb{P}(X_{\lceil n(t_1+t_2)\rceil+1} > u_n) = 1 - F(u_n)$ and $\lim_{n\to\infty} F(u_n) = \{F^{r_n}(u_n)\}^{1/r_n} = 1$ by (9) and the fact that $r_n\to\infty$ as $n\to\infty$. Because $\lim_{n\to\infty}\mathbb{P}(\max_{i\in\{1,\dots,\lceil n(t_1+t_2)\rceil\}} X_i \leq u_n) = H^{\theta\lambda(t_1+t_2)}$, this implies that $\lim_{n\to\infty}\mathbb{P}(\max_{i\in A\cup B} X_i \leq u_n) \geq H^{\theta\lambda(t_1+t_2)}$. Put together, $\lim_{n\to\infty}|\mathbb{P}(\max_{i\in A\cup B} X_i \leq u_n) - \mathbb{P}(\max_{i\in A} X_i \leq u_n)\mathbb{P}(\max_{i\in B} X_i \leq u_n)| = |H(x)^{\theta\lambda(t_1+t_2)} - H(x)^{\theta\{\lambda(t_1)+\lambda(t_2)\}}|$, which is non-zero since $H(x)\in(0,1)$ and $\theta\neq 0$ as well as $\lambda(t_1+t_2)\neq\lambda(t_1)+\lambda(t_2)$. □

Proof of Theorem 4.4. The convergence $D_n^r(u)\to D(u)$ is immediate for $u=0$ and $u=1$, so consider an arbitrary $u\in(0,1)$. By continuity of ψ, we only need to establish the convergence of $\eta_n\psi^{-1}(u^{1/r_n})$. Setting $x_n = 1/\{r_n(1-u^{1/r_n})\}$, the latter expression can be rewritten as $\psi^{-1}\{1-1/(x_n r_n)\}/\psi^{-1}(1-1/r_n)$. To compute the limit of x_n for $n\to\infty$, note that $r_n\to\infty$ because $\eta_n\to\infty$ as $n\to\infty$ by assumption, and $1-\psi(1/t)\to 0$ as $t\to\infty$. Because $t(1-u^{1/t})\to -\log u$ as $t\to\infty$, we obtain that $x_n\to 1/(-\log u)$ as $n\to\infty$. Utilizing Larsson and G. Nešlehová (2011, Proposition 1 (d)) with $\rho\in(0,1]$, we have that $1-\psi(1/\cdot)\in\mathrm{RV}_{-\rho}$ if and only if $\psi^{-1}(1-1/\cdot)\in\mathrm{RV}_{-1/\rho}$. The uniform convergence theorem for regularly varying functions Bingham et al. (1987, Theorem 1.5.2) implies that $\lim_{n\to\infty}\frac{\psi^{-1}\{1-1/(x_n r_n)\}}{\psi^{-1}(1-1/r_n)} = \{(-\log u)^{-1}\}^{-1/\rho} = (-\log u)^{1/\rho}$. Because ψ is continuous, $\psi\{\eta_n\psi^{-1}(u^{1/r_n})\}\to\psi\{(-\log u)^{1/\rho}\}$ as $n\to\infty$. The remaining part of Theorem 4.4 then follows from Theorem 2.2 (i). □

Proof of Corollary 4.7. The fact that D is a distribution function with density d and quantile function D^{-1} follows immediately from properties of ψ and the logarithm. With the change of variables $(-\log u)^{1/\rho} = y$, one obtains the limit $\lim_{u\to 0} d(u)$ as stated. Finally,

$$\lim_{u\to 1} d(u) = \lim_{y\to 0} -\psi'(y)y^{1-\rho}\mathrm{e}^{(y^\rho)}/\rho = \begin{cases} -\psi'(0), & \text{if } \rho = 1,\\ \lim_{y\to 0} -\psi'(y)y^{1-\rho}/\rho, & \text{if } \rho\in(0,1), \end{cases}$$

as claimed. □

Proof of Proposition 4.12. Given that $r_n = 1/(1-\psi(1/\eta_n))$, we have that $r_{\lceil nt\rceil}/r_n = \frac{1/\{1-\psi(1/\eta_{\lceil nt\rceil})\}}{1/\{1-\psi(1/\eta_n)\}} = \frac{1-\psi(1/\eta_n)}{1-\psi(1/\eta_{\lceil nt\rceil})} = \frac{1-\psi(1/\eta_n)}{1-\psi[1/\{\eta_n(\eta_{\lceil nt\rceil}/\eta_n)\}]}$. Setting $t_n = \eta_{\lceil nt\rceil}/\eta_n$, we have, by assumption, that $t_n \to \kappa(t)$. With $1-\psi(1/\cdot) \in \mathrm{RV}_{-\rho}$, $\rho \in (0,1]$, and $\eta_n \to \infty$ we obtain, by regular variation combined with the uniform convergence theorem for regularly varying functions Bingham et al. (1987, Theorem 1.5.2), that $\lim_{n\to\infty} r_{\lceil nt\rceil}/r_n = \lambda(t)$, where $\lambda(t) = (\kappa(t))^\rho$. Clearly, $\lambda\colon (0,\infty) \to (0,\infty)$ is a bijection, and assumption (ii) of Theorem 4.4 implies that $r_n \to \infty$. Because ψ is strict, the copula diagonals $\delta_n(u) = \psi\{\eta_n\psi^{-1}(u)\}$ as well as the limiting distortion $D(u) = \psi\{(-\log u)^{1/\rho}\}$ are continuous and strictly increasing. Now apply Theorem 2.10. □

Proof of Lemma 4.14. From Herrmann et al. (2023, Theorem 3.3), a distribution function D on $[0,1]$ satisfies $G = D \circ H \in \mathrm{GEV}$ for all $H \in \mathrm{GEV}$ if and only if D is of the form $\mathrm{e}^{-\lambda(-\log u)^\gamma}$ for some $\lambda, \gamma > 0$. Equating this expression to $\psi(-\log u)$ for $u \in (0,1)$ and setting $t = -\log(u)$, we obtain that $\psi(t) = \mathrm{e}^{-\lambda t^\gamma} = \mathrm{e}^{-(ct)^{1/\theta}}$ with $\theta = 1/\gamma$ and $c = \lambda^\theta$. And the function $\mathrm{e}^{-(ct)^{1/\theta}}$ is regularly varying at 0 with index $1/\theta$. However, from Larsson and G. Nešlehová (2011, Remark 2), ψ is convex only if $\theta \geq 1$. □

Proof of Theorem 5.1. Under the assumptions of Theorem 5.1 the conditions of Theorem 2.2 (i) hold and hence $(M_n - d^*_{\lceil r_n\rceil})/c^*_{\lceil r_n\rceil}$ converges weakly to $D \circ H$. By the triangle inequality,

$$\begin{aligned}|\mathbb{P}(M_n \leq c^*_{\lceil r_n\rceil}x + d^*_{\lceil r_n\rceil}) - D(H(x))| &\leq \sup_{x\in\mathbb{R}} |D\{F^{r_n}(c^*_{\lceil r_n\rceil}x + d^*_{\lceil r_n\rceil})\} - D\{H(x)\}| \\ &\quad + \sup_{x\in\mathbb{R}} |D^r_n\{F^{r_n}(c^*_{\lceil r_n\rceil}x + d^*_{\lceil r_n\rceil})\} - D\{F^{r_n}(c^*_{\lceil r_n\rceil}x + d^*_{\lceil r_n\rceil})\}|.\end{aligned}$$

By assumption (ii), the second term on the right-hand side is bounded above by $s(n)$. Using Hölder continuity of D in assumption (iii), the first term is bounded above by $K\sup_{x\in\mathbb{R}} |F^{r_n}(c^*_{\lceil r_n\rceil}x + d^*_{\lceil r_n\rceil}) - H(x)|^\kappa$. If r is integer-valued, $\lceil r_n\rceil = r_n$ for all $n \in \mathbb{N}$, and assumption (i) gives the upper bound $K\{\beta^*(r_n)\}^\kappa$. Otherwise, the supremum in the previous display is bounded above by $\sup_{x\in\mathbb{R}} |F^{\lceil r_n\rceil}(c^*_{\lceil r_n\rceil}x + d^*_{\lceil r_n\rceil}) - H(x)| + \sup_{x\in\mathbb{R}} |F^{r_n}(c^*_{\lceil r_n\rceil}x + d^*_{\lceil r_n\rceil}) - F^{\lceil r_n\rceil}(c^*_{\lceil r_n\rceil}x + d^*_{\lceil r_n\rceil})|$. The first term is at most $\beta^*(\lceil r_n\rceil)$ by assumption (i), while the second is at most $3\mathrm{e}^{-1}/r_n$ by Corollary A.6. □

Proof of Corollary 5.3. In view of the discussion immediately preceding Corollary 5.3, the case of monotonically decreasing β^* and a PQD C_n remains to be considered. The Fréchet–Hoeffding inequality Nelsen (2006, Theorem 2.10.12) and the PQD property imply that $u^n \leq u^{\eta_n} \leq u$ for all $u \in [0,1]$ and hence $1 \leq \eta_n \leq n$. As β^* is monotonically decreasing, $\beta^*(\lceil\eta_n\rceil) \geq \beta^*(n)$. □

Proof of Proposition 5.6. Calling on assumption (i) of Theorem 5.1 and the triangle inequality, we can bound $\sup_{x\in\mathbb{R}}|F^{r_n}(c^*_{\lceil r_n\rceil}x+d^*_{\lceil r_n\rceil})-H(x)|$ from above by $\sup_{x\in\mathbb{R}}|F^{r_n}(c^*_{\lceil r_n\rceil}x+d^*_{\lceil r_n\rceil})-F^{\lceil r_n\rceil}(c^*_{\lceil r_n\rceil}x+d^*_{\lceil r_n\rceil})|+\sup_{x\in\mathbb{R}}|F^{\lceil r_n\rceil}(c^*_{\lceil r_n\rceil}x+d^*_{\lceil r_n\rceil})-H(x)|$, where the second term is at most $\beta^*(\lceil r_n\rceil)$. When $\lceil r_n\rceil = r_n$, the first term disappears; otherwise, it is at most $3/(r_n\mathrm{e})$ by Corollary A.6.

For $u=0$ and $u=1$, we immediately have $D_n^r(u)=D(u)$ for all $n\in\mathbb{N}$. For a given $u\in(0,1)$, pick $x\in\mathbb{R}$ such that $u=H(x)$ and set $u_n=F^{r_n}(c^*_{\lceil r_n\rceil}x+d^*_{\lceil r_n\rceil})$; clearly, such an x exists since $H\in\mathrm{GEV}$ is continuous. Hölder continuity of D_n^r now implies that $\sup_{u\in[0,1]}|D_n^r(u)-D(u)|\leq\sup_{u\in[0,1]}|D_n^r(u)-D_n^r(u_n)|+\sup_{u\in[0,1]}|D_n^r(u_n)-D(u)|\leq\sup_{u\in[0,1]}K_n|u-u_n|^{\kappa_n}+\sup_{x\in\mathbb{R}}|\mathbb{P}(M_n\leq c^*_{\lceil r_n\rceil}x+d^*_{\lceil r_n\rceil})-D\circ H(x)|\leq K_n\big\{\sup_{x\in\mathbb{R}}|F^{r_n}(c^*_{\lceil r_n\rceil}x+d^*_{\lceil r_n\rceil})-H(x)|\big\}^{\kappa_n}+\gamma(n)\leq K\big\{\beta^*(\lceil r_n\rceil)+3\mathrm{e}^{-1}\mathbb{1}_{\{r_n\notin\mathbb{N}\}}/r_n\big\}^{\kappa}+\gamma(n)$. The assumptions on r_n, β^* and γ imply that the right-hand side converges to 0 as $n\to\infty$. □

C Connections to and discussion of the results of Ballerini (1994)

In contrast to Wüthrich (2004) and the results in Section 4.1, the results in Ballerini (1994) are not derived based on a suitable regular variation condition of the Archimedean generator (or its inverse). Instead, comparable results are derived based on the asymptotic polynomial growth condition

$$\lim_{t\to\infty} t^{\rho}\{1-\psi(1/t)\}=c \tag{12}$$

for some $\rho\in(0,\infty)$ and $c\in(0,\infty)$; see Ballerini (1994, Theorem 2). However this condition limits the scope of the derived results. While this is already noted in Ballerini (1994, p. 386), the precise extent of this limitation has not been addressed in the literature. In the following discussion, we show that the regular variation condition $1-\psi(1/\cdot)\in\mathrm{RV}_{-\rho}$ used in Section 4.1 is strictly more general than the asymptotic polynomial growth condition (12) of Ballerini (1994).

Proposition C.1
Let f be a positive measurable function, defined on some neighbourhood $[S,\infty)$ of infinity, and $\rho\in\mathbb{R}$ and $c>0$ be two constants. Then $f\in\mathrm{RV}_{-\rho}$ with associated slowly varying function L such that $\lim_{x\to\infty}L(x)=c$ if and only if $\lim_{x\to\infty}x^{\rho}f(x)=c$.

Proof. For necessity, $f\in\mathrm{RV}_{-\rho}$ can be represented as $f(x)=x^{-\rho}L(x)$ for some slowly varying function L; see Bingham et al. (1987, Theorem 1.4.1). By assumption, we have that $\lim_{x\to\infty}x^{\rho}f(x)=\lim_{x\to\infty}x^{\rho}x^{-\rho}L(x)=\lim_{x\to\infty}L(x)=c$. For sufficiency, if $\lim_{x\to\infty}x^{\rho}f(x)=c$, $c>0$, we have by assumption that for any $\lambda>0$ that, $\lim_{x\to\infty}\frac{f(\lambda x)}{f(x)}=\lim_{x\to\infty}\lambda^{-\rho}\frac{(x\lambda)^{\rho}f(\lambda x)}{x^{\rho}f(x)}=\lambda^{-\rho}$, that is $f\in\mathrm{RV}_{-\rho}$. This implies that f is necessarily of the form $f(x)=x^{-\rho}L(x)$, where L is slowly varying; see Bingham et al. (1987, Theorem 1.4.1). Again by assumption, we then have that $c=\lim_{x\to\infty}x^{\rho}f(x)=\lim_{x\to\infty}L(x)$. □

By Proposition C.1, and as already noted in Example 4.11, the class $\mathrm{RV}_{-\rho}$ is strictly larger than the set of all functions with $\lim_{x\to\infty} x^\rho f(x) = c > 0$, since slowly varying functions with finite non-zero limits are only a subset of all slowly varying functions; a simple example of a slowly varying function L such that $\lim_{x\to\infty} L(x)$ does not exist is $L(x) = \log x$; see Bingham et al. (1987, p. 16) for more examples.

In the context of completely monotone Archimedean generators, the following lemma of Ballerini (1994, Lemma 1) can be used to to derive a useful corollary for the constructing of an example where $1 - \psi(1/\cdot) \in \mathrm{RV}_{-\rho}$ but where the asymptotic polynomial growth condition (12) is violated.

Lemma C.2 (Ballerini (1994), Lemma 1)
Denote by $\psi\colon (0,\infty) \to \mathbb{R}$ a real valued function. Suppose $\psi(t) \geq 0$ on $(0,\infty)$ and $\psi'(t) = -\kappa(t)\psi(t)$ for $\kappa\colon (0,\infty) \to \mathbb{R}$ being completely monotone on $(0,\infty)$. Then ψ is completely monotone on $(0,\infty)$.

Lemma C.2 implies the following method for constructing completely monotone Archimedean generators ψ with given regular variation properties of $1 - \psi(1/\cdot)$.

Corollary C.3
Let $f\colon [0,\infty) \to \mathbb{R}$ be a differentiable function.

(i) Then it holds that $f(0) = 0$, $\lim_{t\to\infty} f(t) = \infty$ and that the derivative $f'\colon (0,\infty) \to \mathbb{R}$ is completely monotone if and only if $\psi(t) = \exp\{-f(t)\}$ is a completely monotone generator of an Archimedean copula such that ψ^α is completely monotone for all $\alpha \in (0,\infty)$.

(ii) Consider a completely monotone Archimedean generator of the form $\psi(t) = \exp\{-f(t)\}$. Then the mean of the associated frailty distribution is given by $\lim_{t\to 0} f'(t)$. If, additionally, $f'(1/\cdot) \in \mathrm{RV}_\beta$, $\beta \in \mathbb{R}$, then $1 - \psi(1/\cdot) \in \mathrm{RV}_{-(1-\beta)}$.

Proof.

(i) Consider sufficiency. If $\psi(t) = \exp\{-f(t)\}$, we have $f(t) = -\log \psi(t)$, which directly yields $f(0) = 0$ and $\lim_{t\to\infty} f(t) = \infty$. For the derivative we have $f'(t) = -\psi'(t)/\psi(t) = \{-\log\psi(t)\}'$. By Hofert (2010, Proposition 2.1.5 (5)), f' is completely monotone if and only if ψ^α is completely monotone for all $\alpha \in (0,\infty)$, which holds by assumption.

Now consider necessity. By construction, we have $\psi(0) = 1$ and $\lim_{t\to\infty} \psi(t) = 0$. Furthermore, we have $\psi'(t) = -\exp\{-f(t)\}f'(t) = -\psi(t)f'(t)$. Since, by assumption, f' is completely monotone, Lemma C.2 guarantees that ψ is completely monotone. By Kimberling (1974b), ψ is therefore the generator of an Archimedean copula.

(ii) The mean of the associated frailty distribution is given by $\lim_{t\to 0} -\psi'(t) = \lim_{t\to 0}\psi(t) \cdot f'(t) = \lim_{t\to 0} f'(t)$ since $\psi(0) = 1$. For the regular variation property, we first compute $\frac{\mathrm{d}}{\mathrm{d}t}\psi\{1/(\lambda t)\} = -\psi'\{1/(\lambda t)\}t^{-2}\lambda^{-1}$. Using L'Hôpital's rule and $\psi(0) = 1$, we have $\lim_{t\to\infty} \frac{1-\psi\{1/(\lambda t)\}}{1-\psi(1/t)} = \frac{1}{\lambda}\lim_{t\to\infty}\frac{\psi'\{1/(\lambda t)\}}{\psi'(1/t)} = \frac{1}{\lambda}\lim_{t\to\infty}\frac{\psi\{1/(\lambda t)\}f'\{1/(\lambda t)\}}{\psi(1/t)f'(1/t)} = \frac{1}{\lambda}\lim_{t\to\infty}\frac{f'\{1/(\lambda t)\}}{f'(1/t)}$ for $\lambda > 0$. By assumption, $f'(1/\cdot) \in \mathrm{RV}_\beta$ and so $\lim_{t\to\infty}\frac{1-\psi\{1/(\lambda t)\}}{1-\psi(1/t)} = \lambda^{-1}\lambda^\beta = \lambda^{-1+\beta}$, that is $1 - \psi(1/\cdot) \in \mathrm{RV}_{-(1-\beta)}$. □

We now use Corollary C.3 in the following example to construct a regularly varying Archimedean generator which is not covered by the framework of Ballerini (1994). We therefore close a gap in the literature since the proposed construction of Ballerini (1994, Example 3) based on $\psi_\beta(t) = \exp\big[-\{\log(1+t)\}^{1/\beta}\big], \quad \beta \geq 1$, actually yields (via L'Hôspital's rule) that $\lim_{t\to\infty} t^{1/\beta}\{1-\psi(1/t)\} = 1$ and not that $\lim_{t\to\infty} t^{1/\beta}\{1-\psi(1/t)\} = 0$ as claimed.

Example C.4

Consider the function $f(x) = (1+x)\log(1+1/x)+\log(x)$, $x \in (0,\infty)$, which satisfies $f(x) \geq 0$, $\lim_{x\to 0} f(x) = 0$, $\lim_{x\to\infty} f(x) = \infty$ and $f'(x) = \log(1+1/x)$. Also note that f' is completely monotone. By Corollary C.3, the function $\psi(t) = \exp\{-f(t)\} = 1/\{t(1+1/t)^{1+t}\}$ is therefore a completely monotone Archimedean generator. Furthermore, we have $f'(1/\cdot) \in \mathrm{RV}_0$, that is $f'(1/\cdot)$ is slowly varying; this can be seen from L'Hôpital's rule via $\lim_{x\to\infty} \frac{f'\{1/(\lambda x)\}}{f'(1/x)} = \lim_{x\to\infty} \frac{\log(1+\lambda x)}{\log(1+x)} = \lambda \lim_{x\to\infty} \frac{1+x}{1+\lambda x} = 1$. By Corollary C.3, $1-\psi(1/\cdot) \in \mathrm{RV}_{-1}$. While it is easily seen that $-\psi'(0) < \infty$ implies $\rho = 1$, it is interesting to note that here we have $-\psi'(0) = f'(0) = \infty$ even though $\rho = 1$, demonstrating that the converse implication is not true in general; see also Table 1 and the example provided by Family 23 of Charpentier and Segers (2009, Table 1). More importantly, considering the limit condition in Ballerini (1994), we have that $\lim_{t\to\infty} t^\beta\{1-\psi(1/t)\} \in \{0,\infty\}$ for $\beta \neq 1$, and, for $\beta = 1$ via L'Hôpital's rule, that $\lim_{t\to\infty} t\{1-\psi(1/t)\} = \lim_{t\to\infty} \frac{1}{t^{-1}}\{1 - \frac{t}{(1+t)^{1+1/t}}\} = \lim_{t\to\infty} \frac{-(1+t)^{-1-1/t}\log(1+t)/t}{-t^{-2}} = \lim_{t\to\infty} \frac{t\log(1+t)}{(1+t)^{1+1/t}} = \infty$. Hence the generator $\psi(t) = 1/\{t(1+1/t)^{1+t}\}$ does not satisfy the asymptotic polynomial growth condition (12) of the framework of Ballerini (1994).

As a final step we can compare the limiting distributions from Ballerini (1994) with ours from Theorem 4.4. In contrast to Theorem 4.4, the limiting distributions in Ballerini (1994, Equation (2.2)) reference a positive parameter $c > 0$ linked to the asymptotic polynomial growth condition (12), where ρ is the coefficient of regular variation of $1-\psi(1/\cdot)$. The following remark shows that in our more general framework, this dependence is in fact an artifact that can be interpreted as a rescaling of the copula generator ψ without an effect on the final dependence structure. We illustrate this point with a detailed example in the remaining part of this supplement.

Remark C.5

An Archimedean generator ψ_1 is not uniquely identified in the sense that the generator $\psi_2(t) = \psi_1(ct)$ generates the same copula for all $c > 0$. Our results in Theorem 4.4 about the limiting distribution of M_n in the Archimax case should therefore not depend on the rescaling factor c. However, following the same steps as in Theorem 4.4, it seems counter-intuitive at first that the limiting distribution function D depends on c. Specifically, for $1-\psi_1(1/\cdot) \in \mathrm{RV}_{-\rho}$, consider $\psi_2(t) = \psi_1(\tilde{c}t)$ with $\tilde{c} = c^{-1/\rho}$, where the power $-1/\rho$ of c allows us to connect our discussion to the results derived in Ballerini (1994). Verifying that then also $1-\psi_2(1/\cdot) \in \mathrm{RV}_{-\rho}$, it follows that $\lim_{n\to\infty} D_n^r(u) = \lim_{n\to\infty} \psi_2\{\eta_n \psi_2^{-1}(u^{1/r_n})\} = \psi_1\{\tilde{c}(-\log u)^{1/\rho}\} = \psi_1\{c^{-1/\rho}(-\log u)^{1/\rho}\}$, where the limit on the right-hand side is equal to the distortion implicitly derived in the Archimedean case in Ballerini (1994, Equation (2.2)). To reconcile the fact that ψ_1 leads to the limit $D_1(u) = \psi_1\{(-\log u)^{1/\rho}\}$, while ψ_2 leads to the limit $D_2(u) = \psi_1\{c^{-1/\rho}(-\log u)^{1/\rho}\} = \psi_1\{(-\log u^{1/c})^{1/\rho}\}$, it is important to remember that in the Archimedean case, the stabilizing constants (c_n) and (d_n) depend on the underlying generator. This leads to different pairs of stabilizing sequences for M_n for ψ_1 and ψ_2, and hence implies by the convergence to types theorem, see Resnick (1987, Proposition 0.2), that the limiting distributions are related by an affine transformation. That this is indeed the case can be seen from the fact that the two limits differ by a power transformation of the argument, that is $D_1(u) = D_2(u^c)$. In case of a max-stable distribution as argument, this indeed leads to an affine shift. As discussed in Remark 2.9, for a max-stable distribution function H the constants $a > 0$ and $b \in \mathbb{R}$ such that $H^c(x) = H(ax + b)$ can be calculated explicitly, which hence precisely determines the impact of using stabilizing sequences connected to either ψ_1 or ψ_2. Via the convergence to types theorem, also their relation to each other can be determined.

The following example illustrates this point.

Example C.6

Denote by $F_\lambda(x) = 1 - \mathrm{e}^{-\lambda x}$, $x \geq 0$, the distribution function of an exponential distribution with parameter $\lambda > 0$. Then possible iid stabilizing constants are $c_n^* = 1/\lambda$ and $d_n^* = \log(n)/\lambda$, leading to $\lim_{n\to\infty} F_\lambda^n\,(c_n^* x + d_n^*) = \mathrm{e}^{-\mathrm{e}^{-x}} = \Lambda(x)$, that is $F_\lambda \in \mathrm{MDA}(\Lambda)$. Now consider a sequence (X_i) where the margins are exponentially distributed and the copula C_θ of $(X_1, \ldots, X_n)$ is a Clayton copula with parameter $\theta > 0$. In this case the generator takes the form $\psi(t) = (1 + t)^{-1/\theta}$, and, for $c > 0$, the scale-transformed version is given by $\psi_c(t) = (1 + ct)^{-1/\theta}$. In this case we have $\rho = 1$ for ψ and ψ_c. At the same time, we have $-\psi_c'(0) = -c\psi'(0) = c/\theta < \infty$, and hence, as discussed in Example 4.9, we can use the iid stabilizing constants (c_n^*) and (d_n^*) to obtain $\lim_{n\to\infty} \mathbb{P}\big(\frac{M_n - d_n^*}{c_n^*} \leq x\big) = \psi_c\{-\log \Lambda^{-1/\psi_c'(0)}(x)\} = \psi\{-c \log \Lambda^{\theta/c}(x)\} = \psi\{-\log \Lambda^{\theta}(x)\}$, where, as expected, the result is invariant with respect to the value of $c > 0$ since ψ and ψ_c induce the same copula C_θ.

However, when using the stabilization implied by Theorem 4.4, the result seemingly depends on the value of $c > 0$. In this case, we have for $r_n = 1 - \psi_c(1/n)$ the asymptotic expansion $r_n = \frac{1}{1-\psi_c(1/n)} = \frac{\theta n}{c} + \frac{\theta+1}{2} - \frac{c(\theta^2-1)}{12\theta n} + \mathcal{O}\left(n^{-2}\right)$, which, for large n, suggests the approximation $r_n \approx \theta n/c$. With $\lceil r_n \rceil = \lceil n\theta/c \rceil = c_n(n\theta/c)$, where $c_n = \lceil n\theta/c \rceil\, c/(n\theta) \to 1$, this leads to the (asymptotic) stabilizing constants $c^*_{\lceil r_n \rceil} = 1/\lambda$ and $d^*_{\lceil r_n \rceil} = \log(\lceil r_n \rceil)/\lambda = \log(n)/\lambda + \log(\theta)/\lambda - \log(c)/\lambda + \log(c_n)/\lambda$. By Theorem 4.4, we then have for $x \in \mathbb{R}$ the weak limit $\lim_{n\to\infty} \mathbb{P}\big(\frac{M_n - d^*_{\lceil r_n \rceil}}{c^*_{\lceil r_n \rceil}} \leq x\big) = \psi_c\{-\log \Lambda(x)\} = \psi\{-\log \Lambda^c(x)\}$. The key to understanding the dependence on c is to realize that in this case the stabilizing constants $c^*_{\lceil r_n \rceil}$ and $d^*_{\lceil r_n \rceil}$ also depend on $c > 0$. As such, the appearance of $c > 0$ in the limit is a consequence of the convergence to types theorem (see Resnick (1987, Proposition 0.2)): First noting that $\Lambda^c(x) = \Lambda(x - \log c)$, we also have $c^*_{\lceil r_n \rceil}/c_n^* = 1$

and $(d^*_{\lceil r_n \rceil} - d^*_n)/c^*_n \to \log(\theta) - \log(c)$. The two different limits are now connected via $\psi\big[-\log \Lambda^\theta\{x + \log(\theta) - \log(c)\}\big] = \psi\big[-\log \Lambda\{x - \log(c)\}\big] = \psi\{-\log \Lambda^c(x)\}$, and the appearance of $c > 0$ in the second limit can indeed be attributed to the utilized stabilizing constants.

References

Aczél, J. (1987), A short course on functional equations, D. Reidel Publishing Company.

Anderson, C. W., Coles, S. G., and Hüsler, J. (1997), Maxima of Poisson-like variables and related triangular arrays, *The Annals of Applied Probability*, 7(4), 953–971.

Arellano-Valle, R. B. and Genton, M. G. (2008), On the exact distribution of the maximum of absolutely continuous dependent random variables, *Statistics & Probability Letters*, 78(1), 27–35.

Balakrishnan, N., Bendre, S. M., and Malik, H. J. (1992), General relations and identities for order statistics from non-independent non-identical variables, *Ann. Inst. Statist. Math.*, 44(1), 177–183.

Ballerini, R. (1994), Archimedean copulas, exchangeability, and max-stability, *Journal of Applied Probability*, 31(2), 383–390.

Beirlant, J., Goegebeur, Y., Teugels, J. L., and Segers, J. (2004), Statistics of Extremes: Theory and Applications, 1st ed., John Wiley & Sons, Ltd.

Belzile, Léo Raymond and Nešlehová, Johanna G. (2017), Extremal attractors of Liouville copulas, Submitted.

Berman, S. M. (1962a), Equally correlated random variables, *Sankhya, Ser. A*, 24, 155–156.

Berman, S. M. (1962b), Limiting distribution of the maximum term in sequences of dependent random variables, *The Annals of Mathematical Statistics*, 33(3), 894–908.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1987), Regular Variation, 1st ed., Cambridge University Press.

Buriticá, G., Meyer, N., Mikosch, T., and Wintenberger, O. (2021-05), Some variations on the extremal index, *Zap. Nauchn. Semin. POMI.*

Capéraà, Philippe, Fougères, Anne-Laure, and Genest, Christian (2000), Bivariate distributions with given extreme value attractor, *J. Multivariate Anal.*, 72, 30–49.

Charpentier, A., Fougères, A.-L., Genest, C., and G. Nešlehová, J. (2014), Multivariate Archimax copulas, *Journal of Multivariate Analysis*, 126, 118–136.

Charpentier, A. and Segers, J. (2009), Tails of multivariate Archimedean copulas, *J. Multivariate Anal.*, 100, 1521–1537.

Chatelain, Simon, Fougères, Anne-Laure, and Nešlehová, Johanna G. (2020), Inference for Archimax copulas, *Ann. Statist.*, 48(2), 1025–1051, ISSN: 0090-5364, `doi:10.1214/19-AOS1836`, `https://doi.org/10.1214/19-AOS1836`.

Chavez-Demoulin, V. and Guillou, A. (2018), Extreme quantile estimation for β-mixing time series and applications, *Insurance: Mathematics and Economics*, 83, 59–74.

Cohen, Jonathan P. (1982), The Penultimate Form of Approximation to Normal Extremes, *Advances in Applied Probability*, 14(2), 324–339, ISSN: 00018678, http://www.jstor.org/stable/1426524 (2024-01-11).

Cuadras, C. M. and Augé, J. (1981), A continuous general multivariate distribution and its properties, *Communications in Statistics - Theory and Methods*, 10(4), 339–353.

Darsow, W. F., Nguyen, B., and Olsen, E. T. (1992), Copulas and Markov processes, *Illinois journal of mathematics*, 36(4), 600–642.

Deheuvels, P. (1983), Point processes and multivariate extreme values, *Journal of multivariate analysis*, 13(2), 257–272.

Dudziński, M. and Furmańczyk, K. (2017), Some applications of the Archimedean copulas in the proof of the almost sure central limit theorem for ordinary maxima, *Open Mathematics*, 15(1), 1024–1034.

Durante, F. and Mai, J.-F. (2010), Representation of Exchangeable Sequences by Means of Copulas, *Combining Soft Computing and Statistical Methods in Data Analysis, SMPS 2010, Oviedo, Spain, September 29 - October 1, 2010*, ed. by C. Borgelt et al., vol. 77, Advances in Intelligent and Soft Computing, Springer, 227–232.

Durante, F. and Sempi, C. (2016), Principles of copula theory, vol. 474, CRC press Boca Raton, FL.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997), Modelling Extremal Events for Insurance and Finance, 1st ed., Springer.

Falk, M. (2019), Multivariate Extreme Value Theory and D-Norms, 1st ed., Springer.

Fawcett, L. and Walshaw, D. (2012), Estimating return levels from serially dependent extremes, *Environmetrics*, 23(3), 272–283.

Feidt, A., Genest, C., and Nešlehová, J. (2010), Asymptotics of joint maxima for discontinuous random variables, *Extremes*, 13, 35–53.

Feller, W. (1971), An Introduction to Probability Theory and Its Applications, 2nd ed., vol. 2, Wiley.

Ferreira, H. and Ferreira, M. (2018), Estimating the extremal index through local dependence, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 54(2), 587–605, doi:10.1214/16-AIHP815, https://doi.org/10.1214/16-AIHP815.

Ferreira, M. S. (2018), Heuristic tools for the estimation of the extremal index: A comparison of methods, *REVSTAT: Statistical Journal*, 16(1), 115–136.

Fisher, R. A. and Tippett, L. H. C. (1928), Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Mathematical proceedings of the Cambridge philosophical society*, vol. 24, (2), Cambridge University Press, 180–190.

Fredricks, G. A. and Nelsen, R. B. (1997), Copulas constructed from diagonal sections, *Distributions with Given Marginals and Moment Problems*, ed. by V. Beneš and J. Štěpán, Dordrecht: Springer, 129–136.

Gnedenko, B. (1943), Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire, *Annals of Mathematics*, 44(3), 423–453, (2023-09-29).

Gudendorf, G. and Segers, J. (2010), Extreme-value copulas, *Copula Theory and Its Applications*, ed. by P. Jaworski, F. Durante, W. Härdle, and T. Rychlik, vol. 198, Lecture Notes in Statistics, Springer, Berlin, Heidelberg, 127–145.

Haan, L. de (1984), A spectral representation for max-stable processes, *The Annals of Probability*, 12(4), 1194–1204.

Haan, L. de and Ferreira, A. (2006), Extreme Value Theory. An introduction. 1st ed., Springer.

Haan, L. de, Mercadier, C., and Zhou, C. (2016), Adapting extreme value statistics to financial time series: dealing with bias and serial dependence, *Finance and Stochastics*, 20(2), 321–354.

Hall, P. (1979), On the rate of convergence of normal extremes, *Journal of Applied Probability*, 16(2), 433–439.

Hall, W. J. and Wellner, J. A. (1979), The rate of convergence in law of the maximum of an exponential sample, *Statistica Neerlandica*, 33(3), 151–154.

Herrmann, K., Hofert, M., and G. Nešlehová, J. (2023), Property preserving distortions of max-infinite divisible and generalized extreme value distributions, *Working paper*.

Herrmann, K., Hofert, M., and G. Nešlehová, J. (2024), Supplement to "Limiting Behavior of Maxima under Dependence", *Supplementary Material*.

Hofert, M. (2010), Sampling Nested Archimedean Copulas with Applications to CDO Pricing, PhD thesis, Südwestdeutscher Verlag für Hochschulschriften AG & Co. KG, ISBN 978-3-8381-1656-3.

Hofert, M. (2011), Efficiently sampling nested Archimedean copulas, *Computational Statistics & Data Analysis*, 55, 57–70, `doi:10.1016/j.csda.2010.04.025`.

Hofert, M., Mächler, M., and McNeil, A. J. (2013), Archimedean copulas in high dimensions: Estimators and numerical challenges motivated by financial applications, *Journal de la Société Française de Statistique*, 154(1), 25–63.

Hofert, M., Prasad, A., and Zhu, M. (2022), Multivariate time-series modeling with generative neural networks, *Econometrics and Statistics*, 23, 147–164, `doi:10.1016/j.ecosta.2021.10.011`.

Hofert, M. and Scherer, M. (2011), CDO pricing with nested Archimedean copulas, *Quantitative Finance*, 11(5), 775–787.

Huang, C.-K., North, D., and Zewotir, T. (2017), Exchangeability, extreme returns and Value-at-Risk forecasts, *Physica A: Statistical Mechanics and its Applications*, 477, 204–216, ISSN: 0378-4371.

Huang, X. (1992), Statistics of bivariate extreme values, PhD thesis, Tinbergen Institute Research Series, The Netherlands.

Jaworski, P. (2009), On copulas and their diagonals, *Information Sciences*, 179, 2863–2871.

Kimberling, C. H. (1974a), A Probabilistic Interpretation of Complete Monotonicity, *Aequationes Math.*, 10, 152–164.

Kimberling, C. H. (1974b), A probabilistic interpretation of complete monotonicity, *Aequationes Math*, (10), 152–164.

Kratz, M. and Rootzén, H. (1997), On the rate of convergence for extremes of mean square differentiable stationary normal processes, *Journal of Applied Probability*, 34(4), 908–923.

Larsson, M. and G. Nešlehová, J. (2011), Extremal behavior of Archimedean copulas, *Adv. Appl. Prob.*, 43, 195–216.

Leadbetter, M. R. (1974), On extreme values in stationary sequences, *Probability theory and related fields*, 28(4), 289–303.

Leadbetter, M. R. (1983), Extremes and local dependence in stationary sequences, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65(2), 291–306.

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983), Extremes and related properties of random sequences and processes, 1st, Springer-Verlag.

Leadbetter, M. R. and Rootzén, H. (1988), Extremal Theory for Stochastic Processes, *The Annals of Probability*, 16(2), 431–478, (2022-06-23).

Li, H. (2008), Tail dependence comparison of survival Marshall–Olkin copulas, *Methodology and Computing in Applied Probability*, 10(1), 39–54.

Liebscher, E. (2008), Construction of asymmetric multivariate copulas, *Journal of Multivariate Analysis*, 99(10), 2234–2250.

Loynes, R. M. (1965), Extreme values in uniformly mixing stationary stochastic processes, *The Annals of Mathematical Statistics*, 36(3), 993–999.

Mai, J.-F. (2018), Extreme-value copulas associated with the expected scaled maximum of independent random variables, *Journal of Multivariate Analysis*, 166, 50–61, ISSN: 0047-259X, doi:10.1016/j.jmva.2018.02.005, sciencedirect.com/science/article/pii/S0047259X17307467.

Mai, J.-F. (2019), Canonical spectral representation for exchangeable max-stable sequences, *Extremes*, ISSN: 1572-915X, doi:10.1007/s10687-019-00361-3.

Mai, J.F. and Scherer, M. (2009), Efficiently sampling exchangeable Cuadras–Augé copulas in high dimensions, *Information Sciences*, 179(17), 2872–2877.

Marshall, Albert W. and Olkin, Ingram (1988), Families of multivariate distributions, *J. Amer. Statist. Assoc.*, 83, 834–841.

Mazo, G., Girard, S., and Forbes, F. (2015), A class of multivariate copulas based on products of bivariate copulas, *Journal of Multivariate Analysis*, 140, 363–376.

McCormick, W. P. and Seymour, L. (2001), Rates of convergence and approximations to the distribution of the maximum of chain-dependent sequences, *Extremes*, 4(1), 23–52.

McNeil, A., Frey, R., and Embrechts, P. (2015), Quantitative Risk Management: Concepts, Techniques and Tools, 2nd ed., Princeton University.

McNeil, A. J. and Nešlehová, J. (2009), Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions, *The Annals of Statistics*, 37(5b), 3059–3097.

Mehr, R. I. and Cammack, E. (1980), Principles of insurance, 7th ed., Richard D. Irwin, INC.

Mulinacci, S. (2015), Marshall–Olkin machinery and power mixing: The mixed generalized Marshall–Olkin distribution, *Marshall–Olkin Distributions - Advances in Theory and Applications*, ed. by U. Cherubini, F. Durante, and S. Mulinacci, Cham: Springer, 65–86.

Nagler, T., Krüger, D., and Min, A. (2022), Stationary vine copula models for multivariate time series, *Journal of Econometrics*, 227(2), 305–324.

Nair, K. Aiyappan (1981), Asymptotic Distribution and Moments of Normal Extremes, *The Annals of Probability*, 9(1), 150–153, ISSN: 00911798, http://www.jstor.org/stable/2243186 (2024-01-11).

Nelsen, R. (2006), An Introduction to Copulas, 2nd ed., Springer.

Newell, G. F. (1964), Asymptotic extremes for m-dependent random variables, *The Annals of Mathematical Statistics*, 35(3), 1322–1325.

Pancheva, E. (1985), Limit theorems for extreme order statistics under nonlinear normalization, *Stability Problems for Stochastic Models. 1985: Proceedings of the International Seminar Held in Varna, Bulgaria, May 13-19, 1985*, ed. by V.V. Kalasnikov and V.M. Zolotarev, Springer, 284–309.

Peng, Zuoxiang, Nadarajah, Saralees, and Lin, Fuming (2010), Convergence Rate of Extremes for the General Error Distribution, *Journal of Applied Probability*, 47(3), 668–679, doi:10.1239/jap/1285335402.

Pereira, H. I. (1983), Rate of convergence towards a Fréchet type limit distribution, *Annales scientifiques de l'Université de Clermont-Ferrand 2. Série Probabilités et applications*, 76(1), 67–80.

Resnick, S. I. (1987), Extreme Values, Regular Variation and Point Processes, 1st ed., Springer-Verlag New York.

Ressel, P. (2013), Homogeneous distributions - And a spectral representation of classical mean values and stable tail dependence functions, *Journal of Multivariate Analysis*, 117, 246–256.

Rootzén, H. (1983), The Rate of Convergence of Extremes of Stationary Normal Sequences, *Advances in Applied Probability*, 15(1), 54–80, (2022-06-23).

Royden, H. L. and Fitzpatrick, P. (2010), Real analysis, 4th, Pearson.

Russell, B. T. and Huang, W. K. (2021), Modeling short-ranged dependence in block extrema with application to polar temperature data, *Environmetrics*, 32(3).

Rychlik, T. (1994), Distributions and expectations of order statistics for possibly dependent random variables, *Journal of Multivariate Analysis*, 48, 31–42.

Schönbucher, P. J. and Schubert, D. (2001), Copula-Dependent Default Risk in Intensity Models, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=301968 (2009-12-30).

Schweizer, B. and Sklar, A. (1983), Probabilistic Metric Spaces, New York: North-Holland.

Sklar, A. (1959), Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, 8, 229–231.

Smith, R. L. (1990), Max-stable processes and spatial extremes, 24, Preprint Univ. North Carolina, http://www.stat.unc.edu/faculty/rs/papers/RLSPapers.html (2019-02-21).

Watson, G. S. (1954), Extreme values in samples from m-dependent stationary stochastic processes, *The Annals of Mathematical Statistics*, 798–800.

Wüthrich, M. V. (2004), Extreme value theory and Archimedean copulas, *Scandinavian Actuarial Journal*, 2004(3), 211–228.