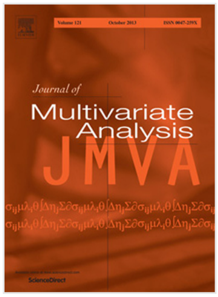
Accepted Manuscript



Properties of extremal dependence models built on bivariate max-linearity

Mónika Kereszturi, Jonathan Tawn

PII: S0047-259X(16)30215-9
DOI: http://dx.doi.org/10.1016/j.jmva.2016.12.001
Reference: YJMVA 4195

To appear in: *Journal of Multivariate Analysis*

Received date: 18 May 2016

Please cite this article as: M. Kereszturi, J. Tawn, Properties of extremal dependence models built on bivariate max-linearity, *Journal of Multivariate Analysis* (2016), http://dx.doi.org/10.1016/j.jmva.2016.12.001

Properties of extremal dependence models built on bivariate max-linearity

Mónika Kereszturi* and Jonathan Tawn

STOR-i Centre for Doctoral Training, Lancaster University, UK

Abstract

Bivariate max-linear models provide a core building block for characterizing bivariate max-stable distributions. The limiting distribution of marginally normalized component-wise maxima of bivariate max-linear models can be dependent (asymptotically dependent) or independent (asymptotically independent). However, for modeling bivariate extremes they have weaknesses in that they are exactly max-stable with no penultimate form of convergence to asymptotic dependence, and asymptotic independence arises if and only if the bivariate max-linear model is independent. In this work we present more realistic structures for describing bivariate extremes. We show that these models are built on bivariate max-linearity but are much more general. In particular, we present models that are dependent but asymptotically independent and others that are asymptotically dependent but have penultimate forms. We characterize the limiting behavior of these models using two new different angular measures in a radial-angular representation that reveal more structure than existing measures.

Keywords: Bivariate extremes; max-linear models; extremal dependence; asymptotic independence.

*m.kereszturi@lancaster.ac.uk

1 Introduction

When modeling extremes of spatial environmental processes we often care about both local dependence and long-range dependence. For example, in an oceanographic application, we would be interested in the relationship between extreme significant wave heights at two locations that might be close by or located far apart. In particular, we want to know how likely it is that both locations are affected by the same storm and have high waves simultaneously; see, e.g., [8]. Since interest lies in the extremes, the standard measures of spatial dependence are not appropriate and alternative dependence measures and models should be used. Here we introduce a family of bivariate distributions, with simple multivariate extensions, that exhibits all the required features of short, medium and long range extremal dependence for spatial applications. This family is shown to capture all possible bivariate distributions with these properties. We propose novel bivariate characterizations of the extremal dependence structure that reveal structure of this family of distributions that standard measures of extremal dependence fail to identify.

First we identify the two core extremal dependence measures. Let X and Y be identically distributed random variables. Then, an intuitive measure of extremal dependence is the tail dependence measure χ, which is defined as the limiting probability that Y is extreme given that X is extreme,

$$\chi = \lim_{z \to z_F} \Pr(Y > z \mid X > z), \tag{1}$$

where z_F is the upper end point of the common marginal distribution. When $\chi > 0$, X and Y are said to be asymptotically dependent (AD) and the value of χ signifies the strength of asymptotic dependence. This means that X and Y can be extreme simultaneously. However, when the variables are asymptotically independent (AI), $\chi = 0$ and hence χ does not contain any information about the sub-asymptotic dependence structure. Coles et al. [1] argue that to give a more complete summary of extremal dependence a second measure is needed to describe the rate of convergence of $\Pr(Y > z \mid X > z)$ to 0. A useful tail dependence measure can be obtained from the Ledford and Tawn [12] joint tail dependence model, which states that

$$\Pr(X > z, Y > z) = \mathcal{L}\{1/\Pr(X > z)\}\{\Pr(X > z)\}^{2/(\bar{\chi}+1)}, \tag{2}$$

where $\mathcal{L}$ is a slowly varying function at infinity and $\bar{\chi} \in (-1, 1]$. The exponent $2/(\bar{\chi}+1)$ determines the decay rate of the joint probability, with smaller $\bar{\chi}$ giving more rapid convergence of χ to 0. The pair ($\chi > 0$; $\bar{\chi} = 1$) signifies AD, for which the value of χ gives a measure of strength of dependence; and ($\chi = 0$; $\bar{\chi} < 1$) signifies AI, for which the value of $\bar{\chi}$ gives the strength of dependence.

Both the dependence measures χ and $\bar{\chi}$, in expressions (1) and (2), are invariant to the marginal distribution. Of course, using the concept of copulas, all dependence measures can be expressed independently of the marginal distributions. However, for some choices of marginal distributions extremal dependence structure properties are more simply expressed than for other marginal choices. For example, much of the traditional multivariate extreme value theory results are expressed for Fréchet marginals, as they lead to the cleanest expressions of results for component-wise maxima and multivariate regular variation [17]. This marginal choice is fine when the variables are AD, however for AI variables this selection leads to an identical limit form whatever the nature of the AI, i.e., whatever $\bar{\chi} < 1$. For AI variables, [7], [9] and [20] all identify that non-degenerate limit distributions, under affine transformations, can be obtained using exponential margins/tails, whereas under their formulations the limits are degenerate for Fréchet margins. Furthermore, in exponential margins results for AD are also non-degenerate. The reason for this extra flexibility in exponential margins is that an affine transformation in that space is a complex non-linear transformation in Fréchet margins; see Section 2.2 of [15]. Therefore, we work in exponential margins to illustrate our novel extremal dependence characterizations and show that if Fréchet margins had been used, the structure we find would not have been apparent using affine transformations.

In the analysis of multivariate data, it is often difficult to make a choice between AD and AI; see, e.g., [3], [11] and [18]. By having a model that has both AD and AI components, we can avoid having to make this key decision. Wadsworth and Tawn [19] combine a max-stable process with an inverted max-stable process to construct a hybrid spatial dependence model. This model can capture both the AD and AI dependence structure but it is restricted in its forms of AD and AI that can be modeled. Here we use the core structure of the Wadsworth and Tawn [19] model as a basis for exploring bivariate extreme value modeling in a new light. Specifically, we develop a distribution that contains both AD and AI components and has the flexibility to capture all dependence forms within very broad classes in each case.

We construct our model using the multivariate max-linear model [2] as the building block. This class of distributions is both mathematically elegant and was the starting point for understanding the formulation of multivariate extremes [16]. In the bivariate case with Fréchet marginal variables X_F and Y_F, the max-linear model takes the following form:

$$X_F = \max_{i=1,\dots,m}(\alpha_i Z_i), \quad Y_F = \max_{i=1,\dots,m}(\beta_i Z_i), \tag{3}$$

where $\alpha_i, \beta_i \in [0,1]$ for all i, m can be finite or infinite, $\Sigma_{i=1}^m \alpha_i = 1$, $\Sigma_{i=1}^m \beta_i = 1$, and $Z_i \sim$ i.i.d. Fréchet, $i = 1,\dots,m$, with distribution function $F_Z(z) = \exp(-1/z)$ for $z > 0$ and density denoted $f_Z(z)$. This model has joint distribution function

$$\Pr(X_F < x, Y_F < y) = \exp\left\{-\sum_{i=1}^m \max\left(\frac{\alpha_i}{x}, \frac{\beta_i}{y}\right)\right\}, \text{ for } x > 0, y > 0,$$

and it is straightforward to show that this satisfies max-stability, since for any $n > 0$, $x > 0$ and $y > 0$,

$$\Pr(X_F < nx, Y_F < ny)^n = \Pr(X_F < x, Y_F < y).$$

Fundamental to our approach is that Deheuvels [5] shows that every multivariate extreme value distribution for minima, with exponential marginals (i.e., with variables (X_F^{-1}, Y_F^{-1})), can be arbitrarily well approximated by a multivariate max-linear model. Fougères et al. [6] showed this property holds for (X_F, Y_F), as well as presenting a broader discussion on alternative representations of multivariate extreme value distributions.

Our paper introduces two bivariate distributions, with exponential margins, that are derived from the max-linear model (3) with Fréchet margins: these are the transformed max-linear model and the inverted max-linear model, denoted by (X_E, Y_E) and $(X_E^{(I)}, Y_E^{(I)})$ respectively. Specifically,

$$(X_E, Y_E) = (-\ln\{1 - \exp(-1/X_F)\}, -\ln\{1 - \exp(-1/Y_F)\}) \tag{4}$$

and

$$(X_E^{(I)}, Y_E^{(I)}) = (1/X_F, 1/Y_F). \tag{5}$$

Here X_E (Y_E) transforms X_F (Y_F) to the exponential margins through a monotone increasing mapping, with repeated use of the probability integral transform, whereas $X_E^{(I)}$ ($Y_E^{(I)}$) transforms X_F (Y_F) to the exponential margins through a monotone decreasing mapping. So marginally both (X_E, Y_E) and $(X_E^{(I)}, Y_E^{(I)})$ are identical, but they differ significantly in their dependence structure and, in particular, their extremal dependence properties. The models (X_F, Y_F) and (X_E, Y_E) have the same copula, while $(X_E^{(I)}, Y_E^{(I)})$ has the same copula as the joint lower tail of (X_F, Y_F). Hence, we refer to (X_E, Y_E) and $(X_E^{(I)}, Y_E^{(I)})$ as having the upper tail and the lower tail copula of (X_F, Y_F), respectively. For both models we explore their joint upper tail, and so focus on different features of the (X_F, Y_F) copula.

The copula of the joint distribution $(X_E^{(I)}, Y_E^{(I)})$ is an example of the class of inverted max-stable models first introduced in Ledford and Tawn [12, 13]. The inverted max-stable distributions are a broad class of AI distributions, covering all values of $\bar{\chi}$ with $0 \le \bar{\chi} < 1$. Heffernan and Tawn [7] found interesting conditional extremal behavior for a sub-family of this class, with much broader structures explored by Papastathopoulos and Tawn [15]. Furthermore, Wadsworth and Tawn [19] explored extensions of the representations of Ledford and Tawn [13] through a series of multivariate regular variation conditions, and found the inverted max-stable distributions to have particular importance in modeling AI. It follows from results in [5] that inverted max-linear models give an arbitrarily good approximation to inverted multivariate extreme value distributions, and so for a study of AI distributions models of the form $(X_E^{(I)}, Y_E^{(I)})$ are of core importance. Next we derive χ and $\bar{\chi}$ for the transformed max-linear and inverted max-linear models.

The joint distribution function of the transformed max-linear model (X_E, Y_E) is

$$\Pr(X_E < x, Y_E < y) = \prod_{i=1}^{m} \min\left[\{1-\exp(-x)\}^{\alpha_i}, \{1-\exp(-y)\}^{\beta_i}\right], \text{ for } x > 0, y > 0. \tag{6}$$

Unlike (X_F, Y_F), this is not max-stable, but this is due to the margin choice not the copula, which remains unchanged. The limiting distribution of normalized component-wise maxima of (6) can be shown to be max-stable, so (X_E, Y_E) is in the domain of attraction of a bivariate extreme value distribution with limiting dependence. For this model it can be shown that $\bar{\chi} = 1$ and $\chi = 2 - \sum_{i=1}^{m} \max(\alpha_i, \beta_i)$, so the variables are AD. On exponential margins, simulations from the max-linear model in (3) give lines of mass, parallel with $X_E = Y_E$, and points scattered around these lines, as shown on Figure 1a, where X_E and Y_E were determined by $X_F = \max(0.7Z_1, 0.2Z_2, 0.1Z_3)$ and $Y_F = \max(0.4Z_1, 0.5Z_2, 0.1Z_4)$. The number of Z_i variables in common between X_F and Y_F determines the number of lines with mass on. In the case of Figure 1a there are two Z_i variables, Z_1 and Z_2, in common between X_F and Y_F, hence there is mass on two lines. The independent scatter of points around the lines is due to the presence of Z_3 in X_F and Z_4 in Y_F.

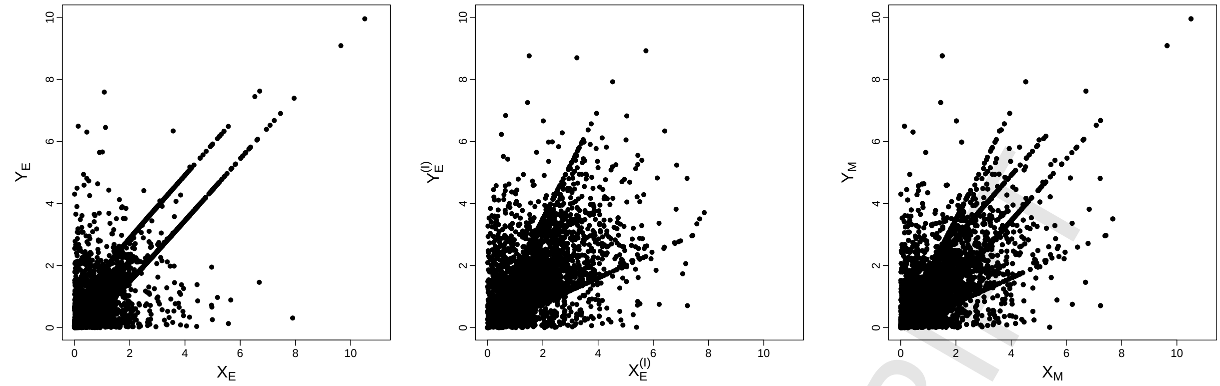
The joint distribution function of the inverted max-linear model $(X_E^{(I)}, Y_E^{(I)})$ is

$$\Pr(X_E^{(I)} < x, Y_E^{(I)} < y) = 1 - \exp(-x) - \exp(-y) + \exp\left\{-\sum_{i=1}^{m} \max(\alpha_i x, \beta_i y)\right\}, \tag{7}$$

for $x > 0$, $y > 0$. For this model, it can be shown that $\chi = 0$ and $\bar{\chi} = \{2/\sum_{i=1}^{m} \max(\alpha_i, \beta_i)\} - 1$, so the variables are AI. Figure 1b shows a random sample from $(X_E^{(I)}, Y_E^{(I)})$ derived from the same max-linear model (X_F, Y_F) as used to illustrate (X_E, Y_E) above. This model gives points on rays and points scattered around these rays. Similarly to (X_E, Y_E), the number of rays is determined by the number of Z_i variables that are common between X_F and Y_F, which in our example is two. Note, that in the inverted max-linear model the point masses are no longer on parallel lines, but on rays ($y = hx$ for $0 < h < \infty$) that meet at the origin. If there exists at least one $i = 1, \ldots, m$ such that $\alpha_i = \beta_i$, then there is a ray with gradient $h = 1$, but despite this the variables are AI.

Combining these two models provides a flexible approach to modeling extremal dependence that can capture both AI and AD. Figure 1c shows an example of a model (X_M, Y_M) that has both AI and AD components. Note, that there is a mass both on parallel lines and on rays in exponential margins, and hence, both AD and AI behaviors are represented. We are interested in the tail behavior of these models, where this feature is most apparent.

Wadsworth and Tawn [19] present a statistical analysis which shows the benefit of this mixture type of model, incorporating AD and AI, over established dependence models. As illustrated in Figure 1, our models put mass on rays and lines, which is inconsistent with most data applications where an assumption of



(a) (b) (c)

Figure 1: Bivariate simulations derived from the max-linear model in (3) with $X_F = \max(0.7Z_1, 0.2Z_2, 0.1Z_3)$ and $Y_F = \max(0.4Z_1, 0.5Z_2, 0.1Z_4)$; (a) transformed max-linear model (X_E, Y_E), (b) inverted max-linear model $(X_E^{(I)},\, Y_E^{(I)})$, and (c) mixture model $(X_M,\, Y_M)$ with $\delta = 0.5$ in (15).

a joint density everywhere is reasonable. Consequently, if these models are fitted using likelihood/Bayesian-based inference they would need almost as many parameters as data points to get a reasonable fit as each line of mass can only explain one data point. Therefore, as currently set up, these are not parsimonious models for likelihood inference but can be used as building blocks for future parsimonious model development. Alternatively, such models can be fitted using other inference criteria which do not depend on the mass on rays/lines. That though is not the focus of this paper. The aim of this paper is to study mathematically the extremal structure of this class of models, the novel tools we use for this are introduced next.

To explore the tail behavior of bivariate distributions with identical marginals the established approach is to adopt a so-called radial-angular representation. We want the radial component, R, to represent how far we are from the origin, and the angular component, W, to represent some form of measure of angle relative to the coordinate axes. This is common practice in multivariate extremes (see, e.g., [4] and [17]). For Fréchet marginals, (X_F, Y_F), these correspond to $R_F = X_F + Y_F$ and $W_F = X_F/(X_F + Y_F)$, although other norms can be used to define these. Then in the limit as $r \to \infty$ the distribution $W_F \mid (R_F > r)$ is non-degenerate if (X_F, Y_F) are AD, but not perfectly dependent, but collapses to mass on $\{0\}$ and $\{1\}$ if the variables are AI. Here the extreme events being considered are those with $R_F > r$.

The key departures to this standard radial-angular approach in our work is that we focus on exponential margins, different combinations of the variables are considered to be extreme, and we use a different dependence variable than W_F. We consider the following radial-angular variables for general bivariate variables $(X,\, Y)$ on exponential margins:

$$R = X + Y, \quad W_D = Y - X, \quad W_I = X/(X + Y).$$

Here two different angular variables W_D and W_I are considered. Also the radial variable R differs from R_F as X and Y are on exponential scale. We will explore these radial-angular variables for the transformed max-linear model (X_E, Y_E) in (4) and the inverted max-linear model $(X_E^{(I)}, Y_E^{(I)})$ in (5).

To help understand the difference in our new radial-angular variables first consider the connection between W_D and W_F. For large X_E we have that $X_E \approx \ln(X_F)$, similarly for Y_E, and so

$$W_F \approx \exp(X_E)/\{\exp(X_E) + \exp(Y_E)\} = 1/\{1 + \exp(Y_E - X_E)\}.$$

Hence, for large X_E and Y_E, W_D is simply a function of W_F, and at first sight it would appear that this choice of radial and angular variable should not reveal any new structure. But conditioning on $R > r$ leads to the selection of different extreme events than $R_F > s$, for any choice of r and s, so different results can arise. Specifically, we will show that the radial and angular representation (R, W_D) gives a non-trivial limit for the distribution of $W_D \mid (R > r)$ as $r \to \infty$ for the transformed max-linear model, but for the inverted max-linear model it gives only mass at $\{-\infty\}$, $\{0\}$ or $\{\infty\}$ depending on the (α_i, β_i), $i = 1, \ldots, m$, values. The latter limit is at odds with results for W_F, as there the associated mass at $\{0\}$, corresponding to $W_F = 1/2$, does not arise. For the radial and angular formulation (R, W_I) the limit distribution of $W_I \mid (R_I > r)$, as $r \to \infty$, for the transformed max-linear model is degenerate, with limit $W_I = 1/2$, and is a non-trivial limit for the inverted max-linear model.

The layout of the article is as follows. In Section 2 we introduce a simple case of the max-linear model given in (3), called the Marshall-Olkin model, and we will use this to derive some of the key tail dependence properties of the model. The mathematical techniques used throughout are based on the techniques shown in this section. Then in Section 3 we derive properties for the general case for both the transformed max-linear and inverted max-linear models. In Section 4 we examine the asymptotic behavior of the upper tail for both of these models. In Section 5 we combine the two models together and study the extremal properties of this formulation. Proofs of the results are given in Section 6. We close with a discussion in Section 7 that discusses multivariate and spatial models extending our bivariate models.

2 Marshall–Olkin model

Let us consider a simple case of model (3). This corresponds to the Marshall and Olkin [14] model, and has the following form:

$$\begin{aligned} X_F &= \max\{\alpha Z_1, (1-\alpha) Z_2\}, \\ Y_F &= \max\{\beta Z_1, (1-\beta) Z_3\}, \end{aligned}$$

where Z_i, $i = 1, 2, 3$, are defined as in (3), and $0 \le \alpha, \beta \le 1$ are known constants. As there is only Z_1 in common between X_F and Y_F, a similar simulation to that shown on Figure 1a would give point mass on a single line, with the rest of the points scattered above and below the line. The variables X_F and Y_F are independent only in the cases when $\alpha = 1$ and $\beta = 0$ or $\alpha = 0$ and $\beta = 1$, otherwise they are dependent.

In order to characterize this model it is useful to define the following three cases: (i) on the line $Y_F = (\beta/\alpha) X_F$, (ii) below the line with $Y_F < (\beta/\alpha) X_F$, and (iii) above the line with $Y_F > (\beta/\alpha) X_F$. In each of these cases there are certain combinations of Z_i's that can lead to them. To have points on the line we need $(X_F, Y_F) = (\alpha Z_1, \beta Z_1)$, which requires $Z_2 \le \alpha Z_1/(1-\alpha)$ and $Z_3 \le \beta Z_1/(1-\beta)$. Below the line we need $(X_F, Y_F) = ((1-\alpha) Z_2, \beta Z_1)$ or $(X_F, Y_F) = ((1-\alpha) Z_2, (1-\beta) Z_3)$ with $(1-\alpha) Z_2/\alpha > (1-\beta) Z_3/\beta$, and above the line $(X_F, Y_F) = (\alpha Z_1, (1-\beta) Z_3)$ or $(X_F, Y_F) = ((1-\alpha) Z_2, (1-\beta) Z_3)$ with $(1-\alpha) Z_2/\alpha < (1-\beta) Z_3/\beta$.

In each case we can derive the probability of being in that case and the density conditional on being in each region. Here we will illustrate the calculations for case (i) when $Y_F = (\beta/\alpha) X_F$; i.e., we want to work out the probability that $X_F < x$ for some $x > 0$ given that $Y_F = (\beta/\alpha) X_F$. We use conditional probability:

$$\Pr\left(X_F < x \,\middle|\, Y_F = \frac{\beta}{\alpha} X_F\right) = \frac{\Pr(X_F < x, Y_F = \frac{\beta}{\alpha} X_F)}{\Pr(Y_F = \frac{\beta}{\alpha} X_F)}. \tag{8}$$

The joint probability in the numerator is

$$\begin{aligned} \Pr\left(X_F < x, Y_F = \frac{\beta}{\alpha} X_F\right) &= \Pr\left\{\alpha Z_1 < x, \alpha Z_1 > (1-\alpha) Z_2, \beta Z_1 > (1-\beta) Z_3\right\} \\ &= \Pr\left(Z_1 < \frac{x}{\alpha}, Z_2 < \frac{\alpha Z_1}{1-\alpha}, Z_3 < \frac{\beta Z_1}{1-\beta}\right). \end{aligned}$$

To calculate this, we can condition on one of the Z's, in this case Z_1, and integrate over the range $Z_1 < x/\alpha$, which gives

$$\begin{aligned}\Pr\left(X_F < x, Y_F = \frac{\beta}{\alpha}X_F\right) &= \int_0^{x/\alpha} \Pr\left(Z_2 < \frac{\alpha z}{1-\alpha}, Z_3 < \frac{\beta z}{1-\beta}\middle| Z_1 = z\right) f_Z(z)dz,\\ &= \int_0^{x/\alpha} e^{-(1-\alpha)/(\alpha z)} e^{-(1-\beta)/(\beta z)} \frac{1}{z^2} e^{-1/z} dz,\\ &= \frac{\alpha\beta}{\alpha+\beta-\alpha\beta} \exp\left(-\frac{\alpha+\beta-\alpha\beta}{\beta x}\right),\end{aligned}$$

where the second equality holds as Z_2 and Z_3 are independent Fréchet random variables. It follows that

$$\Pr\left(Y_F = \frac{\beta}{\alpha}X_F\right) = \frac{\alpha\beta}{\alpha+\beta-\alpha\beta},$$

and hence we have obtained the conditional distribution in (8) as $\exp\{-(\alpha+\beta-\alpha\beta)/(\beta x)\}$ for $x > 0$. To obtain the one-dimensional density of the points on the line we can differentiate this distribution function, which gives

$$f_{X_F}\left(x\middle| Y_F = \frac{\beta}{\alpha}X_F\right) = \frac{\alpha+\beta-\alpha\beta}{\beta x^2} \exp\left(-\frac{\alpha+\beta-\alpha\beta}{\beta x}\right), \text{ for } x > 0,$$

or, equivalently,

$$f_{Y_F}\left(y\middle| Y_F = \frac{\beta}{\alpha}X_F\right) = \frac{\alpha+\beta-\alpha\beta}{\alpha y^2} \exp\left(-\frac{\alpha+\beta-\alpha\beta}{\alpha y}\right), \text{ for } y > 0.$$

See Appendix A for similar calculations for the other two cases using this first principles approach. Deriving the densities as described above is laborious, involving many complex integrals, which makes the calculations hard to extend to the more general case. As the densities seem to have much simpler forms than the distribution functions it seems sensible to work with densities directly. For example, above the line $Y_F > (\beta/\alpha)X_F$ the probability element is

$$\Pr\left(X_F \in dx, Y_F \in dy\middle| Y_F > \frac{\beta}{\alpha}X_F\right) = \frac{\Pr(X_F \in dx, Y_F \in dy, Y_F > \frac{\beta}{\alpha}X_F)}{\Pr(Y_F > \frac{\beta}{\alpha}X_F)}, \tag{9}$$

where $X \in dx$ denotes $X \in (x, x+\delta x)$. Then, there are two possible combinations that lead to this case, $(X_F, Y_F) = (\alpha Z_1, (1-\beta)Z_3)$ and $(X_F, Y_F) = ((1-\alpha)Z_2, (1-\beta)Z_3)$, so the joint probability in the numerator of expression (9) can be broken down into the sum of two probabilities, P_1 and P_2, where

$$\begin{aligned}P_1 &= \Pr\left\{X_F = \alpha Z_1 \in dx, Y_F = (1-\beta)Z_3 \in dy, Y_F > \frac{\beta}{\alpha}X_F\right\},\\ P_2 &= \Pr\left\{X_F = (1-\alpha)Z_2 \in dx, Y_F = (1-\beta)Z_3 \in dy, Y_F > \frac{\beta}{\alpha}X_F\right\}.\end{aligned}$$

Then, it follows that the probability P_1 is equivalent to the joint probability $\Pr\{\alpha Z_1 \in dx, (1-\beta)Z_3 \in dy, Z_2 < x/(1-\alpha)\}$ given that $y > (\beta/\alpha)x$. Hence, using that the Z_i's are independent Fréchet random variables,

$$\begin{aligned}P_1 &= \Pr\left(Z_1 \in \frac{dx}{\alpha}\right)\Pr\left(Z_3 \in \frac{dy}{1-\beta}\right)\Pr\left(Z_2 < \frac{x}{1-\alpha}\right)\mathbf{1}\left(y > \frac{\beta}{\alpha}x\right)\\ &\sim \left(\frac{\alpha}{x^2}e^{-\alpha/x}\right)\left\{\frac{1-\beta}{y^2}e^{-(1-\beta)/y}\right\}\left\{e^{-(1-\alpha)/x}\right\}\mathbf{1}\left(y > \frac{\beta}{\alpha}x\right)\delta x\delta y\\ &= \frac{\alpha(1-\beta)}{x^2y^2}e^{-1/x}e^{-(1-\beta)/y}\ \mathbf{1}\left(y > \frac{\beta}{\alpha}x\right)\delta x\delta y,\end{aligned}$$

as $\delta x \to 0$ and $\delta y \to 0$. Similarly, as $\delta x \to 0$ and $\delta y \to 0$,

$$P_2 \sim \frac{(1-\alpha)(1-\beta)}{x^2y^2}\, e^{-1/x} e^{-(1-\beta)/y} \;\; \mathbf{1}\left(y > \frac{\beta}{\alpha}\, x\right) \delta x \delta y.$$

Hence, by summing P_1 and P_2, as $\delta x \to 0$ and $\delta y \to 0$,

$$\Pr\left(X_F \in dx, Y_F \in dy, Y_F > \frac{\beta}{\alpha} X_F\right) \sim \frac{(1-\beta)}{x^2y^2}\, e^{-1/x} e^{-(1-\beta)/y} \;\; \mathbf{1}\left(y > \frac{\beta}{\alpha}\, x\right) \delta x \delta y.$$

This can be integrated in the region $y > (\beta/\alpha)x$ to obtain the probability

$$\Pr\left(Y_F > \frac{\beta}{\alpha} X_F\right) = \frac{\alpha(1-\beta)}{\alpha + \beta - \alpha\beta}.$$

Hence, we obtain the density, conditionally on being above the line $Y_F > (\beta/\alpha)X_F$, as

$$f_{(X_F,Y_F)}\left(x, y \middle| Y_F > \frac{\beta}{\alpha} X_F\right) = \frac{\alpha + \beta - \alpha\beta}{\alpha x^2 y^2}\, e^{-1/x} e^{-(1-\beta)/y} \;\; \mathbf{1}\left(y > \frac{\beta}{\alpha}\, x\right).$$

Similar calculations can be performed to obtain densities for cases (i) and (ii).

3 General max-linear models

3.1 Set up and densities on Fréchet margins

Our work in this section has considerable parallels with the hitting scenarios and the conditional probability results for max-linear models developed by Wang and Stoev [21]. Here, we go beyond the scope of this paper by calculating conditional densities.

Let us consider the general max-linear model given in expression (3). Without loss of generality, let us assume that the $\alpha_i Z_i$ and $\beta_i Z_i$ terms are ordered such that

$$\begin{aligned} \boldsymbol{\alpha} &= (\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_{k+\ell}, 0, \ldots, 0), \\ \boldsymbol{\beta} &= (\beta_1, \ldots, \beta_k, 0, \ldots, 0, \beta_{k+\ell+1}, \ldots, \beta_m), \end{aligned}$$

i.e., for $i = 1, \ldots, k$, $\alpha_i \neq 0$ and $\beta_i \neq 0$, for $i = k+1, \ldots, k+\ell$, $\alpha_i \neq 0$ and $\beta_i = 0$, and for $i = k+\ell+1, \ldots, m$, $\alpha_i = 0$ and $\beta_i \neq 0$, with $\sum_{i=1}^{k+\ell} \alpha_i = 1$ and $\sum_{i=1}^{k} \beta_i + \sum_{h=k+\ell+1}^{m} \beta_h = 1$. We also assume that $\omega_i := \beta_i/\alpha_i$ are unique for $i = 1, \ldots, k$. In this general case there are k common Z_i variables between X_F and Y_F, hence there is mass on k lines, each with equation $Y_F = (\beta_i/\alpha_i)X_F$, $i = 1, \ldots, k$. If $k = 0$ then X_F and Y_F are independent. Furthermore, without loss of generality, let us assume the following ordering for the first k terms,

$$\frac{\beta_1}{\alpha_1} < \cdots < \frac{\beta_k}{\alpha_k} \;\Leftrightarrow\; 0 < \omega_1 < \cdots < \omega_k < \infty.$$

This notation ensures that the line with mass that has the least steep gradient is $Y_F = \omega_1 X_F$, followed by $Y_F = \omega_2 X_F$, and so on until $Y_F = \omega_k X_F$. Let us also define the following sums,

$$\begin{aligned} \alpha_{\text{sum}} &= \alpha_{k+1} + \cdots + \alpha_{k+\ell}, & \alpha_{\text{sum}}^{(j)} &= \alpha_{\text{sum}} + \Sigma_{i=1}^{j} \alpha_i, \;\text{ for } 0 \le j \le k, \\ \beta_{\text{sum}} &= \beta_{k+\ell+1} + \cdots + \beta_m, & \beta_{\text{sum}}^{(h)} &= \beta_{\text{sum}} + \Sigma_{i=h}^{k} \beta_i, \;\text{ for } 1 \le h \le k+1, \end{aligned}$$

where we define $\sum_{i=1}^{0} x_i = 0$ and $\sum_{i=k+1}^{k} x_i = 0$, which leads to $\alpha_{\text{sum}}^{(0)} = \alpha_{\text{sum}}$ and $\beta_{\text{sum}}^{(k+1)} = \beta_{\text{sum}}$.

In this more general set up it is useful to define four types of 'regions': (i) above the line $Y_F = \omega_k X_F$, (ii) on the line $Y_F = \omega_j X_F$, $j = 1, \ldots, k$, (iii) between the two lines $Y_F = \omega_j X_F$ and $Y_F = \omega_{j+1} X_F$, $j = 1, \ldots, k-1$, and (iv) below the line $Y_F = \omega_1 X_F$. There is one region of type (i) and (iv) each, k regions of type (ii) since there are k lines, and $k-1$ regions of type (iii), since k lines define $k-1$ between-line regions.

The strategy for the derivation of the densities for each of these regions is as in Section 2, with full derivations given in Appendix B. Here we will give the conditional density forms for each of the four region types. The density conditional on being in the region above the line $Y_F = \omega_k X_F$ is

$$f_{(X_F,Y_F)}\left(x, y \,\middle|\, \frac{Y_F}{X_F} > \omega_k\right) = \frac{\alpha_k \beta_{\text{sum}} + \beta_k}{\alpha_k x^2 y^2} \exp\left(-\frac{1}{x} - \frac{\beta_{\text{sum}}}{y}\right),$$

for $y > \omega_k x$. On the line $Y_F = \omega_j X_F$, $j = 1, \ldots, k$, the density for $x > 0$ is

$$f_{(X_F,Y_F)}\left(x, \omega_j x \,\middle|\, \frac{Y_F}{X_F} = \omega_j\right) = \frac{\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)}}{\beta_j x^2} \exp\left(-\frac{\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)}}{\beta_j x}\right).$$

Between two lines $Y_F = \omega_j X_F$ and $Y_F = \omega_{j+1} X_F$, $j = 1, \ldots, k-1$, the conditional density is

$$f_{(X_F,Y_F)}\left(x, y \,\middle|\, \omega_j < \frac{Y_F}{X_F} < \omega_{j+1}\right) = \frac{c_j}{(\alpha_j \beta_{j+1} - \beta_j \alpha_{j+1}) x^2 y^2} \exp\left(-\frac{\alpha_{\text{sum}}^{(j)}}{x} - \frac{\beta_{\text{sum}}^{(j+1)}}{y}\right),$$

for $\omega_j x < y < \omega_{j+1} x$ where $c_j = (\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)})(\alpha_{j+1} \beta_{\text{sum}}^{(j+1)} + \beta_{j+1} \alpha_{\text{sum}}^{(j)})$. Finally, in the region below the line $Y_F = \omega_1 X_F$, the conditional density is

$$f_{(X_F,Y_F)}\left(x, y \,\middle|\, \frac{Y_F}{X_F} < \omega_1\right) = \frac{\alpha_1 + \beta_1 \alpha_{\text{sum}}}{\beta_1 x^2 y^2} \exp\left(-\frac{\alpha_{\text{sum}}}{x} - \frac{1}{y}\right),$$

for $y < \omega_1 x$.

3.2 Densities on exponential margins

3.2.1 Transformed max-linear model

In Section 3.1 we gave densities conditional on being on each line and in the regions defined by the lines on Fréchet margins. Since it is more straightforward to expose the difference between AI and AD on exponential margins, we want to obtain the densities for exponential margins. Hence for each case (i)-(iv) defined in Section 3.1 we will identify the corresponding case on exponential margins and then transform to obtain the densities on the new margins.

On exponential margins for the transformed max-linear model (X_E, Y_E), the line $Y_F = \omega_j X_F$ becomes the curve $Y_E = -\ln\{1 - (1 - e^{-X_E})^{1/\omega_j}\}$ for all $j \in \{1, \ldots, k\}$. For ease of notation, let us define $g_j(X_E) = -\ln\{1 - (1 - e^{-X_E})^{1/\omega_j}\}$ for $j = 1, \ldots, k$. Note that $g_j(X_E) \approx X_E + \ln(\omega_j)$ for large X_E; this asymptotic linearity is useful when we explore the limiting behavior of the model. The cases defined in Section 3.1 become (i) above the curve $Y_E = g_k(X_E)$, (ii) on the curve $Y_E = g_j(X_E)$, $j = 1, \ldots, k$, (iii) between the two curves $Y_E = g_j(X_E)$ and $Y_E = g_{j+1}(X_E)$, $j = 1, \ldots, k-1$, and (iv) below the curve $Y_E = g_1(X_E)$. Note that the transformation to exponential margins means that the lines with mass on are now curves. Furthermore, even asymptotically they are no longer rays that meet at the origin, but parallel lines each with gradient equal to one with intercepts $\ln(\omega_j)$, $j = 1, \ldots, k$. For each region we transform the conditional densities given in Section 3.1 to exponential margins.

Hence, the conditional density in the region above the curve $Y_E = g_k(X_E)$ is:

$$f_{(X_E,Y_E)}\{x,y|Y_E > g_k(X_E)\} = \left(\frac{\alpha_k\beta_{\text{sum}} + \beta_k}{\alpha_k}\right) e^{-x}e^{-y}(1-e^{-y})^{\beta_{\text{sum}}-1}, \tag{10}$$

for $x > 0$ and $y > g_k(x)$. On the curve $Y_E = g_j(X_E)$, $j = 1,\ldots,k$, the conditional density for $x > 0$ is:

$$f_{(X_E,Y_E)}\{x,g_j(x)|Y_E = g_j(X_E)\} = \left(\frac{\alpha_j\beta_{\text{sum}}^{(j+1)} + \beta_j\alpha_{\text{sum}}^{(j)}}{\beta_j}\right) e^{-x}(1-e^{-x})^{(\alpha_j\beta_{\text{sum}}^{(j+1)}+\beta_j\alpha_{\text{sum}}^{(j)})/\beta_j-1}.$$

The conditional density between the curves $Y_E = g_j(X_E)$ and $Y_E = g_{j+1}(X_E)$, for $j = 1,\ldots,k-1$, is:

$$\begin{aligned} f_{(X_E,Y_E)}\{x,y|g_j(X_E) < Y_E < g_{j+1}(X_E)\} = &\left\{\frac{(\alpha_j\beta_{\text{sum}}^{(j+1)} + \beta_j\alpha_{\text{sum}}^{(j)})(\alpha_{j+1}\beta_{\text{sum}}^{(j+1)} + \beta_{j+1}\alpha_{\text{sum}}^{(j)})}{(\alpha_j\beta_{j+1} - \beta_j\alpha_{j+1})}\right\} \\ &\times e^{-x}e^{-y}(1-e^{-x})^{\alpha_{\text{sum}}^{(j)}-1}(1-e^{-y})^{\beta_{\text{sum}}^{(j+1)}-1}, \end{aligned}$$

for $x > 0$ and $g_j(x) < y < g_{j+1}(x)$, $j = 1,\ldots,k-1$. The conditional density below the curve $Y_E = g_1(X_E)$ is

$$f_{(X_E,Y_E)}\{x,y|Y_E < g_1(X_E)\} = \left(\frac{\alpha_1 + \beta_1\alpha_{\text{sum}}}{\beta_1}\right) e^{-x}(1-e^{-x})^{\alpha_{\text{sum}}-1}e^{-y},$$

for $x > 0$ and $y < g_1(x)$.

3.2.2 Inverted max-linear model

Now we turn our attention to the lower tail of the max-linear model (3), i.e., the upper tail of the inverted max-linear model. Similarly to Section 3.2.1, the densities given in Section 3.1 can be transformed to inverted exponential margins $(X_E^{(I)}, Y_E^{(I)})$.

To invert the lower tail of X_E, set $U = 1 - e^{-X_E}$. Then the inversion of U is $U^{(I)} = 1 - U = e^{-X_E}$. Also, $U^{(I)} = 1 - e^{-X_E^{(I)}}$, which leads to $e^{-X_E} = 1 - e^{-X_E^{(I)}}$. Hence, $X_E^{(I)} = -\ln(1-e^{-X_E})$ and by substituting in X_E from expression (4) we get $X_E^{(I)} = 1/X_F$. Similarly, $Y_E^{(I)} = 1/Y_F$.

On the new inverted exponential margins, the line $Y_F = \omega_j X_F$ becomes $Y_E^{(I)} = X_E^{(I)}/\omega_j$ for $j = 1,\ldots,k$. Hence the cases defined in Section 3.1 become (i) below the line $Y_E^{(I)} = X_E^{(I)}/\omega_k$, (ii) on the line $Y_E^{(I)} = X_E^{(I)}/\omega_j$, $j = 1,\ldots,k$, (iii) between the two lines $Y_E^{(I)} = X_E^{(I)}/\omega_{j+1}$ and $Y_E^{(I)} = X_E^{(I)}/\omega_j$, $j = 1,\ldots,k-1$, and (iv) above the line $Y_E^{(I)} = X_E^{(I)}/\omega_j$. Note that the transformation flips the order of the lines, with the line $Y_E^{(I)} = X_E^{(I)}/\omega_1$ having the steepest gradient and $Y_E^{(I)} = X_E^{(I)}/\omega_k$ the least steep.

The conditional density below the line $Y_E^{(I)} = X_E^{(I)}/\omega_k$ has the following form:

$$f_{(X_E^{(I)},Y_E^{(I)})}\left(x,y \mid Y_E^{(I)} < X_E^{(I)}/\omega_k\right) = \frac{\alpha_k\beta_{\text{sum}} + \beta_k}{\alpha_k} e^{-x}e^{-\beta_{\text{sum}}y}, \quad x > 0,\ y < x/\omega_k.$$

On the line $Y_E^{(I)} = X_E^{(I)}/\omega_j$, $j = 1,\ldots,k$, the conditional density for $x > 0$ takes the form:

$$f_{(X_E^{(I)},Y_E^{(I)})}\left(x,x/\omega_j \mid Y_E^{(I)} = X_E^{(I)}/\omega_j\right) = \frac{\alpha_j\beta_{\text{sum}}^{(j+1)} + \beta_j\alpha_{\text{sum}}^{(j)}}{\beta_j} \exp\left(-\frac{\alpha_j\beta_{\text{sum}}^{(j+1)} + \beta_j\alpha_{\text{sum}}^{(j)}}{\beta_j}x\right).$$

Between the two lines $Y_E^{(I)} = X_E^{(I)}/\omega_{j+1}$ and $Y_E^{(I)} = X_E^{(I)}/\omega_j$, $j = 1, \ldots, k-1$, the conditional density is:

$$f_{(X_E^{(I)}, Y_E^{(I)})}\left(x, y \mid X_E^{(I)}/\omega_{j+1} < Y_E^{(I)} < X_E^{(I)}/\omega_j\right) = \frac{c_j}{(\alpha_j \beta_{j+1} - \beta_j \alpha_{j+1})}\, e^{-\alpha_{\text{sum}}^{(j)} x} e^{-\beta_{\text{sum}}^{(j+1)} y}.$$

where $c_j = (\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)})(\alpha_{j+1} \beta_{\text{sum}}^{(j+1)} + \beta_{j+1} \alpha_{\text{sum}}^{(j)})$, $x > 0$ and $x/\omega_j < y < x/\omega_{j+1}$. Finally, above the line $Y_E^{(I)} = X_E^{(I)}/\omega_1$ the conditional density is:

$$f_{(X_E^{(I)}, Y_E^{(I)})}\left(x, y \mid Y_E^{(I)} > X_E^{(I)}/\omega_1\right) = \frac{\alpha_1 + \beta_1 \alpha_{\text{sum}}}{\beta_1} e^{-\alpha_{\text{sum}} x} e^{-y},\ x > 0,\ y > x/\omega_1.$$

4 Angular representation and limiting behavior

In this section we explore the asymptotic behavior of the upper tails of the transformed max-linear model (6) and the inverted max-linear model (7). As discussed in Section 1, we use a radial-angular representation (R, W) to explore the limiting properties of the models. For general exponential marginal variables (X, Y) we define the radial component to be of the form $R = X + Y$. For the angular component we use two different forms: $W_D = Y - X$ and $W_I = X/(X + Y)$ for the reasons given in Section 1. Our aim is to determine the tail behavior of the models in the region $(R > r)$ as $r \to \infty$. So for each type of region $\mathcal{J}$ (identified in the previous sections), and for both forms of W, we will also calculate the joint density of R and W given that $R > r$ to give the conditional probability

$$\Pr(W > w, R > r + t \mid R > r, W \in \mathcal{J}) = \frac{\Pr(W > w, R > r + t \mid W \in \mathcal{J})}{\Pr(R > r \mid W \in \mathcal{J})}, \quad t > 0. \tag{11}$$

Then we can use these results to obtain the conditional probability of being in each region $\mathcal{J}$, given $R > r$ as $r \to \infty$, as

$$\Pr(W \in \mathcal{J} \mid R > r) = \frac{\Pr(R > r \mid W \in \mathcal{J}) \Pr(W \in \mathcal{J})}{\Pr(R > r)},\ \text{for all } \mathcal{J}. \tag{12}$$

4.1 Transformed max-linear model

First, we explore the asymptotic behavior of the upper tail of the transformed max-linear model (6). We use the densities in Section 3.2.1, to obtain the densities in each region on (R, W_D) margins.

For (R, W_D) the curve $Y_E = g_j(X_E)$, $j = 1, \ldots, k$, is $W_D = g_j(X_E) - X_E$, which is approximately $W_D = \ln(\omega_j) := w_j$, with $-\infty < w_j < \infty$, for large R and hence for large X_E. So the case (i) becomes approximately the region $W_D > w_k$ for large R. For finite samples the region is $W_D > g_k(X_E) - X_E$. The joint density conditional on being in this region is obtained from the density in (10) as:

$$f_{(R, W_D)}\{r, w | W_D > g_k(X_E) - X_E\} = \left(\frac{\alpha_k \beta_{\text{sum}} + \beta_k}{2\alpha_k}\right) e^{-r} \{1 - e^{-(r+w)/2}\}^{\beta_{\text{sum}} - 1},\ r > 0,\ w > g_k(X_E) - X_E. \tag{13}$$

We can then calculate the conditional probability in (11) as

$$\Pr\{W_D > w, R > r + t \mid R > r, W_D > g_k(X_E) - X_E\} \approx \left(1 + \frac{t}{r+1}\right) e^{-t} \to e^{-t},\ \text{as } r \to \infty,\ t > 0,\ w > w_k.$$

This shows that $\Pr(W_D > w \mid R > r) \to 1$ for all $w > w_k$. Hence Lemma 1 follows.

Lemma 1. *In the limit as $r \to \infty$, $W_D \mid (R > r, W_D > w_k) \to^p w_k$, and asymptotically $W_D \perp\!\!\!\perp R \mid (R > r, W_D > w_k)$.*

On the curve $W_D = g_j(X_E) - X_E$, $j = 1, \dots, k$, the density is

$$f_R\{r | W_D = g_j(X_E) - X_E\} = c_j e^{-(r-w_j)/2}\{1 - e^{-(r-w_j)/2}\}^{2c_j - 1}, \; r > 0,$$

where $c_j = (\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)})/(2\beta_j)$. Then, the distribution of the points, conditional on being on this curve is

$$\begin{aligned}\Pr\{R > r \mid W_D = g_j(X_E) - X_E\} &= 1 - (1 - e^{w_j/2} e^{-r/2})^{c_j}, \\ &\sim c_j e^{w_j/2} e^{-r/2}, \quad \text{as } r \to \infty,\end{aligned}$$

for $j = 1, \dots, k$.

Lemma 2. *The distribution of the radial points on the line $W_D = w_j$, $j = 1, \dots, k$, has an exponential tail.*

In the region between the curves $W_D = g_j(X_E) - X_E$ and $W_D = g_{j+1}(X_E) - X_E$, $j = 1, \dots, k-1$, the joint density is:

$$f_{(R,W_D)}\{r, w | g_j(X_E) - X_E < W_D < g_{j+1}(X_E) - X_E\} = \frac{c_j}{2} e^{-r}\{1 - e^{-(r-w)/2}\}^{\alpha_{\text{sum}} - 1}\{1 - e^{-(r+w)/2}\}^{\beta_{\text{sum}} - 1},$$

for $r > 0$ and $g_j(X_E) - X_E < w < g_{j+1}(X_E) - X_E$. Then the conditional probability,

$$\Pr\{W_D > w, R > r + t \mid R > r, g_j(X_E) - X_E < W_D < g_{j+1}(X_E) - X_E\} \to \frac{(w_{j+1} - w)e^{-t}}{w_{j+1} - w_j}, \quad \text{as } r \to \infty,$$

for $t > 0$ and $w_j < w < w_{j+1}$. Hence Lemma 3 follows.

Lemma 3. *The limiting angular distribution is uniform in regions between the rays $W_D = w_j$ and $W_D = w_{j+1}$, $j = 1, \dots, k-1$, and independent of the radial variable, which follows a unit exponential distribution.*

Lastly, in the region $W_D < g_1(X_E) - X_E$ the joint density is

$$f_{(R,W_D)}\{r, w | W_D < g_1(X_E) - X_E\} = \left(\frac{\alpha_j + \beta_j \alpha_{\text{sum}}}{2\beta_j}\right) e^{-r}\{1 - e^{-(r-w)/2}\}^{\alpha_{\text{sum}} - 1}, \; r > 0, \; w < g_1(X_E) - X_E.$$

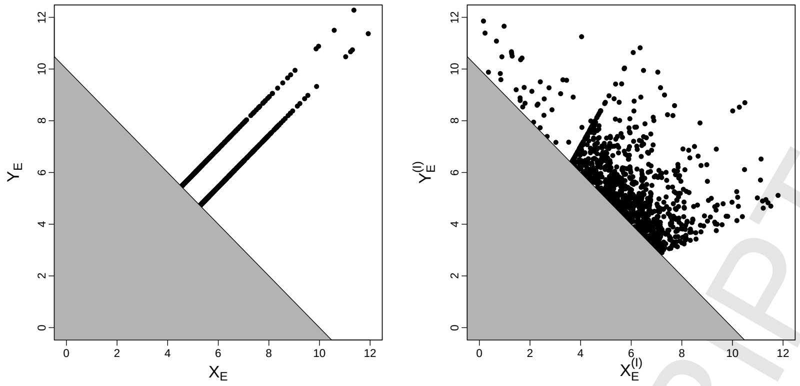
Then the conditional probability,

$$\Pr(W_D > w, R > r + t \mid R > r, W_D < w_1) \sim \frac{(w_1 - w)e^{-t}}{w_1 + r + 1}, \quad \text{as } r \to \infty, \quad t > 0, \quad w < w_1.$$

This suggests that $\Pr(W_D > w \mid R > r, W_D < w_1) \to 0$ as $r \to \infty$ for all $w < w_1$ and the following lemma follows.

Lemma 4. *In the limit as $r \to \infty$, $W_D \mid (R > r, W_D < w_1) \to^p w_1$, and asymptotically $W_D \perp\!\!\!\perp R \mid (R > r, W_D < w_1)$.*

Now we use the results above to calculate the probability of being in each region $\mathcal{J}$, given $R > r$ as $r \to \infty$. Theorems 1 and 2 describe the asymptotic behavior of the conditional probability (12) for angular measures $W = W_D$ and $W = W_I$, respectively. Proofs are deferred to Section 6.



(a) (b)

Figure 2: Asymptotic behavior of the (a) transformed max-linear and (b) inverted max-linear model on exponential margins, for X_F and Y_F defined as in Figure 1.

Theorem 1. *Let $R = X_E + Y_E$, $W_D = Y_E - X_E$ and $w_j = \ln(\beta_j/\alpha_j)$. Then, as $r \to \infty$,*

$$\Pr(W_D \in \mathcal{J} \mid R > r) \to \begin{cases} \dfrac{\alpha_j \exp(w_j/2)}{\sum_{i=1}^{k} \alpha_i \exp(w_i/2)}, & \text{for } \mathcal{J} = \{w_j\}, j = 1, \ldots, k \\ 0, & \text{otherwise.} \end{cases}$$

Theorem 2. *Let $R = X_E + Y_E$ and $W_I = X_E/(X_E + Y_E)$. Then, as $r \to \infty$,*

$$\Pr(W_I < w \mid R > r) \to \begin{cases} 0, & w < 1/2, \\ 1, & w \geq 1/2, \end{cases}$$

i.e., $W_I \mid R > r \to^p 1/2$.

Thus, Theorem 1 shows that in the limit $r \to \infty$, there is only mass on the lines $W_D = w_j$, $j = 1, \ldots, k$, and not in any of the other regions for this model. Figure 2a illustrates this for the max-linear model with the same α and β parameters as in Figure 1. For the other angular form W_I, the mass collapses onto the diagonal, as shown by Theorem 2, and hence W_I is a poor angular measure for exploring the extremal dependence structure for AD variables.

4.2 Inverted max-linear model

Now we explore the asymptotic upper tail behavior of the inverted max-linear model (7). We transform the densities given in Section 3.2.2 to obtain the densities in each region on (R, W_I) margins. On these new margins, the line $Y_E^{(I)} = X_E^{(I)}/\omega_j$ becomes the line $W_I = \omega_j/(1 + \omega_j)$, $j = 1, \ldots, k$. Let us denote $w_j = \omega_j/(1 + \omega_j)$ for $j = 1, \ldots, k$. Note w_j here is different than in Section 4.1. The lines are then ordered such that $0 < w_1 < \cdots < w_k < 1$. Then, the region below the line $Y_E^{(I)} = X_E^{(I)}/\omega_k$ becomes the region $w_k < W_I < 1$. The conditional density in this region is:

$$f_{(R,W_I)}(r, w \mid w_k < W_I < 1) = \left(\frac{\alpha_j \beta_{\text{sum}} + \beta_j}{\alpha_j}\right) r e^{-\beta_{\text{sum}} r} e^{-wr(1-\beta_{\text{sum}})},\ r > 0,\ w_k < w < 1.$$

To determine the limiting behavior for $r \to \infty$ we calculate the conditional probability $\Pr(W_I > w, R > r + t \mid w_k < W_I < 1, R > r)$. We obtain the joint survival function of W_I and R as

$$\Pr(W_I > w, R > r \mid w_k < W_I < 1) = \frac{\alpha_k \beta_{\text{sum}} + \beta_k}{\alpha_k(1 - \beta_{\text{sum}})} \left\{ \frac{e^{-rc(w)}}{c(w)} - e^{-r} \right\}, \quad r > 0, \quad w_k < w < 1, \tag{14}$$

where $c(w) = \beta_{\text{sum}} + w(1 - \beta_{\text{sum}})$. Setting $w = w_k$ in (14) we get the conditional distribution for R given $\omega_k < W_I < 1$. Hence, for $t > 0$ and $w_k < w < 1$,

$$\Pr(W_I > w, R > r + t \mid R > r, w_k < W_I < 1) = \frac{\{c(w)\}^{-1} e^{-(r+t)c(w)} - e^{-(r+t)}}{\{c(w_k)\}^{-1} e^{-rc(w_k)} - e^{-r}}.$$

Now note that $c(w)$ is an increasing function for $w_k < w < 1$ and $c(1) = 1$. Hence, as $r \to \infty$,

$$\Pr(W_I > w, R > r + t \mid R > r, w_k < W_I < 1) \sim \frac{c(w_k)}{c(w)} \, e^{-tc(w)} e^{-r\{c(w) - c(w_k)\}}, \quad w_k < w < 1$$

which as $r \to \infty$ tends to 0 for all $w \in (w_k, 1)$, and equals $e^{-tc(w_k)}$ for $w = w_k$, leading to Lemma 5.

Lemma 5. *$W_I \mid (R > r, w_k < W_I < 1) \to^p w_k$, as $r \to \infty$, and asymptotically W_I is independent of R, which has an exponential tail with rate $c(w_k)$.*

The conditional density of R on the line $W_I = w_j$, $j = 1, \ldots, k$, is

$$f_R(r \mid W_I = w_j) = \left(\frac{\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)}}{\beta_j} \right) \exp\left(-\frac{\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)}}{\beta_j} r \right), \text{ for } r > 0.$$

Lemma 6. *The distribution of $R \mid (W_I = w_j)$, is exponential with rate $(\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)})/\beta_j$, $j = 1, \ldots, k$.*

For the region between the lines $W_I = w_j$ and $W_I = w_{j+1}$ the conditional density is

$$f_{(R,W_I)}(r, w \mid w_j < W_I < w_{j+1}) = \left\{ \frac{(\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)})(\alpha_{j+1} \beta_{\text{sum}}^{(j+1)} + \beta_{j+1} \alpha_{\text{sum}}^{(j)})}{(\alpha_j \beta_{j+1} - \beta_j \alpha_{j+1})} \right\} \times r e^{-\beta_{\text{sum}}^{(j+1)} r} e^{-(\alpha_{\text{sum}}^{(j)} + \beta_{\text{sum}}^{(j)}) w r},$$

for $r > 0$ and $w_j < w < w_{j+1}$. Similarly to above we calculate the conditional probability as

$$\Pr(W_I > w, R > r + t \mid w_j < W_I < w_{j+1}, R > r) = \frac{\{c(w_{j+1})\}^{-1} e^{-(r+t)c(w_{j+1})} - \{c(w)\}^{-1} e^{-(r+t)c(w)}}{\{c(w_{j+1})\}^{-1} e^{-rc(w_{j+1})} - \{c(w_j)\}^{-1} e^{-rc(w_j)}},$$

where $t > 0$, $w_j < w < w_{j+1}$ and $c(w) = w(\alpha_{\text{sum}}^{(j)} + \beta_{\text{sum}}^{(j+1)}) + \beta_{\text{sum}}^{(j+1)}$. Here $c(w)$ is an increasing function for $w_j < w < w_{j+1}$, so $c(w_j) < c(w) < c(w_{j+1})$. Hence, as $r \to \infty$,

$$\Pr(W_I > w, R > r + t \mid w_j < W_I < w_{j+1}, R > r) \sim \frac{c(w_j)}{c(w)} e^{-tc(w)} e^{-r\{c(w) - c(w_j)\}}, \quad w_j < w < w_{j+1},$$

which tends to 0 for $r \to \infty$ for all $w \in (w_j, w_{j+1})$, and tends to $e^{-tc(w_j)}$ for $w = w_j$. Hence, Lemma 7 follows.

Lemma 7. *In the limit as $r \to \infty$, $W_I \mid (R > r, w_j < W_I < w_{j+1}) \to^p w_j$, $j = 1, \ldots, k-1$, and asymptotically $W_I \perp\!\!\!\perp R \mid (R > r, w_j < W_I < w_{j+1})$.*

On (R, W_I) margins the region above the line $Y_E^{(I)} = X_E^{(I)}/\omega_1$ translates to the area represented by $W_I < w_1$. The conditional density in this region is

$$f_{(R,W_I)}(r, w \mid W_I < w_1) = \left(\frac{\alpha_1 + \beta_1\alpha_{\text{sum}}}{\beta_1}\right) e^{-rw(\alpha_{\text{sum}}-1)}e^{-r}, \text{ for } r > 0, w < w_1.$$

Again, we can work out the conditional probability,

$$\Pr(W_I > w, R > r + t \mid W_I < w_1, R > r) = \frac{\{c(w_1)\}^{-1}e^{-(r+t)c(w_1)} - \{c(w)\}^{-1}e^{-(r+t)c(w)}}{\{c(w_1)\}^{-1}e^{-rc(w_1)} - e^{-r}}, \ t > 0, \ w < w_1,$$

where $c(w) = 1 - (1 - \alpha_{\text{sum}})w$. The function $c(w)$ is in this case a decreasing function for $w \in (0, w_1)$, and $c(0) = 1$, so $0 < c(w_1) < c(w) < 1$. This means that as $r \to \infty$,

$$\Pr(W_I > w, R > r + t \mid W_I < w_1, R > r) \to e^{-tc(w)},$$

for all $w \in (0, w_1)$, suggesting that all points in this region will tend to $W_I = 0$ asymptotically as $r \to \infty$.

Lemma 8. *As $r \to \infty$, $W_I \mid (R > r, W_I < w_1) \to^p 0$, and asymptotically W_I is independent of R, which is unit exponential.*

Now we use the above results to calculate the conditional probability (12). Theorems 3 and 4 describe the behavior of the conditional probability of points being in each region given that $R > r$ as $r \to \infty$, i.e., expression (12) with $W = W_I$ and $W = W_D$, respectively, where $\mathcal{J}$ denotes the different regions. Proofs are given in Section 6.

Theorem 3. *Let $R = X_E^{(I)} + Y_E^{(I)}$ and $W_I = X_E^{(I)}/(X_E^{(I)} + Y_E^{(I)})$. Let $\gamma_j = (\alpha_j\beta_{\text{sum}}^{(j+1)} + \beta_j\alpha_{\text{sum}}^{(j)})/(\alpha_j + \beta_j)$, $0 < \gamma_j \leq 1$, for $j = 1, \dots, k$, and $\gamma_{\min} = \min_{j=1,\dots,k}(\gamma_j)$. If there is a unique γ_j value, $j = 1, \dots, k$, equal to $\gamma_{\min}$, i.e., $\gamma_{\min} = \gamma_j$, then, for $t > 0$, as $r \to \infty$,*

a) $\Pr(W_I = w_j, R > r + t \mid R > r) \to (a_j/d_j)e^{-\gamma_j t}$,

b) $\Pr(w_j < W_I < w_{j+1}, R > r + t \mid R > r) \to (b_j/d_j)e^{-\gamma_j t}$,

c) $\Pr(w_{j-1} < W_I < w_j, R > r + t \mid R > r) \to (-b_{j-1}/d_j)e^{-\gamma_j t}$,

d) $\Pr(\{0 < W_I < w_{j-1}\} \cup \{w_j < W_I < 1\}, R > r + t \mid R > r) \to 0$,

where $w_0 = 0$, $w_{k+1} = 1$, $a_j = \alpha_j\beta_j/(\alpha_j + \beta_j)$, $b_j = \alpha_{\text{sum}}^{(j)}\beta_{\text{sum}}^{(j+1)}/(\alpha_{\text{sum}}^{(j)} - \beta_{\text{sum}}^{(j+1)})$ and $d_j = a_j + b_j - b_{j-1}$, $j = 1, \dots, k$.

Theorem 4. *Let $R = X_E^{(I)} + Y_E^{(I)}$ and $W_D = Y_E^{(I)} - X_E^{(I)}$. Let a_j, b_j, d_j, γ_j and $\gamma_{\min}$ be defined as in Theorem 3. If there is a unique γ_j value, $j = 1, \dots, k$, equal to $\gamma_{\min}$, i.e., $\gamma_{\min} = \gamma_j$, then as $r \to \infty$,*

$$W_D \mid R > r \to \begin{cases} +\infty, & \textit{with probability } \mathbf{1}(\alpha_j > \beta_j) + \mathbf{1}(\alpha_j = \beta_j)(-b_{j-1}/d_j), \\ 0, & \textit{with probability } \mathbf{1}(\alpha_j = \beta_j)(a_j/d_j), \\ -\infty, & \textit{with probability } \mathbf{1}(\alpha_j < \beta_j) + \mathbf{1}(\alpha_j = \beta_j)(b_j/d_j). \end{cases}$$

Hence, Theorem 3 shows that asymptotically for $R > r$ and $r \to \infty$, if $\gamma_{\min} = \gamma_j$, there is an exponential density on the jth ray, and a uniform density in the regions between the $(j-1)$th and jth and the jth and $(j+1)$th rays. This is illustrated on Figure 2b for the inverted max-linear model given in Figure 1b.

Note that γ_j is not necessarily unique, so it is possible that $\gamma_{\min} = \gamma_j = \gamma_i$, for i and j distinct integers in $\{1, \ldots, k\}$. If this is the case then mass falls on both the ith and jth rays, and also in the regions on either side of these. For the alternative form W_D for the angular component, the mass collapses to $\{-\infty\}$, $\{0\}$ and $\{\infty\}$, as shown by Theorem 4. This is still the case, even when γ_j is not unique. Note, that even though the inverted max-linear model is AI, we find that there is mass on the diagonal $W_D = 0$ in the case when there exists $i \in \{1, \ldots, k\}$ such that $\alpha_i = \beta_i$. This is due to the fact that we defined the radial and angular components on exponential margins, which gives a different region $R > r$ than the more commonly used Fréchet margins. This illustrates one of the benefits of identifying extremal dependence structure using exponential marginal variables.

5 Mixture distribution

The transformed max-linear model (6) and the inverted max-linear model (7) can be combined into a mixture distribution

$$\begin{pmatrix} X_M \\ Y_M \end{pmatrix} = \begin{cases} \begin{pmatrix} X_E \\ Y_E \end{pmatrix} & \text{with probability } \delta \\ \\ \begin{pmatrix} X_E^{(I)} \\ Y_E^{(I)} \end{pmatrix} & \text{with probability } 1-\delta \end{cases} \tag{15}$$

where $\delta \in [0, 1]$, and (X_E, Y_E) and $(X_E^{(I)}, Y_E^{(I)})$ represent a transformed max-linear model and an inverted max-linear model, respectively, on exponential margins. The statistical importance of the mixture model (15) is most easily seen by studying the sub-asymptotic behavior of χ defined by expression (1). Specifically let

$$\chi(z) = \Pr(Y > z \mid X > z),$$

so $\chi(z) \to \chi$ as $z \to \infty$. For the transformed max-linear model (6) it follows that $\chi = 2 - \sum_{i=1}^m \max(\alpha_i, \beta_i)$ and that for large z

$$\chi_E(z) \approx \chi + \frac{(2-\chi)(1-\chi)}{2} \exp(-z).$$

So here $\chi_E(z)$ converges to $\chi > 0$ at a fixed rate of decay. In contrast, for the inverted max-linear model (7) $\chi = 0$, but

$$\chi_{E^{(I)}}(z) = \{\exp(-z)\}^{\sum_{i=1}^m \max(\alpha_i, \beta_i) - 1}.$$

Here $\chi_{E^{(I)}}(z)$ converges to $\chi = 0$ at a rate of decay depending on the parameters of the underlying max-linear model, but there is no flexibility in the constant multiplier of this rate term. However for the mixture model (15) we have

$$\chi_M(z) \approx \delta\chi + (1-\delta)\{\exp(-z)\}^{1-\chi} + \delta \frac{(2-\chi)(1-\chi)}{2} \exp(-z),$$

where $\chi = 2 - \sum_{i=1}^m \max(\alpha_i, \beta_i)$. Thus, here there is AD, but also a penultimate behavior that has flexibility in both its rate and coefficient features. Hence, although this mixture model is slightly artificial in its construction it has a sufficiently flexible form to be able to capture all natures of the leading and penultimate forms of extremal dependence.

We use results from Section 4 to deduce asymptotic properties of this mixture distribution. Here too, we will use the two different angular form representations W_D and W_I. Let X_M and Y_M be random variables, on exponential margins, from the mixture distribution (15), and let us define the radial and angular variables $R = X_M + Y_M$, $W_D = Y_M - X_M$, and $W_I = X_M/(X_M + Y_M)$. Then, using the angular form W_D, it follows from Theorems 1 and 4 that, asymptotically for $R > r$ as $r \to \infty$, if there are no pairs (α_j, β_j) such that $\alpha_j = \beta_j$, $j = 1, \ldots, k$, then there is mass totalling δ on the lines $W_D = w_h$, $h = 1, \ldots, k$, and $(1-\delta)$ mass either at $\{-\infty\}$ or at $\{+\infty\}$. If there is a pair (α_j, β_j) such that $\alpha_j = \beta_j$, then there is $(1-\delta)(-b_{j-1}/d_j)$ mass at $\{+\infty\}$, $(1-\delta)(-b_j/d_j)$ mass at $\{-\infty\}$, $\delta\{\alpha_h \exp(w_h/2)\}/\{\sum_{i=1}^k \alpha_i \exp(w_i/2)\}$ mass on each line $W_D = w_h$, $h \neq j$, and $\delta\{\alpha_h \exp(w_h/2)\}/\{\sum_{i=1}^k \alpha_i \exp(w_i/2)\} + (1-\delta)(a_j/d_j)$ mass on the diagonal $W_D = w_j = 0$. This is summarised in Theorem 5.

Theorem 5. *Let $R = X_M + Y_M$, $W_D = Y_M - X_M$, and $w_h = \ln(\beta_h/\alpha_h)$, $h = 1, \ldots, k$. Then, for a_j, b_j, d_j and γ_j, $j = 1, \ldots, k$, defined as in Theorem 3, we have the following for $r \to \infty$,*

$$W_D \mid R > r \to \begin{cases} +\infty, & \text{with probability } (1-\delta)\{\mathbf{1}(\alpha_j > \beta_j) + \mathbf{1}(\alpha_j = \beta_j)(-b_{j-1}/d_j)\}, \\ w_h, & \text{with probability } \delta\left\{\frac{\alpha_h \exp(w_h/2)}{\sum_{i=1}^k \alpha_i \exp(w_i/2)}\right\} + (1-\delta)\mathbf{1}(\alpha_j = \beta_j)(a_j/d_j), \text{ for } h = 1, \ldots, k, \\ -\infty, & \text{with probability } (1-\delta)\{\mathbf{1}(\alpha_j < \beta_j) + \mathbf{1}(\alpha_j = \beta_j)(b_j/d_j)\}. \end{cases}$$

Using the second angular form W_I, it follows from Theorems 2 and 3 that asymptotically for $R > r$ as $r \to \infty$, if $\gamma_{\min} = \gamma_j$ and $\alpha_j \neq \beta_j$, then there is mass totalling $(1-\delta)$ on the jth ray and the two regions adjacent to this ray, and δ mass on the diagonal ray $W_I = 1/2$. If $\alpha_j = \beta_j$, then there is $(1-\delta)a_j/d_j + \delta$ mass on the diagonal ray $W_I = 1/2$, and $(1-\delta)b_j/d_j$ and $-(1-\delta)b_{j-1}/d_j$ mass in the two regions on either side of this ray, respectively. See Theorem 6 for details.

Theorem 6. *Let $R = X_M + Y_M$ and $W_I = X_M/(X_M + Y_M)$, and a_j, b_j, d_j and γ_j, $j = 1, \ldots, k$, defined as in Theorem 3. Then, if $\gamma_{\min} = \gamma_j$, we have the following as $r \to \infty$,*

a) $\Pr(W_I = w_j \mid R > r) \to (1-\delta)a_j/d_j + \delta\mathbf{1}(\alpha_j = \beta_j)$,

b) $\Pr(W_I = 1/2 \mid R > r) \to \delta + (1-\delta)(a_j/d_j)\mathbf{1}(\alpha_j = \beta_j)$,

c) $\Pr(w_j < W_I < w_{j+1} \mid R > r) \to (1-\delta)b_j/d_j$,

d) $\Pr(w_{j-1} < W_I < w_j \mid R > r) \to -(1-\delta)b_{j-1}/d_j$,

e) $\Pr(\{0 < W_I < w_{j-1}\} \cup \{w_j < W_I < 1\} \mid R > r) \to 0$.

6 Proofs

6.1 Proof of Theorem 1

The probability $\Pr(R > r)$ can be written in the following way using total probability:

$$\begin{aligned} \Pr(R > r) &= \sum_{\mathcal{J}} \Pr(R > r \mid W_D \in \mathcal{J}) \Pr(W_D \in \mathcal{J}) \\ &= \Pr(R > r \mid W_D > w_k) \Pr(W_D > w_k) + \sum_{j=1}^{k} \Pr(R > r \mid W_D = w_j) \Pr(W_D = w_j) \\ &+ \sum_{j=1}^{k-1} \Pr(R > r \mid w_j < W_D < w_{j+1}) \Pr(w_j < W_D < w_{j+1}) + \Pr(R > r \mid W_D < w_1) \Pr(W_D < w_1), \end{aligned}$$

where each of the product terms in this sum can be derived using results from Section 4.1. We will illustrate the derivation of the elements of the first term $\Pr(R > r \mid W_D > w_k)\Pr(W_D > w_k)$. First, note that $\Pr(W_D > w_k) \approx \Pr\{Y_F > (\beta_k/\alpha_k)X_F\}$ for large X_F and Y_F. Hence,

$$\Pr(W_D > w_k) \approx \frac{\alpha_k\beta_{\text{sum}}}{\alpha_k\beta_{\text{sum}} + \beta_k},$$

using results from Appendix B.1. To obtain $\Pr(R > r \mid W_D > w_k)$ we first integrate expression (13) with respect to w:

$$\begin{aligned} f_R(r \mid W_D > w_k) &= \int_{w_k}^{r} f_{(R,W_D)}(r, w \mid W_D > w_k)dw, \\ &\approx \frac{\alpha_k\beta_{\text{sum}} + \beta_k}{2\alpha_k} e^{-r}(r - w_k). \end{aligned}$$

Then integrate this with respect to r to obtain the conditional probability:

$$\begin{aligned} \Pr(R > r \mid W_D > w_k) &= \int_{r}^{\infty} f_R(r \mid W_D > w_k)dr, \\ &\approx \frac{\alpha_k\beta_{\text{sum}} + \beta_k}{2\alpha_k} e^{-r}(r - w_k + 1). \end{aligned}$$

The other three product terms can be derived similarly, leading to

$$\begin{aligned} \Pr(R > r) \approx{} & \frac{\beta_{\text{sum}}}{2}(r - w_k + 1)\exp(-r) + \sum_{j=1}^{k} \frac{\alpha_j}{2} \exp\left(\frac{w_j}{2}\right) \exp\left(-\frac{r}{2}\right) \\ & + \sum_{j=1}^{k-1} \frac{\alpha_{\text{sum}}^{(j)}\beta_{\text{sum}}^{(j+1)}}{2}(w_{j+1} - w_j)\exp(-r) + \frac{\alpha_{\text{sum}}}{2}(r + w_1 + 1)\exp(-r), \text{ for large } r. \end{aligned}$$

For $r \to \infty$ we can write

$$\Pr(R > r) \sim \exp\left(-\frac{r}{2}\right) \sum_{j=1}^{k} \frac{\alpha_j}{2} \exp\left(\frac{w_j}{2}\right),$$

since the other terms all contain $\exp(-r)$ and they go to zero faster as $r \to \infty$.

For the region $W_D > w_k$, substituting into (12), we get for large r

$$\Pr(W_D > w_k \mid R > r) \approx \frac{\beta_{\text{sum}}(r - w_k + 1)\exp(-r)}{\sum_{j=1}^{k} \alpha_j \exp(w_j/2)\exp(-r/2)}.$$

Since the numerator tends to zero faster than the denominator, $\Pr(W_D > w_k \mid R > r) \to 0$ as $r \to \infty$. This is the same for the region $W_D < w_1$ and the regions $w_j < W_D < w_{j+1}$ for all $j = 1, \ldots, k-1$, i.e., $\Pr(W_D < w_1 \mid R > r) \to 0$ and $\Pr(w_j < W_D < w_{j+1} \mid R > r) \to 0$ for all $j = 1, \ldots, k-1$ as $r \to \infty$.

For the case when $W_D = w_j$, $j = 1, \ldots, k$, both the numerator and denominator have exponent term $\exp(-r/2)$, hence,

$$\Pr(W_D = w_j \mid R > r) \to \frac{\alpha_j \exp(w_j/2)}{\sum_{i=1}^{k} \alpha_i \exp(w_i/2)}, \text{ as } r \to \infty.$$

$\square$

6.2 Proof of Theorem 2

It follows from Theorem 1 that, asymptotically, W_D is on one of the k lines, i.e., on exponential margins, $Y_E = -\ln[-1 - \{1 - \exp(-X_E)\}^{1/\omega_j}] \approx X_E + \ln(\omega_j)$, $j = 1, \ldots, k$, for large X_E. Hence, for large R,

$$W_I \approx \frac{1}{2} - \frac{\ln(\omega_j)}{2R}, \quad j = 1, \ldots, k.$$

Hence, using Theorem 1, it follows that $W_I \mid R > r \to^p 1/2$, $j = 1, \ldots, k$, as $r \to \infty$. □

6.3 Proof of Theorem 3

The probability $\Pr(W_I \in \mathcal{J} \mid R > r)$ is equivalent to the expression given in (12) with $W = W_I$. The joint probability in the numerator can be calculated in each case using methods similar to those in previous sections; see, e.g., equation (14) for the case when $W_I > w_k$. Then we can use total probability to calculate $\Pr(R > r)$ as shown below:

$$\begin{aligned}
\Pr(R > r) &= \sum_{\mathcal{J}} \Pr(R > r \mid W_I \in \mathcal{J}) \Pr(W_I \in \mathcal{J}) \\
&= \Pr(R > r \mid W_I > w_k) \Pr(W_I > w_k) + \sum_{j=1}^{k} \Pr(R > r \mid W_I = w_j) \Pr(W_I = w_j) \\
&+ \sum_{j=1}^{k-1} \Pr(R > r \mid w_j < W_I < w_{j+1}) \Pr(w_j < W_I < w_{j+1}) + \Pr(R > r \mid W_I < w_1) \Pr(W_I < w_1) \\
&= \frac{\beta_{\text{sum}}}{1 - \beta_{\text{sum}}} \left\{\gamma_k^{-1} \exp(-\gamma_k r) - \exp(-r)\right\} + \sum_{j=1}^{k} \frac{\alpha_j \beta_j}{\alpha_j + \beta_j} \gamma_j^{-1} \exp(-\gamma_j r) \\
&+ \sum_{j=1}^{k-1} \frac{\alpha_{\text{sum}}^{(j)} \beta_{\text{sum}}^{(j+1)}}{\alpha_{\text{sum}}^{(j)} - \beta_{\text{sum}}^{(j+1)}} \left\{\gamma_j^{-1} \exp(-\gamma_j r) - \gamma_{j+1}^{-1} \exp(-\gamma_{j+1} r)\right\} \\
&+ \frac{\alpha_{\text{sum}}}{1 - \alpha_{\text{sum}}} \left\{\gamma_1^{-1} \exp(-\gamma_1 r) - \exp(-r)\right\},
\end{aligned}$$

where $\gamma_j = (\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)})/(\alpha_j + \beta_j)$, $0 < \gamma_j \le 1$, for $j = 1, \ldots, k$. For large r, $\Pr(R > r)$ approximately becomes

$$\begin{aligned}
\Pr(R > r) \approx& \frac{\beta_{\text{sum}}}{1 - \beta_{\text{sum}}} \gamma_k^{-1} \exp(-\gamma_k r) + \sum_{j=1}^{k} \frac{\alpha_j \beta_j}{\alpha_j + \beta_j} \gamma_j^{-1} \exp(-\gamma_j r) \\
&+ \sum_{j=1}^{k-1} \frac{\alpha_{\text{sum}}^{(j)} \beta_{\text{sum}}^{(j+1)}}{\alpha_{\text{sum}}^{(j)} - \beta_{\text{sum}}^{(j+1)}} \left\{\gamma_j^{-1} \exp(-\gamma_j r) - \gamma_{j+1}^{-1} \exp(-\gamma_{j+1} r)\right\} \\
&+ \frac{\alpha_{\text{sum}}}{1 - \alpha_{\text{sum}}} \gamma_1^{-1} \exp(-\gamma_1 r), \qquad (16)
\end{aligned}$$

since the terms containing $\exp(-r)$ are smaller than the terms containing the exponential terms $\exp(-\gamma_j r)$ for all $j = 1, \ldots, k-1$. Note that using the notation defined in Theorem 3, we can write (16) as

$$\Pr(R > r) \approx \sum_{j=1}^{k} (a_j + b_j - b_{j-1}) \gamma_j^{-1} e^{-\gamma_j r}. \qquad (17)$$

For each region the probability in (12) will tend to zero as $r \to \infty$, unless the exponent term in the numerator is the same as the largest exponent term in expression (16). Hence, the result in Theorem 3 follows. $\square$

6.4 Proof of Theorem 4

For the inverted max-linear distribution on exponential margins, $Y_E^{(I)} = X_E^{(I)}/\omega_j$, $j = 1, \ldots, k$, where $0 < \omega_j < \infty$. On (R, W_D) margins this becomes, $W_D = R\delta_j$, where $\delta_j = (1-\omega_j)/(1+\omega_j)$, $j = 1, \ldots, k$. Hence, as $r \to \infty$,

$$W_D \mid R > r \to \begin{cases} \infty, & \text{for } \omega_j < 1, \\ 0, & \text{for } \omega_j = 1, \\ -\infty, & \text{for } \omega_j > 1. \end{cases}$$

To work out the mass at $\{-\infty\}$, $\{0\}$ and $\{\infty\}$ we calculate the probabilities $\Pr(W_D < 0 \mid R > r)$, $\Pr(W_D = 0 \mid R > r)$ and $\Pr(W_D > 0 \mid R > r)$. By conditional probability,

$$\Pr(W_D < 0 \mid R > r) = \frac{\Pr(R > r \mid W_D < 0)\Pr(W_D < 0)}{\Pr(R > r)} \tag{18}$$

where $\Pr(R > r)$ is the same as in expression (17), and $\Pr(W_D < 0)$ can be easily obtained as $\Pr(W_D < 0) = \Pr(X_F < Y_F) = \sum_{i=1}^{k} \mathbf{1}(\alpha_i = \beta_i)\beta_{\text{sum}}^{(j+1)}/(\alpha_{\text{sum}}^{(j)} + \beta_{\text{sum}}^{(j+1)})$. Then we calculate the conditional probability $\Pr(R > r \mid W_D < 0)$ as the following sum,

$$\begin{aligned}\Pr(R > r \mid W_D < 0) &= \Pr(R > r \mid W_D < 0, W_D < R\delta_k)\Pr(W_D < R\delta_k \mid W_D < 0) \\ &+ \sum_{j=i+1}^{k} \Pr(R > r \mid W_D < 0, W_D = R\delta_j)\Pr(W_D = R\delta_j \mid W_D < 0) \\ &+ \sum_{j=i}^{k-1} \Pr(R > r \mid W_D < 0, R\delta_{j+1} < W_D < R\delta_j)\Pr(R\delta_{j+1} < W_D < R\delta_j \mid W_D < 0),\end{aligned}$$

where we assumed that there is an $i \in \{1, \ldots, k\}$ such that $\omega_i = 1$. Using results from Section 3.2.2, we have

$$\begin{aligned}\Pr(R > r, W_D < 0) &\approx b_k\gamma_k^{-1}e^{-\gamma_k r} + \sum_{j=i+1}^{k} a_j\gamma_j^{-1}e^{-\gamma_j r} + \sum_{j=i}^{k-1} b_j\left\{\gamma_j^{-1}e^{-\gamma_j r} - \gamma_{j+1}^{-1}e^{-\gamma_{j+1} r}\right\}, \\ &= b_i\gamma_i^{-1}e^{-\gamma_i r} + \sum_{j=i+1}^{k}(a_j + b_j - b_{j-1})\gamma_j^{-1}e^{-\gamma_j r},\end{aligned}$$

where a_j, b_j and γ_j, $j = 1, \ldots, k$, are defined as in Theorem 3. The probabilities $\Pr(W_D = 0 \mid R > r)$ and $\Pr(W_D > 0 \mid R > r)$ can be calculated similarly. Substituting into expression (18) we obtain the results in Theorem 4. $\square$

7 Conclusions

In this paper we have characterized the asymptotic behavior of models built on bivariate max-linearity, using two different angular measures defined in exponential marginal space. We found that the limiting behavior of our three models (transformed max-linear, inverted max-linear and mixture) can be either asymptotically dependent or asymptotically independent. At finite levels, however, they feature points on rays of the form $y = hx$, $0 < h < \infty$, points on lines of the form $y = h + x$, $-\infty < h < \infty$, and independent points scattered in the regions defined by these rays and lines.

Simulation from the max-linear model (3) is straightforward by sampling Z_j, $j = 1, \ldots, m$, independently from a Fréchet distribution and simply calculating X_F and Y_F, subject to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ values being known. Then we can transform to margins (X_E, Y_E) or $(X_E^{(I)}, Y_E^{(I)})$ to obtain samples on exponential margins from the transformed max-linear or the inverted max-linear models. Assuming δ is also known, we can easily sample from the mixture distribution (15), by sampling from the transformed max-linear model with probability δ and the inverted max-linear model with probability $1 - \delta$. Simulation from the conditional distribution $Y_F \mid X_F$ is also straightforward if $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are known, using methods described in [21], so conditional simulation follows easily for our three models. For a detailed description of the simulation algorithm the reader is referred to [10].

This paper has been restricted to bivariate models, but the formulation is straightforward to extend to multivariate cases. Specifically, consider a d-dimensional max-linear model, with Fréchet margins, where

$$X_{F,j} = \max_{i=1,\ldots,m} (\alpha_{ij} Z_i) \text{ for } j = 1, \ldots, d,$$

where $Z_1, \ldots, Z_m$ are independent and identically distributed Fréchet variables and $\alpha_{ij} \geq 0$ with $\sum_{i=1}^m \alpha_{ij} = 1$ for all $j = 1, \ldots, d$. The multivariate transformed max-linear and inverted max-linear models follow using multivariate analogues of transformations (4) and (5), respectively. The extreme values from these joint distributions can be studied by multivariate extensions of our radial and angular transformations, in particular using

$$R = \sum_{j=1}^{d} X_j, \quad \mathbf{W}_D = (X_2 - X_1, \ldots, X_d - X_1), \quad \mathbf{W}_I = (X_2/R, \ldots, X_d/R),$$

where $(X_1, \ldots, X_d)$ are on exponential margins. We expect to obtain similar findings to the bivariate case with a range of asymptotic independence and asymptotic dependence over different subsets of the variables. Similarly, both the joint and conditional simulation algorithms can be easily extended to the multivariate case.

Our work has potential to be useful in spatial applications of extreme value theory since there is often need to model bivariate dependence for both local dependence and long-range dependence in this setting. Generally, extreme events at locations close by are expected to occur simultaneously, as they are likely to be affected by the same underlying process. Hence, it seems natural to model these as asymptotically dependent. On the other hand, extreme events at locations far apart are unlikely to occur together as the chance of both locations being affected by the same event is reduced; thus asymptotically independent models seem more appropriate in this case. In practice, it is necessary to have a model that can move through the two types and different levels of extremal dependence; e.g., to model sites close by as asymptotically dependent, with dependence decreasing with distance, and asymptotic independence for locations further apart. To achieve a smooth transition between the two types of dependence we need a model that has both components. This paper introduced a new model that incorporates both types of dependence, and that easily lends itself to spatial applications, since we can allow the mass on the rays and lines to vary smoothly with distance or some measure of the strength of dependence between locations. One way to do this could be to have both α_i and β_i $(i = 1, \ldots, m)$ decay with distance at different rates. The development of such models and their statistical inference is the topic for further work.

Acknowledgements

Mónika Kereszturi gratefully acknowledges the financial support of EPSRC through the Centre for Doctoral Training in Statistics and Operational Research in partnership with industry (EPSRC EP/H023151/1) and of Shell Research Limited. We also thank the referees for thoughtful comments that greatly improved the paper.

Appendices

A Derivation of conditional densities for the Marshall–Olkin model

A.1 Case (ii) - Below the line with $Y_F < \frac{\beta}{\alpha} X_F$

In this case, we want to work out the distribution of the points in the region below the line. A similar approach to the one described in Section 2 can be taken, but here we need to consider two possible combinations of Z_i's that can give $Y_F < \frac{\beta}{\alpha} X_F$. So, by conditional probability,

$$\Pr\left(X_F > x, Y_F > y \middle| Y_F < \frac{\beta}{\alpha} X_F\right) = \frac{\Pr(X_F > x, Y_F > y, Y_F < \frac{\beta}{\alpha} X_F)}{\Pr(Y_F < \frac{\beta}{\alpha} X_F)}.$$

As noted before, there are two possible combinations that lead to this case: $(X_F, Y_F) = ((1-\alpha)Z_2, \beta Z_1)$, and $(X_F, Y_F) = ((1-\alpha)Z_2, (1-\beta)Z_3)$ with $(1-\alpha)Z_2/\alpha > (1-\beta)Z_3/\beta$. So, the joint probability can be broken down into the sum of the two cases, such that

$$\Pr\left(X_F > x, Y_F > y, Y_F < \frac{\beta}{\alpha} X_F\right) = P_1 + P_2,$$

where

$$\begin{aligned} P_1 &= \Pr\left\{X_F > x, Y_F > y, X_F = (1-\alpha)Z_2, Y_F = \beta Z_1\right\}, \\ P_2 &= \Pr\left\{X_F > x, Y_F > y, X_F = (1-\alpha)Z_2, Y_F = (1-\beta)Z_3\right\}. \end{aligned}$$

Then,

$$\begin{aligned} P_1 &= \Pr\{(1-\alpha)Z_2 > x, \beta Z_1 > y, (1-\alpha)Z_2 > \alpha Z_1, \beta Z_1 > (1-\beta)Z_3\} \\ &= \Pr\left(Z_2 > \frac{x}{1-\alpha}, Z_1 > \frac{y}{\beta}, Z_2 > \frac{\alpha Z_1}{1-\alpha}, Z_3 < \frac{\beta Z_1}{1-\beta}\right) \\ &= \int_{y/\beta}^{\infty} \Pr\left\{Z_2 > \max\left(\frac{x}{1-\alpha}, \frac{\alpha z}{1-\alpha}\right), Z_3 < \frac{\beta Z_1}{1-\beta} \middle| Z_1 = z\right\} f_Z(z) dz \\ &= \mathbf{1}\left(y < \frac{\beta}{\alpha} x\right) \left\{\beta - \frac{\alpha\beta}{\alpha+\beta-\alpha\beta} - \beta e^{-1/y} - \left(\beta - \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}\right) e^{-(\alpha+\beta-\alpha\beta)/(\beta x)} + \beta e^{-(1-\alpha)/x} e^{-1/y}\right\}, \end{aligned}$$

and

$$\begin{aligned} P_2 &= \Pr\{(1-\alpha)Z_2 > x, (1-\beta)Z_3 > y, (1-\alpha)Z_2 > \alpha Z_1, (1-\beta)Z_3 > \beta Z_1, \beta(1-\alpha)Z_2 > \alpha(1-\beta)Z_3\} \\ &= \mathbf{1}\left(y < \frac{\beta}{\alpha} x\right) \left\{\frac{\beta(1-\alpha)(1-\beta)}{\alpha+\beta-\alpha\beta} - (1-\beta)e^{-1/y} - \frac{\beta(1-\alpha)(1-\beta)}{\alpha+\beta-\alpha\beta} e^{-(\alpha+\beta-\alpha\beta)/(\beta x)} \right. \\ &\quad \left. + (1-\beta)e^{-(1-\alpha)/x} e^{-1/y}\right\}, \end{aligned}$$

where the last equalities in the derivations of P_1 and P_2 follow after extensive calculations. Summing P_1 and P_2, we get

$$\Pr\left(X_F > x, Y_F > y, Y_F < \frac{\beta}{\alpha} X_F\right) = \mathbf{1}\left(y < \frac{\beta}{\alpha} x\right)\left\{\frac{(1-\alpha)\beta}{\alpha+\beta-\alpha\beta} - e^{-1/y} - \frac{(1-\alpha)\beta}{\alpha+\beta-\alpha\beta} e^{-(\alpha+\beta-\alpha\beta)/(\beta x)} + e^{-(1-\alpha)/x} e^{-1/y}\right\}.$$

Similarly for the denominator,

$$\Pr\left(Y_F < \frac{\beta}{\alpha} X_F\right) = \tilde{P}_1 + \tilde{P}_2,$$

where

$$\tilde{P}_1 = \Pr\left\{X_F = (1-\alpha)Z_2, Y_F = \beta Z_1\right\} = \beta - \frac{\alpha\beta}{\alpha+\beta-\alpha\beta},$$
$$\tilde{P}_2 = \Pr\left\{X_F = (1-\alpha)Z_2, Y_F = (1-\beta)Z_3\right\} = \frac{\beta(1-\alpha)(1-\beta)}{\alpha+\beta-\alpha\beta}.$$

Hence,

$$\Pr\left(Y_F < \frac{\beta}{\alpha} X_F\right) = \frac{(1-\alpha)\beta}{\alpha+\beta-\alpha\beta}.$$

Substituting into the conditional probability formula we finally obtain the joint survival function conditional on being in the region below the line, as

$$\Pr\left(X_F > x, Y_F > y \middle| Y_F < \frac{\beta}{\alpha} X_F\right) = \mathbf{1}\left(y < \frac{\beta}{\alpha} x\right)\left\{1 - \frac{\alpha+\beta-\alpha\beta}{(1-\alpha)\beta} e^{-1/y} - e^{-(\alpha+\beta-\alpha\beta)/(\beta x)} + \frac{\alpha+\beta-\alpha\beta}{(1-\alpha)\beta} e^{-(1-\alpha)/x} e^{-1/y}\right\}.$$

Differentiating, we obtain the conditional density in the region below the line as:

$$f_{(X_F,Y_F)}\left(x, y \middle| Y_F < \frac{\beta}{\alpha} X_F\right) = \frac{\alpha+\beta-\alpha\beta}{\beta x^2 y^2} e^{-(1-\alpha)/x} e^{-1/y} \ \mathbf{1}\left(y < \frac{\beta}{\alpha} x\right).$$

A.2 Case (iii) - Above the line with $Y_F > \frac{\beta}{\alpha} X_F$

In this case, we want to work out the distribution of the points conditional on being in the region above the line. The calculations in this case are very similar to those in Section A.1 so we will give less detail. Again, by conditional probability,

$$\Pr\left(X_F > x, Y_F > y \middle| Y_F > \frac{\beta}{\alpha} X_F\right) = \frac{\Pr(X_F > x, Y_F > y, Y_F > \frac{\beta}{\alpha} X_F)}{\Pr(Y_F > \frac{\beta}{\alpha} X_F)},$$

and there are two possible combinations that lead to this case: $(X_F, Y_F) = (\alpha Z_1, (1-\beta)Z_3)$ and $(X_F, Y_F) = ((1-\alpha)Z_2, (1-\beta)Z_3)$ with $(1-\alpha)Z_2/\alpha < (1-\beta)Z_3/\beta$.

Similarly to Section A.1, we can calculate the following marginal and joint probabilities:

$$\Pr\left(Y_F > \frac{\beta}{\alpha} X_F\right) = \frac{\alpha(1-\beta)}{\alpha+\beta-\alpha\beta},$$

and

$$\Pr\left(X_F > x, Y_F > y, Y_F > \frac{\beta}{\alpha} X_F\right) = \mathbf{1}\left(y > \frac{\beta}{\alpha} x\right) \left\{ \frac{\alpha(1-\beta)}{\alpha+\beta-\alpha\beta} - e^{-1/x} - \frac{\alpha(1-\beta)}{\alpha+\beta-\alpha\beta} e^{-(\alpha+\beta-\alpha\beta)/(\alpha y)} + e^{-1/x} e^{-(1-\beta)/y} \right\}.$$

Substituting into the conditional probability formula we obtain the conditional distribution:

$$\Pr\left(X_F > x, Y_F > y \middle| Y_F > \frac{\beta}{\alpha} X_F\right) = \mathbf{1}\left(y > \frac{\beta}{\alpha} x\right) \left\{ 1 - \frac{\alpha+\beta-\alpha\beta}{\alpha(1-\beta)} e^{-1/x} - e^{-(\alpha+\beta-\alpha\beta)/(\alpha y)} + \frac{\alpha+\beta-\alpha\beta}{\alpha(1-\beta)} e^{-1/x} e^{-(1-\beta)/y} \right\}.$$

Differentiating, we obtain the conditional density of the points in the region above the line as

$$f_{(X_F,Y_F)}\left(x, y \middle| Y_F > \frac{\beta}{\alpha} X_F\right) = \frac{\alpha+\beta-\alpha\beta}{\alpha x^2 y^2} e^{-1/x} e^{-(1-\beta)/y} \; \mathbf{1}\left(y > \frac{\beta}{\alpha} x\right).$$

B Derivation of density formulas in the general case

B.1 Type (i) - Above the line $Y_F = \omega_k X_F$

From the condition that $Y_F > \omega_k X_F$, it can be established that the pairs that can lead to this case are combinations of the following: $X_F = \alpha_i Z_i$ where $i = 1, \ldots, k+\ell$, and $Y_F = \beta_h Z_h$ where $h = k+\ell+1, \ldots, m$. Hence,

$$\Pr\left(X_F \in dx, Y_F \in dy, Y_F > \omega_k X_F\right) = \sum_{h=k+\ell+1}^{m} \sum_{i=1}^{k+\ell} \Pr\left(\alpha_i Z_i \in dx, \beta_h Z_h \in dy, Y_F > \omega_k X_F\right).$$

The Z_i's are independent Fréchet random variables, hence

$$\begin{aligned}
&\Pr\left(X_F \in dx, Y_F \in dy, Y_F > \omega_k X_F\right) \\
&= \mathbf{1}(y > \omega_k x) \sum_{h=k+\ell+1}^{m} \sum_{i=1}^{k+\ell} \left[\Pr\left(Z_i \in \frac{dx}{\alpha_i}\right) \Pr\left(Z_h \in \frac{dy}{\beta_h}\right) \prod_{p=1,\{p\neq i\}}^{k} \Pr\left\{ Z_p < \min\left(\frac{x}{\alpha_p}, \frac{y}{\beta_p}\right) \right\} \right. \\
&\left. \times \prod_{p=k+1,\{p\neq i\}}^{k+\ell} \Pr\left(Z_p < \frac{x}{\alpha_p}\right) \prod_{p=k+\ell+1,\{p\neq h\}}^{m} \Pr\left(Z_p < \frac{y}{\beta_p}\right) \right] dxdy \\
&= \mathbf{1}(y > \omega_k x) \sum_{h=k+\ell+1}^{m} \sum_{i=1}^{k+\ell} \left\{ \frac{\alpha_i}{x^2} e^{-\alpha_i/x} \frac{\beta_h}{y^2} e^{-\beta_h/y} \prod_{p=1,\{p\neq i\}}^{k+\ell} \left(e^{-\alpha_p/x}\right) \prod_{p=k+\ell+1,\{p\neq h\}}^{m} \left(e^{-\beta_p/y}\right) \right\} dxdy \\
&= \mathbf{1}(y > \omega_k x) \sum_{h=k+\ell+1}^{m} \sum_{i=1}^{k+\ell} \left\{ \frac{\alpha_i \beta_h}{x^2 y^2} \prod_{p=1}^{k+\ell} \left(e^{-\alpha_p/x}\right) \prod_{p=k+\ell+1}^{m} \left(e^{-\beta_p/y}\right) \right\} dxdy \\
&= \mathbf{1}(y > \omega_k x) \frac{\beta_{\text{sum}}}{x^2 y^2} e^{-1/x} e^{-\beta_{\text{sum}}/y} dxdy,
\end{aligned} \tag{19}$$

where the third equality follows since $x/\alpha_p < y/\beta_p$, $\forall p = 1, \ldots, k$. The marginal can be obtained simply by integrating (19) over the region

$$\begin{aligned}\Pr\left(Y_F > \omega_k X_F\right) &= \int_0^\infty \int_{\omega_k x}^\infty \left(\frac{\beta_{\text{sum}}}{x^2 y^2}\, e^{-1/x} e^{-\beta_{\text{sum}}/y}\right) dy dx \\ &= \frac{\alpha_k \beta_{\text{sum}}}{\beta_k + \alpha_k \beta_{\text{sum}}}.\end{aligned}$$

Hence, the conditional density is

$$\begin{aligned}f_{(X_F, Y_F)}(x, y \mid Y_F > \omega_k X_F) &= \frac{\Pr\left(X_F \in dx, Y_F \in dy, Y_F > \omega_k X_F\right)}{\Pr\left(Y_F > \omega_k X_F\right)} \\ &= \mathbf{1}\left(y > \omega_k x\right) \frac{\beta_k + \alpha_k \beta_{\text{sum}}}{\alpha_k x^2 y^2}\, e^{-1/x} e^{-\beta_{\text{sum}}/y}.\end{aligned}$$

B.2 Type (ii) - On the line $Y_F = \omega_j X_F$, $j = 1, \ldots, k$

This case only occurs if $X_F = \alpha_j Z_j$ and $Y_F = \beta_j Z_j$. Hence,

$$\Pr\left(X_F \in dx, Y_F \in dy, Y_F = \omega_j X_F\right) = \Pr\left(\alpha_j Z_j \in dx, \beta_j Z_j \in dy, Y_F = \omega_j X_F\right).$$

Similarly to the approach in Section B.1, this can be written as

$$\begin{aligned}&\Pr\left(X_F \in dx, Y_F \in dy, Y_F = \omega_j X_F\right) \\ &= \mathbf{1}\left(y = \omega_j x\right) \Pr\left(Z_j \in \frac{dx}{\alpha_j}\right) \prod_{p=1}^{j-1} \Pr\left(Z_p < \frac{x}{\alpha_p}\right) \prod_{p=j+1}^{k} \Pr\left(Z_p < \frac{y}{\beta_p}\right) \\ &\times \prod_{p=k+1}^{k+\ell} \Pr\left(Z_p < \frac{x}{\alpha_p}\right) \prod_{p=k+\ell+1}^{m} \Pr\left(Z_p < \frac{y}{\beta_p}\right) dx dy \\ &= \mathbf{1}\left(y = \omega_j x\right) \frac{\alpha_j}{x^2} e^{-\alpha_j/x} \prod_{p=1}^{j-1} \left(e^{-\alpha_p/x}\right) \prod_{p=j+1}^{k} \left(e^{-\beta_p/y}\right) \prod_{p=k+1}^{k+\ell} \left(e^{-\alpha_p/x}\right) \prod_{p=k+\ell+1}^{m} \left(e^{-\beta_p/y}\right) dx dy \\ &= \mathbf{1}\left(y = \omega_j x\right) \frac{\alpha_j}{x^2} \exp\left(-\frac{\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)}}{\beta_j x}\right) dx dy.\end{aligned}$$

By integrating this from 0 to ∞ we can obtain the marginal as

$$\Pr\left(Y_F = \omega_j X_F\right) = \frac{\alpha_j \beta_j}{\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)}}.$$

Hence, the conditional density is

$$f_{(X_F, Y_F)}(x, y \mid Y_F = \omega_j X_F) = \mathbf{1}\left(y = \omega_j x\right) \frac{\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)}}{\beta_j x^2} \exp\left(-\frac{\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)}}{\beta_j x}\right).$$

B.3 Type (iii) - Between the two lines $Y_F = \omega_j X_F$ and $Y_F = \omega_{j+1} X_F$, $j = 1, \ldots, k-1$

From the condition that $\omega_j X_F < Y_F < \omega_{j+1} X_F$, it follows that the pairs that lead to this case are combinations of $X_F = \alpha_i Z_i$ where $i = 1, \ldots, j, k+1, \ldots, k+\ell$, and $Y_F = \beta_h Z_h$ where $h = j+1, \ldots, m$. Hence,

$$\begin{aligned}&\Pr\left(X_F \in dx, Y_F \in dy, \omega_j X_F < Y_F < \omega_{j+1} X_F\right)\\ &\qquad = \sum_{h\in\mathcal{H}}\sum_{i\in\mathcal{I}} \Pr\left(\alpha_i Z_i \in dx, \beta_h Z_h \in dy, \omega_j X_F < Y_F < \omega_{j+1} X_F\right),\end{aligned}$$

where $\mathcal{I} = \{1, \ldots, j, k+1, \ldots, k+\ell\}$ and $\mathcal{H} = \{j+1, \ldots, m\}$. Then, due to the independence of the Z_i's this can be written as

$$\begin{aligned}&\Pr\left(X_F \in dx, Y_F \in dy, \omega_j X_F < Y_F < \omega_{j+1} X_F\right)\\ &= \mathbf{1}\left(\omega_j x < y < \omega_{j+1} x\right) \sum_{h\in\mathcal{H}}\sum_{i\in\mathcal{I}} \left\{ \Pr\left(Z_i \in \frac{dx}{\alpha_i}\right) \Pr\left(Z_h \in \frac{dy}{\beta_h}\right) \prod_{p=1,\{p\neq i\}}^{j} \Pr\left(Z_p < \frac{x}{\alpha_p}\right)\right.\\ &\times \left.\prod_{p=j+1,\{p\neq h\}}^{k} \Pr\left(Z_p < \frac{y}{\beta_p}\right) \prod_{p=k+1,\{p\neq i\}}^{k+\ell} \Pr\left(Z_p < \frac{x}{\alpha_p}\right) \prod_{p=k+\ell+1,\{p\neq h\}}^{m} \Pr\left(Z_p < \frac{y}{\beta_p}\right)\right\} dxdy\\ &= \mathbf{1}\left(\omega_j x < y < \omega_{j+1} x\right) \sum_{h\in\mathcal{H}}\sum_{i\in\mathcal{I}} \left\{ \frac{\alpha_i}{x^2} e^{-\alpha_i/x} \frac{\beta_h}{y^2} e^{-\beta_h/y} \prod_{p=1,\{p\neq i\}}^{j} \left(e^{-\alpha_p/x}\right)\right.\\ &\times \left.\prod_{p=j+1,\{p\neq h\}}^{k} \left(e^{-\beta_p/y}\right) \prod_{p=k+1,\{p\neq i\}}^{k+\ell} \left(e^{-\alpha_p/x}\right) \prod_{p=k+\ell+1,\{p\neq h\}}^{m} \left(e^{-\beta_p/y}\right)\right\} dxdy\\ &= \mathbf{1}\left(\omega_j x < y < \omega_{j+1} x\right) \frac{\alpha_{\text{sum}}^{(j)} \beta_{\text{sum}}^{(j+1)}}{x^2 y^2} e^{-\alpha_{\text{sum}}^{(j)}/x} e^{-\beta_{\text{sum}}^{(j+1)}/y} dxdy.\end{aligned}$$

Integrating this over the range $0 < x < \infty$ and $\omega_j x < y < \omega_{j+1} x$, we obtain the marginal as

$$\Pr\left(\omega_j X_F < Y_F < \omega_{j+1} X_F\right) = \frac{\alpha_{\text{sum}}^{(j)} \beta_{\text{sum}}^{(j+1)} (\alpha_j \beta_{j+1} - \alpha_{j+1}\beta_j)}{(\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)})(\alpha_{j+1}\beta_{\text{sum}}^{(j+1)} + \beta_{j+1}\alpha_{\text{sum}}^{(j)})}.$$

Using conditional probability, the conditional density is,

$$\begin{aligned}&f_{(X_F,Y_F)}(x, y \mid \omega_j X_F < Y_F < \omega_{j+1} X_F)\\ &\qquad = \mathbf{1}\left(\omega_j x < y < \omega_{j+1} x\right) \left\{ \frac{(\alpha_j \beta_{\text{sum}}^{(j+1)} + \beta_j \alpha_{\text{sum}}^{(j)})(\alpha_{j+1}\beta_{\text{sum}}^{(j+1)} + \beta_{j+1}\alpha_{\text{sum}}^{(j)})}{(\alpha_j \beta_{j+1} - \alpha_{j+1}\beta_j) x^2 y^2} \right\} e^{-\alpha_{\text{sum}}^{(j)}/x} e^{-\beta_{\text{sum}}^{(j+1)}/y}.\end{aligned}$$

B.4 Type (iv) - Below the line $Y_F = \omega_1 X_F$

From the condition that $Y_F < \omega_1 X_F$, it can be established that the pairs that can lead to this case are combinations of $X_F = \alpha_i Z_i$ and $Y_F = \beta_h Z_h$ where $i = k+1, \ldots, k+\ell$ and $h = 1, \ldots, k, k+\ell+1, \ldots, m$. Hence,

$$\Pr\left(X_F \in dx, Y_F \in dy, Y_F < \omega_1 X_F\right) = \sum_{h\in\mathcal{H}}\sum_{i\in\mathcal{I}} \Pr\left(\alpha_i Z_i \in dx, \beta_h Z_h \in dy, Y_F < \omega_1 X_F\right),$$

where $\mathcal{I} = \{k+1, \ldots, k+\ell\}$ and $\mathcal{H} = \{1, \ldots, k, k+\ell+1, \ldots, m\}$. The Z_i's are independent Fréchet random variables, hence

$$
\begin{aligned}
&\Pr\left(X_F \in dx, Y_F \in dy, Y_F < \omega_1 X_F\right) \\
&= \mathbf{1}\left(y < \omega_1 x\right) \sum_{h\in\mathcal{H}} \sum_{i\in\mathcal{I}} \left\{ \Pr\left(Z_i \in \frac{dx}{\alpha_i}\right) \Pr\left(Z_h \in \frac{dy}{\beta_h}\right) \prod_{p=1,\{p\neq h\}}^{k} \Pr\left(Z_p < \frac{y}{\beta_p}\right) \right. \\
&\times \left. \prod_{p=k+1,\{p\neq i\}}^{k+\ell} \Pr\left(Z_p < \frac{x}{\alpha_p}\right) \prod_{p=k+\ell+1,\{p\neq h\}}^{m} \Pr\left(Z_p < \frac{y}{\beta_p}\right) \right\} dxdy \\
&= \mathbf{1}\left(y < \omega_1 x\right) \sum_{h\in\mathcal{H}} \sum_{i\in\mathcal{I}} \left\{ \frac{\alpha_i}{x^2} e^{-\alpha_i/x} \frac{\beta_h}{y^2} e^{-\beta_h/y} \prod_{p=1,\{p\neq h\}}^{k} \left(e^{-\beta_p/y}\right) \prod_{p=k+1,\{p\neq i\}}^{k+\ell} \left(e^{-\alpha_p/x}\right) \right. \\
&\times \left. \prod_{p=k+\ell+1,\{p\neq h\}}^{m} \left(e^{-\beta_p/y}\right) \right\} dxdy \\
&= \mathbf{1}\left(y < \omega_1 x\right) \frac{\alpha_{\mathrm{sum}}}{x^2 y^2} e^{-\alpha_{\mathrm{sum}}/x} e^{-1/y} dxdy.
\end{aligned}
$$

The marginal can be obtained by integrating this over the region.

$$
\Pr\left(Y_F < \omega_1 X_F\right) = \frac{\alpha_{\mathrm{sum}}\beta_1}{\alpha_{\mathrm{sum}}\beta_1 + \alpha_1}.
$$

Hence, the conditional density is

$$
f_{(X_F,Y_F)}(x, y \mid Y_F < \omega_1 X_F) = \mathbf{1}\left(y < \omega_1 x\right) \frac{\alpha_{\mathrm{sum}}\beta_1 + \alpha_1}{\beta_1 x^2 y^2} e^{-\alpha_{\mathrm{sum}}/x} e^{-1/y}.
$$

References

[1] S. G. Coles, J. E. Heffernan, and J. A Tawn. Dependence measures for extreme value analyses. *Extremes*, 2:339–365, 1999.

[2] R. A. Davis and S. I. Resnick. Basic properties and prediction of max-ARMA processes. *Adv. App. Prob.*, 21:781–803, 1989.

[3] A. C. Davison, R. Huser, and E. Thibaud. Geostatistics of dependent and asymptotically independent extremes. *Math. Geosci.*, 45:511–529, 2013.

[4] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer, 2006.

[5] P. Deheuvels. Point processes and multivariate extreme values. *J. Multivariate Anal.*, 272:257–272, 1983.

[6] A.-L. Fougères, C. Mercadier, and J.P. Nolan. Dense classes of multivariate extreme value distributions. *J. Mult. Anal.*, 116:109–129, 2013.

[7] J. E. Heffernan and J. A. Tawn. A conditional approach for multivariate extreme values. *J. Roy. Statist. Soc. B*, 66:497–546, 2004.

[8] P. Jonathan, K. C. Ewans, and D. Randell. Joint modelling of extreme ocean environments incorporating covariate effects. *Coastal Eng.*, 79:22–31, 2013.

[9] C. Keef, I. Papastathopoulos, and J. A. Tawn. Estimation of the conditional distribution of a vector variable given that one of its components is large: additional constraints for the Heffernan and Tawn model. *J. Mult. Anal.*, 115:396–404, 2013.

[10] M. Keresztúri. *Assessing and Modelling Extremal Dependence in Spatial Extremes.* PhD thesis, Lancaster University, Lancaster, 2016. In preparation.

[11] M. Keresztúri, J.A. Tawn, and P. Jonathan. Assessing extremal dependence of North Sea storm severity. *Oce. Eng.*, 118:242–259, 2016.

[12] A. W. Ledford and J. A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187, 1996.

[13] A. W. Ledford and J. A. Tawn. Modelling dependence within joint tail regions. *J. Roy. Statist. Soc. B*, 59:475–499, 1997.

[14] A. W. Marshall and I. Olkin. A multivariate exponential distribution. *J. Amer. Statist. Assoc.*, 62: 30–44, 1967.

[15] I. Papastathopoulos and J. A. Tawn. Conditioned limit laws for inverted max-stable processes. *J. Mult. Anal.*, 150:214–228, 2016.

[16] J. Pickands. Multivariate extreme value distributions. In *Bulletin of the International Statistical Institute: Proceedings of the 43rd Session (Buenos Aires)*, pages 859–878. Vorburg, Netherlands: ISI, 1981.

[17] S. I. Resnick. *Extreme Values, Regular Variation, and Point Processes.* Springer, 1987.

[18] E. Thibaud, R. Mutzner, and A. C. Davison. Threshold modelling of extreme spatial rainfall. *Water Resources Research*, 49:4633–4644, 2013.

[19] J. L. Wadsworth and J. A. Tawn. Dependence modelling for spatial extremes. *Biometrika*, 99:253–272, 2012.

[20] J. L. Wadsworth and J. A. Tawn. A new representation for multivariate tail probabilities. *Bernoulli*, 19:2689–2714, 2013.

[21] Y. Wang and S. A. Stoev. Conditional sampling for spectrally discrete max-stable random fields. *Adv. Appl. Prob.*, 43:461–483, 2011.