An alternative point process framework for modelling multivariate extreme values

Alexandra Ramos*

Faculdade de Economia da Universidade do Porto,

Rua Dr Roberto Frias, 4200-464-Porto, Portugal

and

CMUP,

Rua do Campo Alegre, 687, 4169-007 Porto, Portugal.

Anthony Ledford†

Oxford-Man Institute of Quantitative Finance, University of Oxford,

Eagle House, Walton Well Road, Oxford, OX2 6ED, UK

and

AHL Research, Man Investments, Sugar Quay, Lower Thames Street, London EC3R 6DU

October 15, 2009

Abstract

Classical techniques for analysing multivariate extremes can often be framed in terms of the point process representation of de Haan (1985). Amongst other things, this representation provides a characterisation of the limiting distribution of the normalised componentwise maxima of independent and identically distributed unit Fréchet variables, i.e. the class of multivariate extreme value distributions. The dependence structures accommodated within this class correspond only to asymptotic dependence or to exact independence, and so are rather restrictive.

In this paper, an alternative limiting point process representation is studied that holds regardless of whether the underlying data generation mechanism is asymptotically dependent or asymptotically independent. Through the use of the usual pseudo-polar coordinates, we characterise the intensity function of this point process in terms of the coefficient of tail dependence $\eta \in (0, 1]$ and a non-negative measure that has to satisfy a simple normalisation condition but is otherwise arbitrary. We use this point process representation to derive an analogue of the standard componentwise maxima result that holds for both asymptotically dependent and asymptotically independent cases. We illustrate our results using a flexible parametric example and provide methods for simulating from both the limiting point process and the limiting componentwise maxima distribution.

*aramos@fep.up.pt

†ALedford@maninvestments.com

Keywords: Extreme value theory, multivariate extreme values, point processes, asymptotic dependence, asymptotic independence, coefficient of tail dependence, componentwise maxima, joint survivor function, simulation.

ams: 60G70; 62E20

1 Introduction

1.1 Classical background

The classical results of multivariate extremes can often be framed in terms of the point process representation of de Haan (1985), see for example Coles and Tawn (1991, 1994). To illustrate this, in the following we concentrate on the 2-dimensional case and consider independent and identically distributed (i.i.d.) bivariate random variables $(X_1, Y_1), \ldots, (X_n, Y_n)$ with unit Fréchet distributed margins, so that $\Pr(X \leq x) = \Pr(Y \leq x) = \exp(-1/x)$ on $x > 0$, and joint distribution function $F(x, y)$ where F is assumed to be in the domain of attraction of a bivariate extreme value (BEV) distribution (Resnick, 1987). The choice of unit Fréchet distributed margins admits no loss of generality here, as probability integral transformations may be used to derive corresponding results for arbitrary marginal distributions. Defining the componentwise maxima $M_{X,n} = \max(X_1, \ldots, X_n)$ and $M_{Y,n} = \max(Y_1, \ldots, Y_n)$, it is clear $\Pr(M_{X,n}/n \leq x) = \Pr(M_{Y,n}/n \leq x) = \exp(-1/x)$ so that $M_{X,n}/n$ and $M_{Y,n}/n$ are both exactly unit Fréchet distributed, or more loosely, a normalisation of n^{-1} is required in order to stabilise the componentwise maxima.

Following de Haan (1985), it can be shown that the point process $\mathcal{P}_n$ defined by

$$\mathcal{P}_n = \left\{ \left(\frac{X_i}{n}, \frac{Y_i}{n} \right) ;\ i = 1, \ldots, n \right\}$$

converges weakly in the limit as $n \to \infty$ to a non-homogeneous Poisson process $\mathcal{P}$ on $[0, \infty) \times [0, \infty) \setminus (0, 0)$. Furthermore, by changing variables to the pseudo-radial and angular coordinates $R = X + Y$ and $W = X/R$, it can be shown that $\mathcal{P}$ has a point intensity μ that factorises as

$$\mu(\mathrm{d}r \times \mathrm{d}w) = r^{-2}\,\mathrm{d}r\,\mathrm{d}H(w)$$

where H is a non-negative measure on $[0, 1]$ that is arbitrary apart from having to satisfy

$$\int_0^1 w\,\mathrm{d}H(w) = \int_0^1 (1 - w)\ \mathrm{d}H(w) = 1. \tag{1.1}$$

An equivalent but more straightforward way to express the conditions (1.1) is that $H/2$ is the cumulative probability distribution function of any random variable on $[0, 1]$ that has mean equal to $1/2$.

As an illustration of how $\mathcal{P}$ can be used to derive classical multivariate extreme value results, we use it here to derive the general characterisation of a BEV distribution with unit Fréchet distributed margins. Following the set up above, let G denote any non-

-degenerate limiting bivariate distribution function of $(M_{X,n}/n, M_{Y,n}/n)$ and define the set $A_{xy} = \{(0,x] \times (0,y]\}$ for fixed positive x and y. Then

$$\begin{aligned}
G(x,y) &= \lim_{n\to\infty} \Pr\left(\frac{M_{X,n}}{n} \le x, \frac{M_{Y,n}}{n} \le y\right) = \lim_{n\to\infty} \Pr(\text{all points of } \mathcal{P}_n \text{ are in } A_{xy}) \\
&= \lim_{n\to\infty} \Pr(\text{no points of } \mathcal{P}_n \text{ are outside } A_{xy}) \\
&= \Pr(\text{no points of } \mathcal{P} \text{ are outside } A_{xy}) \\
&= \exp(-1 \times \text{Expected number of points of } \mathcal{P} \text{ outside } A_{xy}) \\
&= \exp\left\{-\int_0^1 \int_{r_{xy}(w)}^{\infty} r^{-2}\,\mathrm{d}r\,\mathrm{d}H(w)\right\} \qquad (1.2)
\end{aligned}$$

where $r_{xy}(w) = \min\{x/w, y/(1-w)\}$. On performing the r-integration in equation (1.2) we obtain $G(x,y) = \exp\{-V(x,y)\}$ where V is a function that determines the dependence structure of $G(x,y)$ and is given by

$$V(x,y) = \int_0^1 \max\left(\frac{w}{x}, \frac{1-w}{y}\right) \mathrm{d}H(w) \qquad (1.3)$$

for H satisfying the regularity conditions (1.1). This is the classical representation of a BEV distribution with unit Fréchet distributed margins, see de Haan and Resnick (1977) and Pickands (1981). It is easy to see from equation (1.3) that V is a homogeneous function of order -1, that is $V(nx, ny) = n^{-1}V(x,y)$. Hence we have that G is max-stable since $G^n(nx, ny) = G(x,y)$, and G is also in its own domain of attraction since if each $(X_i, Y_i) \sim G$ in the above then $(M_{X,n}/n, M_{Y,n}/n) \sim G$ for each $n \ge 1$.

1.2 Limitations of the classical theory

The notions of asymptotic dependence and asymptotic independence are important for describing the limitations of the classical theory and so are now defined. Letting κ denote the limiting conditional probability in

$$\lim_{u\to\infty} \Pr(X > u | Y > u) = \kappa$$

then X and Y are said to be asymptotically dependent if $\kappa > 0$ and asymptotically independent if $\kappa = 0$ (Sibuya, 1960). Asymptotic independence is an important feature that arises in both practical data modelling applications and the study of the dependence features of most classical families of distributions, see de Haan and de Ronde (1998), Ledford and Tawn (1997), Capéraà *et al.* (2000) and Heffernan (2000).

When the components of the underlying (X_i, Y_i) pairs in Section 1.1 are asymptotically dependent then the de Haan (1985) limiting point process representation provides a sound foundation for modelling the observed joint tail dependence structure, and consequently constructions derived from this representation such as the class of BEV distributions are well motivated and typically useful in practice. This is not the case when the (X_i, Y_i) components are asymptotically independent because then the limiting point process $\mathcal{P}$ is degenerate as the measure $H(w)$ is identically zero on the interior of $[0,1]$ and places unit masses at $w = 0$ and $w = 1$. In this case equations (1.2) and (1.3) give that the resulting BEV distribution has exactly independent margins since $G(x, y) = \exp\{-(1/x + 1/y)\}$. The fact that the limiting point process and hence constructions derived from it such as the class of BEV distributions all correspond either to exact independence or asymptotic dependence leads to non-regular inference issues, see Tawn (1988), Ledford and Tawn (1996) and Ramos and Ledford (2005), and considerable difficulties in situations where asymptotic independence is apparent, as neither exact independence nor asymptotic dependence may be adequate for describing the observed joint tail dependence structure, see Ledford and Tawn (1996, 1997) and Coles *et al.* (1999). The crux of the problem is that the currently available underlying limit theory is degenerate for cases of asymptotic independence. It is our view that an alternative limit theory is required that provides useful results for both asymptotic dependence and asymptotic independence. This is what is provided in this paper.

1.3 Alternatives to the classical approach

Over the last decade or so there has been progress on developing extremal dependence models that accommodate both asymptotic dependence and asymptotic independence. Ledford and Tawn (1997) proposed a method based on joint survivor functions that could be applied to both asymptotically dependent and asymptotically independent data but overlooked some necessary regularity conditions and used a parametric form that was not guaranteed to correspond to a valid probability model. More recently, Ramos and Ledford (2009) extended the Ledford and Tawn (1997) results by establishing the required regularity conditions, developing a flexible parametric model that satisfied these and deriving a censored likelihood approach for estimation. Coles and Pauli (2002) developed a different approach based on a parametric copula that describes dependence throughout the entire domain rather than just the extremes, and Heffernan and Tawn (2004) derived a method with a more empirical flavour based on combining normalised limiting conditional distributions. However none of these developments has provided a modelling framework with generality and depth comparable to those of the de Haan (1985) point process and its encapsulation of other classical extreme results such as the characterisation (1.2). Resnick (2002) and Maulik and Resnick (2004) also gave a characterization and representation of distributions possessing a property they called hidden regular variation, which is a property of the subfamily of distributions possessing both multivariate regular variation and asymptotic independence, and is based on an elaboration the coefficient of tail dependence of Ledford and Tawn (1996, 1997). In these two works hidden regular variation is represented in terms of a non-negative measure called the hidden angular (or spectral) measure, which was estimated by Heffernan and Resnick (2005) using the rank transform. We construct a similar representation via our point process approach and thereby obtain a simple normalization condition that the hidden angular measure must satify. This condition is missing in the hidden regular variation papers cited above.

The starting point for the developments of this paper is the limiting Poisson process result Ledford and Tawn (1997) outlined but did not explore in any detail. Focusing on the bivariate case, we derive regularity conditions on the intensity function of this point process in Section 2. In Section 3 we use our point process results to derive the analogue of the classical componentwise maxima results given in Section 1.1. A flexible parametric example is introduced in Section 4 that is used to illustrate the new point process and componentwise maxima results, and also to demonstrate how the joint tail modelling approach of Ramos and Ledford (2009) can be derived from the point process results developed in this paper. Methods for simulating from both the underlying point process and the limiting distribution of componentwise maxima are given in Section 5. In Section 6 we provide some concluding remarks and show how the bivariate results discussed in this paper can be extended to the multivariate case. In the appendices we provide some proofs and discuss a possible extension of our results that may be useful for further relating our approach to the classical extreme values one.

2 A unified point process framework

In this section we build on the limiting Poisson process results of Ledford and Tawn (1997) and show that the intensity function which arises can be written in terms of a non-negative measure that is arbitrary apart from having to satisfy some regularity conditions. Our treatment unifies the asymptotically dependent and asymptotically independent cases. The convergence of the relevant sequence of point processes to a Poisson process on the cone $(0, \infty)^2$ and the convergence of the associated intensity measure can also be obtained from the results of Resnick (2002), Maulik and Resnick (2004) and Heffernan and Resnick (2005).

2.1 Derivation

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ denote independent and identically distributed bivariate random variables with unit Fréchet distributed margins and joint survivor function that is bivariate regularly varying with index $-1/\eta$ for some $\eta \in (0, 1]$ so that

$$\overline{F}(x, y) = \Pr(X > x, Y > y) = \frac{\mathcal{L}(x, y)}{(xy)^{1/(2\eta)}} \tag{2.1}$$

where $\mathcal{L}$ is a bivariate slowly varying (BSV) function. Following Ledford and Tawn (1996, 1997) we refer to the parameter $\eta \in (0, 1]$ as the coefficient of tail dependence. The representation (2.1) accommodates a wide spectrum of dependence behaviour, including asymptotic independence, asymptotic dependence, complete dependence and exact independence. As a modelling framework, it also provides a smooth transition between complete dependence and exact independence, and allows negative dependence between the marginal extremes, yielding a very flexible and broadly applicable statistical model, see Ramos and Ledford (2009). Let g denote the limit function of $\mathcal{L}$, so that for all $(x, y) \in (0, \infty) \times (0, \infty)$ and $c > 0$ we have

$$g(x, y) = \lim_{r \to \infty} \frac{\mathcal{L}(rx, ry)}{\mathcal{L}(r, r)} \quad \text{and} \quad g(cx, cy) = g(x, y). \tag{2.2}$$

A consequence of equation (2.2) is that for any $x > 0$ and $y > 0$

$$g(x, y) = g\{x/(x + y), y/(x + y)\} = g\{w, 1 - w\} \equiv g^*(w) \tag{2.3}$$

where $w = x/(x+y) \in (0,1)$, and thus it is clear that the bivariate function $g(x,y)$ depends only on the univariate quantity $w \in (0,1)$.

Following Ledford and Tawn (1997), we consider the sequence of point processes

$$\mathcal{P}_n^* = \left\{ \left(\frac{X_i}{b_n}, \frac{Y_i}{b_n} \right) : i = 1, \dots, n \right\} \tag{2.4}$$

where the normalising constants b_n are defined implicitly as satisfying $n\overline{F}(b_n, b_n) = 1$, so that typically $b_n = O(n^\eta)$. Clearly when $\eta < 1$ the normalising constants satisfy $b_n = o(n)$ and are therefore too light to stabilise the componentwise maxima $M_{X,n}$ and $M_{Y,n}$ within $\mathcal{P}_n^*$, see Section 1. In contrast, when the componentwise maxima are asymptotically dependent then $b_n = O(n)$ and $M_{X,n}$ and $M_{Y,n}$ are stabilised. The impact of this distinction is that $\mathcal{P}_n^*$ converges weakly to a non-homogeneous Poisson process $\mathcal{P}^*$ in the limit as $n \to \infty$ with domain $(0,\infty) \times (0,\infty)$ when $\eta < 1$ and domain $\{[0,\infty) \times [0,\infty)\} \backslash \{(0,0)\}$ when the componentwise maxima are asymptotically dependent. The first case is discussed in more detail in Ledford and Tawn (1997) and the latter case in de Haan (1985), however for our purposes we will restrict attention to the domain $(0,\infty) \times (0,\infty)$ as then our results will hold regardless of whether the componentwise maxima are asymptotically dependent or asymptotically independent.

We start by deriving the intensity function $\Lambda_\eta\{(x,\infty) \times (y,\infty)\}$ of $\mathcal{P}^*$ for any $x > 0$ and $y > 0$, that is the expected number of points of $\mathcal{P}^*$ within a set of the form $\{(x,\infty) \times (y,\infty)\}$. By definition, and substituting according to equations (2.1) and (2.2), we have

$$\begin{aligned} \Lambda_\eta\{(x,\infty) \times (y,\infty)\} &= \lim_{n\to\infty} n\overline{F}(b_n x, b_n y) \\ &= \lim_{n\to\infty} n\overline{F}(b_n, b_n) \, \frac{\overline{F}(b_n x, b_n y)}{\overline{F}(b_n, b_n)} = \frac{g(x,y)}{(xy)^{1/(2\eta)}} \end{aligned} \tag{2.5}$$

since $n\overline{F}(b_n, b_n) = 1$. Substituting into equation (2.5) according to the pseudo-polar coordinates $r = x + y$ and $w = x/r$ and exploiting equation (2.3), it is clear that

$$\Lambda_\eta\{(x,\infty) \times (y,\infty)\} = r^{-1/\eta} g^*(w) / \{w(1-w)\}^{1/(2\eta)}$$

and thus for any $x > 0$ and $y > 0$ we have $\Lambda_\eta\{(x,\infty) \times (y,\infty)\}$ factorises into separate terms involving r and w. We now derive an alternative expression for $\Lambda_\eta\{(x,\infty) \times (y,\infty)\}$ as the integral of its point intensity μ_η over the set $\{(x,\infty) \times (y,\infty)\}$ and then exploit the above factorisation by changing variables to (r,w), and so obtain

$$\begin{aligned} \Lambda_\eta\{(x,\infty) \times (y,\infty)\} &= \int_{\{(x,\infty)\times(y,\infty)\}} \mu_\eta(\mathrm{d}x \times \mathrm{d}y) \\ &= \int_{w\in(0,1)} \int_{r=r'_{xy}(w)}^{\infty} r^{-(1+1/\eta)} \, \mathrm{d}r \, \mathrm{d}H_\eta(w) \qquad (2.6) \\ &= \eta \int_{(0,1)} \left\{ \min\left(\frac{w}{x}, \frac{1-w}{y} \right) \right\}^{1/\eta} \mathrm{d}H_\eta(w) \qquad (2.7) \end{aligned}$$

where $r'_{xy}(w) = \max\{x/w, y/(1-w)\}$ and $H_\eta(w)$ is a non-negative measure on $(0,1)$ (see Ramos and Ledford, 2009 or Beirlant *et al.*, 2004) that is arbitrary apart from having to satisfy a simple normalisation condition that we now derive. Since $\lim_{n\to\infty} n\overline{F}(b_n, b_n) = 1$ implies that the expected number of points of $\mathcal{P}^*$ within $(1,\infty)\times(1,\infty)$ is 1 also, then η and H_η are related via

$$1 = \Lambda_\eta\{(1,\infty)\times(1,\infty)\} = \eta \int_{(0,1)} \{\min(w, 1-w)\}^{1/\eta}\, \mathrm{d}H_\eta(w),$$

or equivalently H_η must satisfy the following normalisation condition:

$$\eta^{-1} = \int_0^{1/2} w^{1/\eta}\, \mathrm{d}H_\eta(w) + \int_{1/2}^{1} (1-w)^{1/\eta}\, \mathrm{d}H_\eta(w). \tag{2.8}$$

The measure H_η is a particular case of the hidden angular measure considered by Resnick (2002) and Maulik and Resnick (2004), however these papers omit the normalisation condition (2.8)

The results obtained above add significant rigour to the treatment provided by Ledford and Tawn (1997) and extend the hidden regular variation results of Resnick and collaborators to include the integral constraint (2.8). The normalisation condition (2.8) is identical to that obtained in Ramos and Ledford (2009) but is different to that which arises in the de Haan (1985) classical point process treatment of multivariate extremes, see equation (1.1). Condition (2.8), although necessary, is not sufficient to ensure that an (η, H_η)-pair defines the intensity of a point process that can arise via construction (2.4) from a bivariate random variable (X, Y) with unit Frechét distributed marginals. As in Theorem 1 of Ramos and Ledford (2009), additional conditions are required for this so that the marginal tails of the point process are not heavier than those of a unit Frechét variable. See Appendix A for details.

In order to provide a brief illustration we here examine a simple discrete model. Following Ramos and Ledford (2009), suppose initially that H_η has two atoms of equal mass a at positions w_1 and $(1-w_1)$ in the interior of $(0,1)$ so that $0 < w_1 < 1/2$. The normalisation condition (2.8) implies $a = (2\eta)^{-1} w_1^{-1/\eta}$. Generalising this to the case of an asymmetric discrete measure with positive masses $a_1, \ldots, a_q, \ldots, a_n$ at positions $0 < w_1 < \cdots < w_q < \cdots < w_n < 1$, where q is such that $w_{q-1} \le 1/2 < w_q$, the normalisation condition (2.8) is satisfied if and only if

$$\sum_{i=1}^{q-1} a_i w_i^{1/\eta} + \sum_{i=q}^{n} a_i (1-w_i)^{1/\eta} = \eta^{-1}.$$

The additional conditions in Appendix A are satisfied for any $\eta \in (0,1)$. We do not examine this model any further in this paper but instead concentrate later on a smooth and flexible parametric form for H_η that is more suited to applications.

3 Componentwise maxima revisited

In this section we study an analogue of result (1.2) based on the sequence of point processes $\mathcal{P}_n^*$ of Section 2. This topic was examined briefly by Ledford and Tawn (1997), however the more detailed treatment given here makes clear some issues that are not apparent within their treatment. Our approach is to obtain the limiting distribution of normalised componentwise maxima of points that occur within a region bounded away from the axes and then examine the behaviour of this distribution as the bounds approach the axes.

Under the same setup as Section 2, we additionally define $R_\varepsilon = \{(x, y) : x > \varepsilon, y > \varepsilon\}$ and $R_\varepsilon(x, y) = R_\varepsilon \setminus \{(\varepsilon, x) \times (\varepsilon, y)\}$, and let $M_{X,n,\varepsilon}$ and $M_{Y,n,\varepsilon}$ denote the componentwise maxima of those points out of $(X_1, Y_1), \ldots, (X_n, Y_n)$ that occur within R_ε. Then for any $x > \varepsilon$ and $y > \varepsilon$ we have

$$
\begin{aligned}
\Pr&\left(\frac{M_{X,n,\varepsilon b_n}}{b_n} \le x, \frac{M_{Y,n,\varepsilon b_n}}{b_n} \le y\right) \\
&= \Pr\{\text{no } (X_i, Y_i) \text{ pair occurs within } R_{\varepsilon b_n}(x b_n, y b_n)\} \\
&= \Pr\{\text{every } (X_i, Y_i) \text{ pair occurs outside } R_{\varepsilon b_n}(x b_n, y b_n)\} \\
&= \left\{1 - \overline{F}(x b_n, \varepsilon b_n) - \overline{F}(\varepsilon b_n, y b_n) + \overline{F}(x b_n, y b_n)\right\}^n \\
&= \left[1 - \frac{1}{n}\left\{\frac{\overline{F}(x b_n, \varepsilon b_n)}{\overline{F}(b_n, b_n)} + \frac{\overline{F}(\varepsilon b_n, y b_n)}{\overline{F}(b_n, b_n)} - \frac{\overline{F}(x b_n, y b_n)}{\overline{F}(b_n, b_n)}\right\}\right]^n
\end{aligned}
$$

since $n\overline{F}(b_n, b_n) = 1$. Substituting according to equation (2.5) and taking the limit as $n \to \infty$ gives

$$
\begin{aligned}
\lim_{n\to\infty} \Pr&\left(\frac{M_{X,n,\varepsilon b_n}}{b_n} \le x, \frac{M_{Y,n,\varepsilon b_n}}{b_n} \le y\right) \\
&= \exp\left[-\left\{\frac{g(x, \varepsilon)}{(x\varepsilon)^{1/(2\eta)}} + \frac{g(\varepsilon, y)}{(\varepsilon y)^{1/(2\eta)}} - \frac{g(x, y)}{(xy)^{1/(2\eta)}}\right\}\right] \qquad (3.1) \\
&= \exp\{-1 \times \text{Expected number of points of } \mathcal{P}^* \text{ that occur in } R_\varepsilon(x, y)\}. \qquad (3.2)
\end{aligned}
$$

To examine the terms in equation (3.1) as $\varepsilon \to 0$ we recall equations (2.5) and (2.7) and so obtain

$$
\frac{g(x, y)}{(xy)^{1/(2\eta)}} = \eta \int_{(0,1)} \left\{\min\left(\frac{w}{x}, \frac{1-w}{y}\right)\right\}^{1/\eta} \mathrm{d}H_\eta(w). \qquad (3.3)
$$

From equation (3.3) we therefore have

$$
\frac{g(x, \varepsilon)}{(x\varepsilon)^{1/(2\eta)}} = \eta \int_0^{x/(x+\varepsilon)} \left(\frac{w}{x}\right)^{1/\eta} \mathrm{d}H_\eta(w) + \eta \int_{x/(x+\varepsilon)}^1 \left(\frac{1-w}{\varepsilon}\right)^{1/\eta} \mathrm{d}H_\eta(w)
$$

and hence

$$
\lim_{\varepsilon\to 0} \frac{g(x, \varepsilon)}{(x\varepsilon)^{1/(2\eta)}} = \eta x^{-1/\eta} \int_{(0,1)} w^{1/\eta}\, \mathrm{d}H_\eta(w) + \eta \lim_{\varepsilon\to 0} \varepsilon^{-1/\eta} \int_{x/(x+\varepsilon)}^1 (1-w)^{1/\eta}\, \mathrm{d}H_\eta(w).
$$

Similarly, equation (3.3) gives

$$
\lim_{\varepsilon\to 0} \frac{g(\varepsilon, y)}{(\varepsilon y)^{1/(2\eta)}} = \eta \lim_{\varepsilon\to 0} \varepsilon^{-1/\eta} \int_0^{\varepsilon/(\varepsilon+y)} w^{1/\eta}\, \mathrm{d}H_\eta(w) + \eta y^{-1/\eta} \int_{(0,1)} (1-w)^{1/\eta}\, \mathrm{d}H_\eta(w).
$$

Thus provided the measure H_η satisfies the normalisation condition (2.8) and is finite then both limits on the right hand side of the above equations equal zero for any $x > 0$ and $y > 0$ respectively, and therefore taking the limit of equation (3.1) as $\varepsilon \to 0$ gives a non-degenerate result and

$$G_\eta(x,y) \equiv \lim_{\varepsilon\to 0}\lim_{n\to\infty} \Pr\left(\frac{M_{X,n,\varepsilon b_n}}{b_n} \le x, \frac{M_{Y,n,\varepsilon b_n}}{b_n} \le y\right) = \exp\{-V_\eta(x,y)\} \tag{3.4}$$

where

$$\begin{aligned} V_\eta(x,y) &= \text{Expected number of points of } \mathcal{P}^* \text{ that occur in } \lim_{\varepsilon\to 0} R_\varepsilon(x,y) \\ &= \eta \int_{(0,1)} \left\{\max\left(\frac{w}{x}, \frac{1-w}{y}\right)\right\}^{1/\eta} \mathrm{d}H_\eta(w). \end{aligned} \tag{3.5}$$

It is straightforward to see that the dependence function V_η defined in equation (3.5) is homogeneous of order $-1/\eta$, that is, $V_\eta(tx,ty) = t^{-1/\eta}V_\eta(x,y)$ for all $t > 0$. Thus we have $G_\eta^n(n^\eta x, n^\eta y) = G_\eta(x,y)$ and so G_η is max-stable. Examining now the marginal behaviour implied by equations (3.4) and (3.5) we obtain

$$\begin{aligned} \lim_{\varepsilon\to 0}\lim_{n\to\infty} \Pr\left(M_{X,n,\varepsilon b_n}/b_n \le x\right) &= \exp\left\{-V_\eta(1,\infty)\, x^{-1/\eta}\right\} \quad \text{and} \\ \lim_{\varepsilon\to 0}\lim_{n\to\infty} \Pr\left(M_{Y,n,\varepsilon b_n}/b_n \le y\right) &= \exp\left\{-V_\eta(\infty,1)\, y^{-1/\eta}\right\}, \end{aligned}$$

that is, Fréchet distributions with shape parameter η and scale parameters equal to $V_\eta(1,\infty) = \eta\int_0^1 w^{1/\eta}h_\eta(w)\,\mathrm{d}w$ and $V_\eta(\infty,1) = \eta\int_0^1 (1-w)^{1/\eta}\, h_\eta(w)\,\mathrm{d}w$ respectively. So, unlike for the classical results given in Section 1.1, the limiting marginal distributions here are not unit Fréchet. However, by adopting different normalising conditions in the point process, it is possible to obtain a limiting distribution that has Fréchet margins with the same shape parameter but scale parameters equal to 1. Specifically the point process

$$\mathcal{P}_n^{\#} = \left\{\left(\frac{X_i}{V_\eta^\eta(1,\infty)b_n}, \frac{Y_i}{V_\eta^\eta(\infty,1)b_n}\right) : i = 1,\ldots,n\right\}$$

achieves this.

Equations (3.4) and (3.5) give the joint distribution of componentwise maxima for pairs of variables which are simultaneously large regardless of whether the underlying dependence structure is asymptotically dependent or asymptotically independent, and thus considerably extend the standard componentwise maxima result given in equation (1.2). It is hoped that this asymptotically motivated framework will provide versatile and improved modelling and extrapolation capabilities compared to the classical BEV paradigm, especially in cases where neither exact independence nor asymptotic dependence is appropriate. The above also clarifies the point process results of Ledford and Tawn (1997) and makes clear the convergence issues. We remark that the limiting point process $\mathcal{P}^*$ of Section 2, in contrast to the de Haan (1985) approach, is based on a normalisation that stabilises the joint tail rather than the componentwise maxima and a domain that excludes the axes $x = 0$ and $y = 0$. If these key differences are omitted, then we are back in the classical BEV case and result (1.2) holds. This point seems to have been overlooked in the statement of the result given in Section 3.2.1 of Kotz and Nadarajah (2000).

The constraint of H_η being finite causes some restriction since, as noted by Maulik and Resnick (2004), the class of regularly varying multivariate distributions with infinite hidden angular measure is at least as large as the class of distributions on $[1, \infty)$ which have infinite mean. Thus an infinite measure H_η which is valid for the point process $\mathcal{P}^*$ of Section 2 will yield the degenerate limit of zero as $\varepsilon \to 0$ for the componentwise maxima distribution in equation (3.4). We will see an explicit example of this later. This emphasises that the point process representation in Section 2 accommodates a wider class of dependence behaviour than that captured by the componentwise maxima results of this section. Geometrically, what is happening when the componentwise maxima results are degenerate is that $M_{X,n,\varepsilon b_n}$ and $M_{Y,n,\varepsilon b_n}$ grow too fast to be stabilised by the b_n normalisation. We illustrate this later through simulation results for a particular parametric example.

3.1 Obtaining H_η from V_η

In equation (3.5) the dependence function V_η is defined in terms of a given measure H_η. Restricting ourselves here to the class of models for which V_η is differentiable, we now show how H_η, or more specifically its measure density h_η, can be recovered from a given V_η. Our approach is similar to that of Coles and Tawn (1991). Using, for example, equations (2.7) and (3.5), it is clear that

$$\Lambda_\eta\{(x,\infty) \times (y,\infty)\} = V_\eta(x,\infty) + V_\eta(\infty,y) - V_\eta(x,y), \tag{3.6}$$

so taking the mixed derivative of equation (3.6) with respect to x and y we obtain

$$\frac{\partial^2 V_\eta}{\partial x\, \partial y} = -r^{-(2+1/\eta)} h_\eta(w) \tag{3.7}$$

where $r = x + y$ and $w = x/(x + y)$. Thus provided the density h_η exists, equation (3.7) describes the relationship between the pseudo-angular measure H_η and the exponent measure V_η. This extends part of the Coles and Tawn (1991) result to the asymptotically independent case, however their result is more wide-ranging as it also covers the relationship between the lower dimensional pseudo-radial and exponent measures in higher dimensional cases.

4 A parametric example

In order to examine the properties of estimators and to undertake applied statistical modelling it is useful to have a tractable parametric model that is flexible enough to exhibit a wide range of behaviours and also allows straightforward simulation. Here we introduce such a model via a particular parametric H_η family. This model will be the focus for much of the remainder of the paper.

4.1 The η-asymmetric logistic model

This is similar to the BEV-asymmetric logistic dependence structure of classical bivariate extremes, see Tawn (1988). As defined in Ramos and Ledford (2009), consider H_η with density h_η given by

$$h_\eta(w) = \frac{\eta-\alpha}{\alpha\eta^2 N_\varrho}\left\{(\varrho w)^{-1/\alpha} + \left(\frac{1-w}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta-2} \{w\,(\,1-w)\}^{-(1+1/\alpha)} \tag{4.1}$$

for $w \in (0,1)$ where $N_\varrho = \varrho^{-1/\eta}+\varrho^{1/\eta}-(\varrho^{-1/\alpha}+\varrho^{1/\alpha})^{\alpha/\eta}$ and $\eta \in (0,1]$, $\varrho > 0$ and $\alpha \in (0,1]$. It is straightforward to show that h_η satisfies both the normalisation condition (2.8) and the conditions mentioned in Appendix A.

4.2 Point process intensity

For the η-asymmetric logistic model with measure density h_η satisfying equation (4.1), equation (2.7) gives that the intensity function of the limiting Poisson point process $\mathcal{P}^*$ satisfies

$$\Lambda_\eta\left\{(x,\infty)\times(y,\infty)\right\} = N_\varrho^{-1}\left[(\varrho x)^{-1/\eta} + \left(\frac{y}{\varrho}\right)^{-1/\eta} - \left\{(\varrho x)^{-1/\alpha} + \left(\frac{y}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}\right] \tag{4.2}$$

where $x > 0$ and $y > 0$.

4.3 Limiting distribution of componentwise maxima

For the η-asymmetric logistic model defined by equation (4.1), the limiting componentwise maxima distribution given in equation (3.4) satisfies

$$\begin{aligned} G_\eta(x,y) &= \lim_{\varepsilon\to 0}\lim_{n\to\infty} \Pr\left(M_{X,n,\varepsilon b_n} \le x b_n, M_{Y,n,\varepsilon b_n} \le y b_n\ \right) \\ &= \begin{cases} \exp\left[-N_\varrho^{-1}\left\{(\varrho x)^{-1/\alpha} + \left(\frac{y}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}\right] & \text{for } \alpha < \eta, \\ 0 & \text{for } \alpha \ge \eta. \end{cases} \end{aligned} \tag{4.3}$$

The marginal distributions of G_η are easily obtained from equation (4.3) by substituting $x = \infty$ or $y = \infty$ and hence satisfy

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} \Pr\left(M_{X,n,\varepsilon b_n} \le x b_n\right) = \begin{cases} \exp\left(-N_\varrho^{-1}\varrho^{-1/\eta}x^{-1/\eta}\right) & \text{for } \alpha < \eta, \\ 0 & \text{for } \alpha \ge \eta, \end{cases} \quad \text{and}$$

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} \Pr\left(M_{Y,n,\varepsilon b_n} \le y b_n\right) = \begin{cases} \exp\left(-N_\varrho^{-1}\varrho^{1/\eta}y^{-1/\eta}\right) & \text{for } \alpha < \eta, \\ 0 & \text{for } \alpha \ge \eta. \end{cases}$$

The degenerate $\alpha \ge \eta$ case in the above arises when the measure H_η is infinite, as discussed in Section 3. We return to this in Section 5.4.

4.4 Limiting model for the joint tail

Let A_{st} denote a set of the form $\{(x', y') : x' > s, y' > t\}$ for some fixed $s > 0$ and $t > 0$. Following the notation of Section 2 and writing $b_n A_{st}$ for $\{(x', y') : x' > b_n s, y' > b_n t\}$, then on $x > s$ and $y > t$ we have that

$$\begin{aligned}
\lim_{n\to\infty} \Pr\{(X,Y) \in b_n A_{xy} | (X,Y) \in b_n A_{st}\} &= \lim_{n\to\infty} \frac{\Pr(X > b_n x, Y > b_n y)}{\Pr(X > b_n s, Y > b_n t)} \\
&= \frac{\Lambda_\eta \{(x,\infty)\times(y,\infty)\}}{\Lambda_\eta \{(s,\infty)\times(t,\infty)\}}. \qquad (4.4)
\end{aligned}$$

The normalisation condition (2.8) and the conditions in Appendix A ensure the existence of a bivariate random variable (X, Y) with unit Fréchet distributed margins such that equations (2.7) and (4.4) hold.

Substituting the η-asymmetric logistic intensity function given in equation (4.2) into equation (4.4) we obtain the limiting joint tail model

$$\lim_{n\to\infty} \Pr\{(X,Y) \in b_n A_{xy} | (X,Y) \in b_n A_{st}\} = \\
N_{st\varrho}^{-1} \left[(\varrho x)^{-1/\eta} + \left(\frac{y}{\varrho}\right)^{-1/\eta} - \left\{ (\varrho x)^{-1/\alpha} + \left(\frac{y}{\varrho}\right)^{-1/\alpha} \right\}^{\alpha/\eta} \right] \qquad (4.5)$$

for $x > s$ and $y > t$ where $N_{st\varrho} = (\varrho s)^{-1/\eta} + (t/\varrho)^{-1/\eta} - \left\{(\varrho s)^{-1/\alpha} + (t/\varrho)^{-1/\alpha}\right\}^{\alpha/\eta}$. This is the basis of the modelling approach and likelihood methods adopted by Ramos and Ledford (2009) who chose the $s = t = 1$ special case of this more general result. Clearly, since the $\alpha \geq \eta$ case gives rise to a non-degenerate model in equation (4.5) and a degenerate limiting distribution in equation (4.3), the joint tail modelling approach of this section has more generality and wider applicability than an approach based on the componentwise maxima results.

5 Simulation

Methods for simulating from both the limiting point process and the limiting distribution of componentwise maxima are examined here. Most discussion focuses on the η-asymmetric logistic model introduced in Section 4.

5.1 The η-asymmetric logistic limiting point process

The simulation scheme here is similar to that in Ramos and Ledford (2009). Differentiating equation (4.5) yields the joint density

$$\frac{\eta - \alpha}{\alpha \eta^2 N_{st\varrho}} \left\{ (\varrho x)^{-1/\alpha} + \left(\frac{y}{\varrho}\right)^{-1/\alpha} \right\}^{\alpha/\eta - 2} (xy)^{-(1+1/\alpha)} \qquad (5.1)$$

for $x > s$, $y > t$, $N_{st\varrho}$ as defined in Section 4.4 and fixed $s > 0$ and $t > 0$. Similarly to Shi *et al.* (1992), we construct (X, Y) pairs with joint density (5.1) by transforming a pair of easily simulated variables, V and Z say. This is achieved via $X^{-1} = \varrho Z^{\eta} \cos^{2\alpha} V$ and $\varrho Y^{-1} = Z^{\eta} \sin^{2\alpha} V$ where V and Z satisfy

$$V = \begin{cases} \arccos\left[\left\{1 - N_{st\varrho}(\varrho s)^{1/\eta} U_1\right\}^{\frac{\eta}{2(\eta-\alpha)}}\right] & \text{if } 0 < U_1 \le u^* \\ \arcsin\left\{\left(\left(\frac{t}{\varrho}\right)^{1/\eta}\left[\begin{array}{l} N_{st\varrho}U_1 - (\varrho s)^{-1/\eta} + \\ \qquad \left\{(\varrho s)^{-1/\alpha} + \left(\frac{t}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta} \end{array}\right]\right)^{\frac{\eta}{2(\eta-\alpha)}}\right\} & \text{if } u^* < U_1 < 1 \end{cases} \tag{5.2}$$

and

$$Z = \begin{cases} U_2(\varrho s)^{-1/\eta} \cos^{-2\alpha/\eta} V & \text{if } 0 < U_1 \le u^* \\ U_2 \left(\frac{t}{\varrho}\right)^{-1/\eta} \sin^{-2\alpha/\eta} V & \text{if } u^* < U_1 < 1 \end{cases} \tag{5.3}$$

where U_1 and U_2 are independent uniformly distributed on $[0, 1]$ and

$$u^* = N_{st\varrho}^{-1}\left[(\varrho s)^{-1/\eta} - (\varrho s)^{-1/\alpha}\left\{(\varrho s)^{-1/\alpha} + \left(\frac{t}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta - 1}\right].$$

A derivation is given in Appendix B. The above construction generalises the Ramos and Ledford (2009) simulation method which deals with the $s = t = 1$ case only. Illustrative data sets generated using this scheme together with contours of the underlying joint density are given in Figure 1. Note that $\alpha > 1$ in equation (5.1) is a valid density function however this is not a density that can arise for (X_i, Y_i) with unit Fréchet marginals in definition (2.4) since the constraints in Appendix A are then violated.

5.2 More general limiting point process simulation

The approach described here is similar to that of Nadarajah (1999) and enables simulation of the limiting point process $\mathcal{P}^*$ over a region bounded away from the axes and the origin. From equation (2.6), the conditional density of the pseudo-polar transforms (R, W) of (X, Y) in region $A = \{(r, w) : r > r_0, w \in (k, 1-k)\}$ where $r_0 > 0$ and $0 < k < 1$ satisfies

$$f\{r, w | (R, W) \in A\} = \frac{r^{-(1+1/\eta)}\, \mathrm{d}H_\eta(w)}{\int_k^{1-k}\int_{r_0}^{\infty} r^{-(1+1/\eta)}\, \mathrm{d}r \mathrm{d}H_\eta(w)} = \frac{r_0^{1/\eta} r^{-(1+1/\eta)}}{\eta} \frac{\mathrm{d}H_\eta(w)}{\int_k^{1-k} \mathrm{d}H_\eta(w)} \tag{5.4}$$

for $r > r_0$ and $w \in (k, 1-k)$. The factorisation in equation (5.4) means that the R- and W-components can be simulated independently. Since the R-component has distribution function $1 - (r_0/r)^{1/\eta}$ it can be generated via $R = r_0 U^{-\eta}$ where U is uniformly distributed on $[0, 1]$. Simulation of the W-component may be undertaken by inverting its distribution function $\int_k^w \mathrm{d}H_\eta(w') / \int_k^{1-k} \mathrm{d}H_\eta(w')$, or alternatively via the rejection method (see Ripley, 1987, Section 3.2) when H_η has bounded measure density h_η over $(0, 1)$. In this latter case the W-component has density $h_\eta^*(w) = h_\eta(w) / \int_k^{1-k} h_\eta(w')\, \mathrm{d}w'$ for $w \in (k, 1-k)$, so to apply the rejection method we require a density $g(w)$ from which it is straightforward to simulate and a constant M such that $h_\eta^*(w)/g(w) \le M$ for $w \in (k, 1-k)$. In practice it is often convenient to take $g(w)$ as uniform on $(k, 1-k)$ and $M = \sup\left\{(1-2k)h_\eta^*(w)\right\}$, see Nadarajah (1999). Simulation of (X, Y) over a region A_0 bounded away from the axes and the origin but not of the form of A may be undertaken by first applying the procedures above for a region A with $A_0 \subset A$ and then discarding those points falling outside A_0.

5.3 The η-asymmetric logistic limiting componentwise maxima distribution

Our approach here is similar to the methods given by Shi *et al.* (1992) and Stephenson (2003) for generating from the logistic and asymmetric logistic multivariate extreme value distributions. Let (X, Y) denote a random variable with joint distribution function G_η as in equation (4.3) in the non-degenerate case where $\alpha < \eta$. Then we may write $X^{-1} = \varrho\,(N_\varrho Z)^\eta \cos^{2\alpha} V$ and $\varrho Y^{-1} = (N_\varrho Z)^\eta \sin^{2\alpha} V$ where Z and V are independent with easily characterised distributions: V may be represented as $\arcsin(U^{1/2})$ where U is uniform on $[0, 1]$, while Z is a $\left(1 - \frac{\alpha}{\eta}\right) : \frac{\alpha}{\eta}$ mixture of a unit exponential random variable and the sum of two independent unit exponential random variables. See the derivation in Appendix C. Figure 2 shows examples of points generated using this approach with the corresponding densities superimposed.

5.4 Simulation based examination of the $\alpha \ge \eta$ case

In Section 4.3 we saw that G_η for the η-asymmetric logistic model is degenerate for the case $\alpha \ge \eta$, i.e. when the parametric measure H_η is infinite, whereas the point process $\mathcal{P}^*$ is non-degenerate for all $\eta \in (0, 1]$ and $\alpha \in (0, 1]$. This suggests that when $\alpha \ge \eta$ the componentwise maxima of those points in region $R_{\varepsilon b_n}$ grow faster than b_n and therefore occur in separate observations. This effect can be seen in the simulation results in Figure 1, since as α increases the points tend to concentrate towards the boundaries and furthermore tend to be larger in magnitude than the points towards the $w = 1/2$ region of the domain. This behaviour is also consistent with the results in Section 4.4 since by writing $x = s$ or $y = t$ in equation (4.5) it is straightforward to show that the marginal variables there have shape parameter α when $\alpha \ge \eta$, corresponding to marginal tails heavier than the joint tail which has shape parameter η.

6 Conclusions and extension to the multivariate case

We have derived non-degenerate point process results for bivariate tails that hold under the mild regularity condition of bivariate regular variation and regardless of whether the underlying data generation mechanism exhibits asymptotic dependence or asymptotic independence. This is a considerable extension of the available point process theory of de Haan (1985) which is degenerate for asymptotically independent cases. We have exploited our point process framework to derive non-degenerate limit results for componentwise maxima that again hold for both asymptotically dependent and asymptotically independent cases. This provides a clear extension of the currently available models for componentwise maxima. The normalisation condition (2.8) also extends existing hidden regular variation results.

The point process approach developed here has important consequences for statistical data analysis and provides an alternative theoretical underpinning of the modelling framework detailed by Ramos and Ledford (2009). In contrast to existing classical approaches, which tend to be based on weak convergence results concerning the limiting behaviour of normalised componentwise maxima and are non-degenerate only for asymptotically dependent cases, our methods focus directly on the tail behaviour of the joint survivor function of the underlying observations and capture within the same apparatus both the asymptotically dependent and asymptotically independent cases.

We have illustrated our results using the flexible η-asymmetric logistic model of Ramos and Ledford (2009) and have provided efficient methods for simulating from both the limiting point process and the limiting componentwise maxima distribution associated with this model. Likelihood based estimation is straightforward using this parameterisation and is discussed in Ramos and Ledford (2009) where tests of asymptotic dependence versus asymptotic independence ($H_0 : \eta = 1$ versus $H_1 : \eta \leq 1$), tests of symmetry versus asymmetry ($H_0 : \varrho = 1$) and tests of ray-independence ($H_0 : \alpha = 2\eta$) are all examined. Deriving further parametric models and semi-parametric methods that satisfy both the normalisation condition (2.8) and the conditions in Appendix A remains an important open topic for further research.

6.1 Extension to the multivariate case

It is relatively straightforward to extend the previous results to deal with higher dimensions. Let $\boldsymbol{X}$ denote a d-dimensional random variable with unit Fréchet marginal distributions and joint survivor function that is multivariate regularly varying in d-dimensions with index $-1/\eta$, so that for all $\boldsymbol{x} = (x_1, \ldots, x_d)$ with each $x_i > 0$ we have

$$\lim_{u\to\infty} \frac{\Pr(\boldsymbol{X} > u\boldsymbol{x})}{\Pr(\boldsymbol{X} > u\mathbf{1})} = g(\boldsymbol{x}) \left(\prod_{i=1}^{d} x_i\right)^{-1/(d\eta)} \quad \text{where } g(c\boldsymbol{x}) = g(\boldsymbol{x}) \text{ for all } c > 0.$$

Then for b_n satisfying $n \Pr(\boldsymbol{X} > b_n\mathbf{1}) = 1$, we have that the point process

$$\mathcal{P}_n^{\otimes} = \left\{ \left(\frac{X_{1j}}{b_n}, \ldots, \frac{X_{dj}}{b_n} \right) : j = 1, \ldots, n \right\}$$

converges weakly in the limit as $n \to \infty$ to a non-homogeneous Poisson process $\mathcal{P}^{\otimes}$ on $(0, \infty)^d$ with intensity function

$$\Lambda_\eta^{\otimes}\{(x_1, \infty) \times \cdots \times (x_d, \infty)\} = \eta \int_{S_d} \min_{1\leq i\leq d} \left(\frac{w_i}{x_i}\right)^{1/\eta} \mathrm{d}H_\eta(\boldsymbol{w}) \tag{6.1}$$

where $w_i = x_i / \sum_{j=1}^{d} x_j$ for $i = 1, \ldots, d$ and $H_\eta(\boldsymbol{w})$ is a non-negative measure on the open $(d-1)$-dimensional unit simplex

$$S_d = \left\{ \boldsymbol{w} \in \mathbb{R}^d : \sum w_i = 1, \text{ each } w_i > 0 \right\}$$

that is arbitrary apart from having to satisfy the normalisation condition

$$\eta^{-1} = \int_{S_d} \min_{1\le i\le d} w_i^{1/\eta}\, \mathrm{d}H_\eta(\boldsymbol{w}). \tag{6.2}$$

Defining in the obvious notation $\boldsymbol{R}_{\varepsilon b_n} = \{\boldsymbol{x} : \boldsymbol{x} > \varepsilon b_n \mathbf{1}\}$ and denoting by $\boldsymbol{M}_{n,\varepsilon b_n}$ the d-vector of componentwise maxima of those points which occur within $\boldsymbol{R}_{\varepsilon b_n}$, we have

$$\begin{aligned} G_\eta(\boldsymbol{x}) \equiv \lim_{\varepsilon\to 0}\lim_{n\to\infty} \Pr\left(\frac{\boldsymbol{M}_{n,\varepsilon b_n}}{b_n} \le \boldsymbol{x}\right) &= \exp\left[-\eta\int_{S_d}\left\{\max_{1\le j\le d}\left(\frac{w_j}{x_j}\right)\right\}^{1/\eta}\mathrm{d}H_\eta(\boldsymbol{w})\right] \\ &= \exp\{-V_\eta(\boldsymbol{x})\} \end{aligned} \tag{6.3}$$

which is non-degenerate provided that H_η is finite.

The extension of result (3.7) to the multivariate case can be obtained from equation (6.1), the inclusion-exclusion principle and the definition of V_η in equation (6.3), giving

$$\begin{aligned} &\Lambda_\eta^{\otimes}\{(x_1,\infty)\times\cdots\times(x_d,\infty)\} = \\ &= \sum_{i=1}^{d}\eta\int_{S_d}\left(\frac{w_i}{x_i}\right)^{1/\eta}\mathrm{d}H_\eta(\boldsymbol{w}) - \sum_{i=1}^{d-1}\sum_{j=i+1}^{d}\eta\int_{S_d}\left\{\max\left(\frac{w_i}{x_i},\frac{w_j}{x_j}\right)\right\}^{1/\eta}\mathrm{d}H_\eta(\boldsymbol{w}) + \\ &\qquad + \cdots + (-1)^{d+1}\eta\int_{S_d}\max_{1\le i\le d}\left(\frac{w_i}{x_i}\right)^{1/\eta}\mathrm{d}H_\eta(\boldsymbol{w}) \\ &= \sum_{i=1}^{d} V_\eta(\infty,\ldots,\infty,x_i,\infty,\ldots,\infty) - \sum_{i=1}^{d-1}\sum_{j=i+1}^{d} V_\eta(\infty,\ldots,\infty,x_i,\infty,\ldots,\infty,x_j,\infty,\ldots,\infty) + \\ &\qquad + \cdots + (-1)^{d+1}V_\eta(\boldsymbol{x}). \end{aligned} \tag{6.4}$$

So taking the mixed derivative of equation (6.4) with respect to x_1, …, x_d, we obtain

$$\frac{\partial^d V_\eta}{\partial x_1\cdots\partial x_d} = -r^{-(d+1/\eta)} h_\eta(\boldsymbol{w})$$

where $r = \sum_{j=1}^{d} x_j$ and $w_i = x_i/r$ for $i = 1,\ldots,d$.

Example: To illustrate the above, we provide a multivariate version of the η-asymmetric logistic model of Section 4. Consider the d-dimensional measure density

$$h_\eta(\boldsymbol{w}) = \frac{\prod_{i=1}^{d-1}(i\eta-\alpha)}{\eta^d\alpha^{d-1}N_\varrho}\left\{\sum_{i=1}^{d}\left(\frac{w_i}{\varrho_i}\right)^{-1/\alpha}\right\}^{\alpha/\eta-d} \times \left(\prod_{i=1}^{d} w_i\right)^{-1/\alpha-1}$$

on $\boldsymbol{w}\in S_d$ where $\varrho_i > 0$, $\prod_{i=1}^{d}\varrho_i = 1$, $\eta,\alpha\in(0,1]$ and $N_\varrho = \sum_{b\in B}(-1)^{|b|+1}\left(\sum_{i\in b}\varrho_i^{1/\alpha}\right)^{\alpha/\eta}$ where B represents the set of all non-empty subsets of $\{1,\ldots,d\}$ and $|b|$ represents the number of elements in the set b. Then $h_\eta(\boldsymbol{w})$ as above satisfies the normalisation condition (6.2). For example, when $d = 3$ we have that $h_\eta(w_1, w_2)$ is given by

$$h_\eta(w_1, w_2) = \frac{(\eta-\alpha)(2\eta-\alpha)}{\eta^3\alpha^2 N_\varrho}\left\{(\varrho_1\varrho_2 w_1)^{-1/\alpha} + \left(\frac{w_2}{\varrho_1}\right)^{-1/\alpha} + \left(\frac{1-w_1-w_2}{\varrho_2}\right)^{-1/\alpha}\right\}^{\alpha/\eta-3} \times \{w_1 w_2 (1-w_1-w_2)\}^{-1/\alpha-1}$$

where $\eta, \alpha \in (0,1]$, $\varrho_1 > 0$, $\varrho_2 > 0$, and

$$N_\varrho = (\varrho_1\varrho_2)^{-1/\eta} + \varrho_1^{1/\eta} + \varrho_2^{1/\eta} - \left\{(\varrho_1\varrho_2)^{-1/\alpha} + \varrho_1^{1/\alpha}\right\}^{\alpha/\eta} - \left\{(\varrho_1\varrho_2)^{-1/\alpha} + \varrho_2^{1/\alpha}\right\}^{\alpha/\eta} - \left\{\varrho_1^{1/\alpha} + \varrho_2^{1/\alpha}\right\}^{\alpha/\eta} + \left\{(\varrho_1\varrho_2)^{-1/\alpha} + \varrho_1^{1/\alpha} + \varrho_2^{1/\alpha}\right\}^{\alpha/\eta}.$$

A Sufficient conditions for an (η, H_η)-pair satisfying (2.8) to define a valid limit intensity for $\mathcal{P}_n^*$ in (2.4)

These conditions require that for a large u there exists a small $\lambda = \Pr(X > us, Y > ut) \in (0,1)$ for some fixed $s > 0$ and $t > 0$, such that

$$\begin{array}{l} \exp(-1/x) + \lambda\Lambda_\eta\{(x/u, \infty) \times (1, \infty)\}/\Lambda_\eta\{(s,\infty)\times(t,\infty)\} \text{ and} \\ \exp(-1/x) + \lambda\Lambda_\eta\{(1, \infty) \times (x/u, \infty)\}/\Lambda_\eta\{(s,\infty)\times(t,\infty)\} \end{array}$$

are both monotonic increasing in $x > us$ and $x > ut$, respectively. These are similar to the conditions in Theorem 1 of Ramos and Ledford (2009) where the $s = t = 1$ special case is considered. Under the above conditions, and providing (2.8) holds, then it is straightforward to show by a minor extension of the proof of Theorem 1 of Ramos and Ledford (2009) that there exists a bivariate random variable (X, Y) with unit Fréchet margins that satisfies equation (4.4). Simulating independent copies of (X, Y), we construct the sequence of point processes $\mathcal{P}_n^*$ as defined by equation (2.4). Kallenberg's theorem (Resnick (1987), Proposition 3.22) can then be used to show that $\mathcal{P}_n^*$ converges weakly (i.e. in distribution) to a non-homogeneous Poisson process, $\mathcal{P}^*$ say. By construction $\mathcal{P}^*$ has the required intensity on $(0,\infty)\times(0,\infty)$.

B The simulation scheme in Section 5.1

Changing variables in the conditional density function of (X, Y) for $x > s$ and $y > t$ in equation (5.1) according to $X^{-1} = \varrho Z^\eta \cos^{2\alpha} V$ and $\varrho Y^{-1} = Z^\eta \sin^{2\alpha} V$, the density of (Z, V) is given by

$$f_{ZV}(z, v) = \frac{\eta - \alpha}{\eta N_{st\varrho}} 2 \sin v \cos v$$

with domain

$$\left\{ \begin{array}{lll} 0 < z \le (\varrho s)^{-1/\eta} \cos^{-2\alpha/\eta} v & \text{if} & 0 < v \le v^* \text{ and} \\ 0 < z \le (t/\varrho)^{-1/\eta} \sin^{-2\alpha/\eta} v & \text{if} & v^* < v < \pi/2 \end{array} \right.$$

where $v^* = \arctan \varrho^{1/\alpha}$. Noting that V has distribution function

$$F_V(v) = \left\{ \begin{array}{ll} N_{st\varrho}^{-1} (\varrho s)^{-1/\eta} \left(1 - \cos^{2(1-\alpha/\eta)} v\right) & \text{if } 0 < v \le v^* \\ N_{st\varrho}^{-1} \left[(\varrho s)^{-1/\eta} - \left\{ (\varrho s)^{-1/\alpha} + \left(\frac{t}{\varrho}\right)^{-1/\alpha} \right\}^{\alpha/\eta} + \left(\frac{t}{\varrho}\right)^{-1/\eta} \sin^{2(1-\alpha/\eta)} v \right] & \text{if } v^* < v < \pi/2 \end{array} \right.$$

and that the conditional variable $Z|V = v$ is uniformly distributed with density function

$$f_{Z|V=v}(z) = \left\{ \begin{array}{ll} (\varrho s)^{1/\eta} \cos^{2\alpha/\eta} v & \text{if } 0 < v \le v^* \\ \left(\frac{t}{\varrho}\right)^{1/\eta} \sin^{2\alpha/\eta} v & \text{if } v^* < v < \pi/2, \end{array} \right.$$

it follows that Z and V can be represented as in equations (5.2) and (5.3).

C The simulation scheme in Section 5.3

Let f_η denote the density function associated with G_η in equation (4.3). Then

$$f_\eta(x,y) = N_\varrho^{-1}\eta^{-2}\exp\left[-N_\varrho^{-1}\left\{(\varrho x)^{-1/\alpha}+\left(\frac{y}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}\right](xy)^{-(1+1/\alpha)}$$
$$\times\left\{(\varrho x)^{-1/\alpha}+\left(\frac{y}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta-2}\left[N_\varrho^{-1}\left\{(\varrho x)^{-1/\alpha}+\left(\frac{y}{\varrho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}+\frac{\eta-\alpha}{\alpha}\right].$$

The transformation $X^{-1}=\varrho\,(N_\varrho Z)^\eta\cos^{2\alpha}V$ and $\varrho Y^{-1}=(N_\varrho Z)^\eta\sin^{2\alpha}V$ is such that the bivariate random variable (Z,V) has density

$$f_{ZV}(z,v)=\exp(-z)\left\{\frac{\alpha}{\eta}z+\left(1-\frac{\alpha}{\eta}\right)\right\}\sin 2v$$

for $z>0$ and $0<v<\pi/2$. From this representation the simulation approach of Section 5.3 follows immediately.

D The role of open and closed integration intervals

The point process derivation in Section 2 relies on the bivariate regular variation of $\overline{F}(x,y)$. Since by equation (2.1) this is a property that requires both $x>0$ and $y>0$, it is clear that our previous results cannot be extended to a domain that includes the axes unless additional assumptions are made. Equivalently, when our results are expressed in terms of the (r,w)-pseudo-polar coordinates, then it is the properties of the w-component over the open interval $w\in(0,1)$ rather than the closed interval $w\in[0,1]$ that are captured by our analysis. Accordingly, the w-integrations up until now have all been over open intervals. In this section we discuss one possible reason why making the leap from open- to closed-intervals for the w-integrations might be useful, however at this stage we do not have any well reasoned theoretical justification for this leap.

We suppose that $H_\eta(w)$ is as before but additionally has finite masses at the end-points of the w-interval, that is at $w=0$ and $w=1$. Due to the zeros of the integrand at $w=0$ and $w=1$, there is no impact on the normalisation condition (2.8) when the domain of integration is replaced by the closed interval $[0,1]$. This is not the case for the componentwise maxima results of Section 4.3, where replacing the integration region in equation (3.5) with a closed interval enables a direct connection to existing classical bivariate extremes results to be made. To illustrate this, we will examine the relationship between the classical asymmetric logistic bivariate extreme value (BEV) distribution (Tawn, 1988) and this extended version of our componentwise maxima representation. We define H_η as follows: for $w\in(0,1)$ we assume the alternatively parameterised version of the η-asymmetric logistic density measure of Section 4.1 given by

$$h_\eta(w)=\frac{\eta-\alpha}{\alpha\eta^2\theta\phi N_{\theta\phi}}\left\{\left(\frac{w}{\theta}\right)^{-1/\alpha}+\left(\frac{1-w}{\phi}\right)^{-1/\alpha}\right\}^{\alpha/\eta-2}\left\{\frac{w}{\theta}\left(\frac{1-w}{\phi}\right)\right\}^{-(1+1/\alpha)}$$

where $N_{\theta\phi} = \theta^{1/\eta} + \phi^{1/\eta} - (\theta^{1/\alpha} + \phi^{1/\alpha})^{\alpha/\eta}$, and for $w = 0$ and $w = 1$ respectively we add masses of $(1 - \phi^{1/\eta})/(\eta N_{\theta\phi})$ and $(1 - \theta^{1/\eta})/(\eta N_{\theta\phi})$. Note that H_η as defined here satisfies the normalising condition (2.8) regardless of whether the regions of integration are open or closed at $w = 0$ and $w = 1$.

Performing the integration in equation (3.5) over the closed interval $[0, 1]$ gives

$$G_\eta(x, y) = \begin{cases} \exp\left\{-N_{\theta\phi}^{-1}\left[\left\{\left(\frac{x}{\theta}\right)^{-1/\alpha} + \left(\frac{y}{\phi}\right)^{-1/\alpha}\right\}^{\alpha/\eta} + \frac{1-\theta^{1/\eta}}{x^{1/\eta}} + \frac{1-\phi^{1/\eta}}{y^{1/\eta}}\right]\right\} & \text{for } \alpha < \eta, \\ 0 & \text{for } \alpha \geq \eta. \end{cases} \tag{D.1}$$

Setting $\eta = 1$ here yields a joint distribution with the classical asymmetric logistic BEV dependence structure and common scaled unit Fréchet marginal distributions where, for example, $G_\eta(x, \infty) = \exp(-N_{\theta\phi}^{-1} x^{-1})$. Thus under this extended version of our previous results the classical asymmetric logistic BEV dependence structure arises though the $\eta = 1$ special case of the measure H_η given above. Simulation from G_η in equation (D.1) in the non-degenerate case for general η is straightforward. Taking $Z_1 = N_{\theta\phi} X^{1/\eta}$ and $Z_2 = N_{\theta\phi} Y^{1/\eta}$, it follows that Z_1 and Z_2 are unit Fréchet random variables with classical asymmetric logistic BEV dependence structure with dependence parameters $\alpha^* = \alpha/\eta \in (0, 1]$, $\theta^* = \theta^{1/\eta} \in [0, 1]$ and $\phi^* = \phi^{1/\eta} \in [0, 1]$. Points from this latter distribution can be generated using the methods given by Stephenson (2003).

References

[1] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley, Chichester.

[2] Capéraà, P., Fougères, A. -L., and C. Genest (2000). Bivariate distributions with given extreme value attractor. *J. Mult. Anal*, **72**, 30-49.

[3] Coles, S. G. and Tawn, J. A. (1991). Modelling extreme multivariate events. *J. R. Statist. Soc., B*, **53**, 377-392.

[4] Coles, S. G. and Tawn, J. A. (1994). Statistical methods for multivariate extremes: an application to structural design (with discussion). *Appl. Statist.*, **43**, 1-48.

[5] Coles, S. G., Hefernan, J. E. and Tawn, J. A. (1999). Dependence Measures for Extreme Value Analyses. *Extremes*, **2**, 339-365.

[6] Coles, S. G. and Pauli, F. (2002). Models and inference for uncertainty in extremal dependence. *Biometrika*, **89** (1), 183-196.

[7] de Haan, L. and Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrsch. Theor.*, **40**, 317-337.

[8] de Haan, L. (1985). Extremes in higher dimensions: the model and some statistics. In *Proc. 45th Sess. Int. Statist. Inst.*, paper 26.3. The Hague: International Statistical Institute.

[9] de Haan, L. and de Ronde, J. (1998). Sea and wind: multivariate extremes at work. *Extremes*, **1**(1), 7-45.

[10] Heffernan, J. E. (2000). A directory of coefficients of tail dependence. *Extremes*, **3**(3), 279-290.

[11] Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *J. R. Statist. Soc.*, B, **66**(3), 497-546.

[12] Heffernan, J. E. and Resnick, S. I. (2005). Hidden regular variation and the rank transform. *Adv. Appl. Probab.* **37** (2), 393-414.

[13] Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions.* Imperial College Press.

[14] Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, **83**, 169-187.

[15] Ledford, A. W. and Tawn, J. A. (1997). Modelling dependence within joint tail regions. *J. R. Statist. Soc.*, B, **59**, 475-499.

[16] Maulik, K. and Resnick, S. I. (2004). Characterizations and examples of hidden regular variation. *Extremes*, **7**(1), 31-67.

[17] Nadarajah, S. (1999). Simulation of multivariate extreme values. *J. Statist. Comput. Simul.*, **62**, 395-410.

[18] Pickands, J. (1981). Multivariate extreme value distribution. *Bull. Int. Statist. Inst.*, **49**, 859-878.

[19] Ramos, A. and Ledford, A. W. (2005). Regular score tests of independence in multivariate extreme values. *Extremes*, **8**(1-2), 5-26.

[20] Ramos, A. and Ledford, A. W. (2009). A new class of models for bivariate joint tails. *J. R. Statist. Soc., B*, 71(1), 219-241.

[21] Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag.

[22] Resnick, S.I. (2002). Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, **5**(4), 303-336.

[23] Ripley, B. D. (1987). *Stochastic Simulation.* New York: John Wiley and Sons.

[24] Shi, D., Smith, R. L. and Coles, S. G. (1992). Joint versus marginal estimation for bivariate extremes. Tech. Rep. 2074, Department of Statistics, University of North Carolina, Chapel Hill.

[25] Sibuya, M. (1960). Bivariate extreme statistics. *Ann. Inst. Statist. Math.*, **11**, 195-210.

[26] Stephenson, A. G. (2003). Simulating multivariate extreme value distributions of logistic type. *Extremes*, **6**(1), 49-59.

[27] Tawn, J. A. (1988). Bivariate extreme value theory: models and estimation, *Biometrika*, **75**, 397-415.

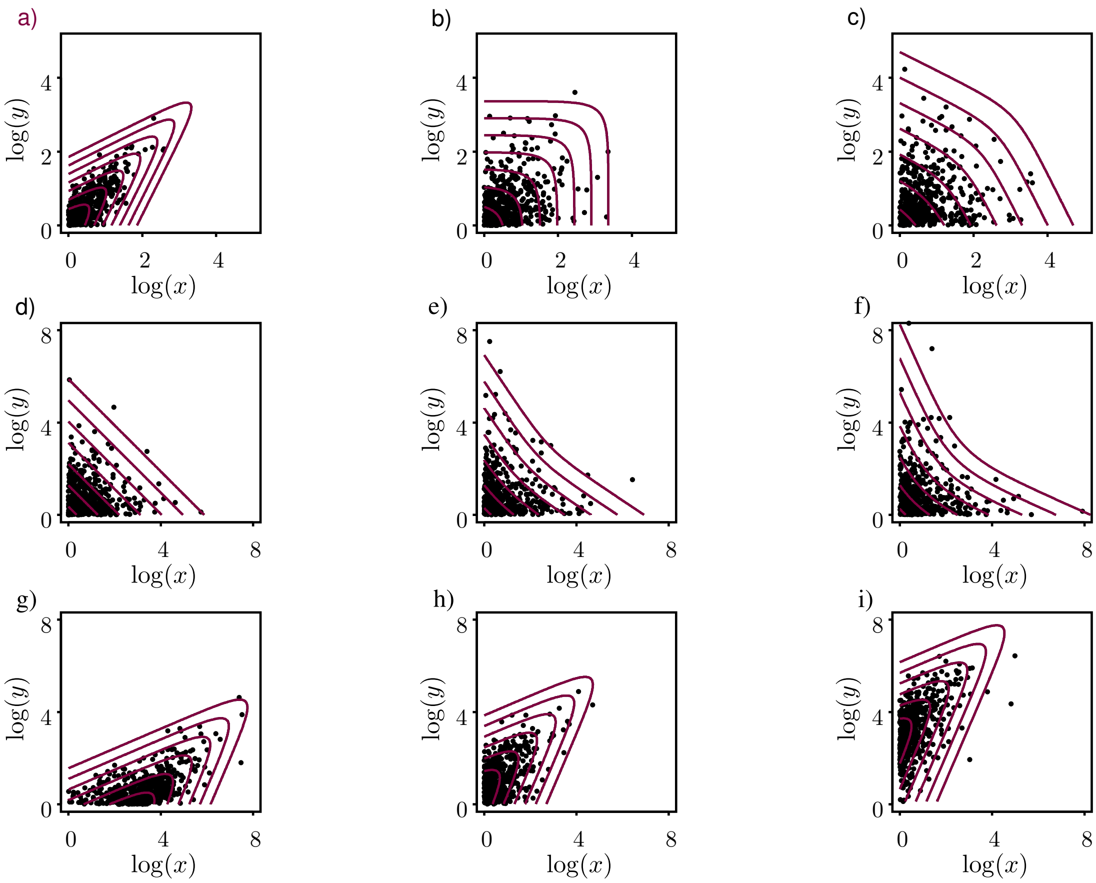

Figure 1: Simulated (X, Y) pairs from the η-asymmetric logistic point process in region $A = \{(x, y) : x > 1, y > 1\}$ obtained using the algorithm of Section 5.1, with density contours superimposed. Panels a)–f) all have $\varrho = 1$ so correspond to symmetric cases and take $\eta = 0.4$ throughout, with $\alpha = \{0.2, 0.4, 0.6, 0.8, 1, 1.3\}$ respectively. Panels g)–i) each have $\alpha = 0.4$ and $\eta = 0.7$, and show the effect of changing asymmetry with $\varrho = \{0.2, 1.5, 5\}$ respectively. Note the axes are not all to the same scale. The $\alpha = 1.3$ panel illustrates an $\alpha > 1$ case of equation (5.1), see Section 5.1.

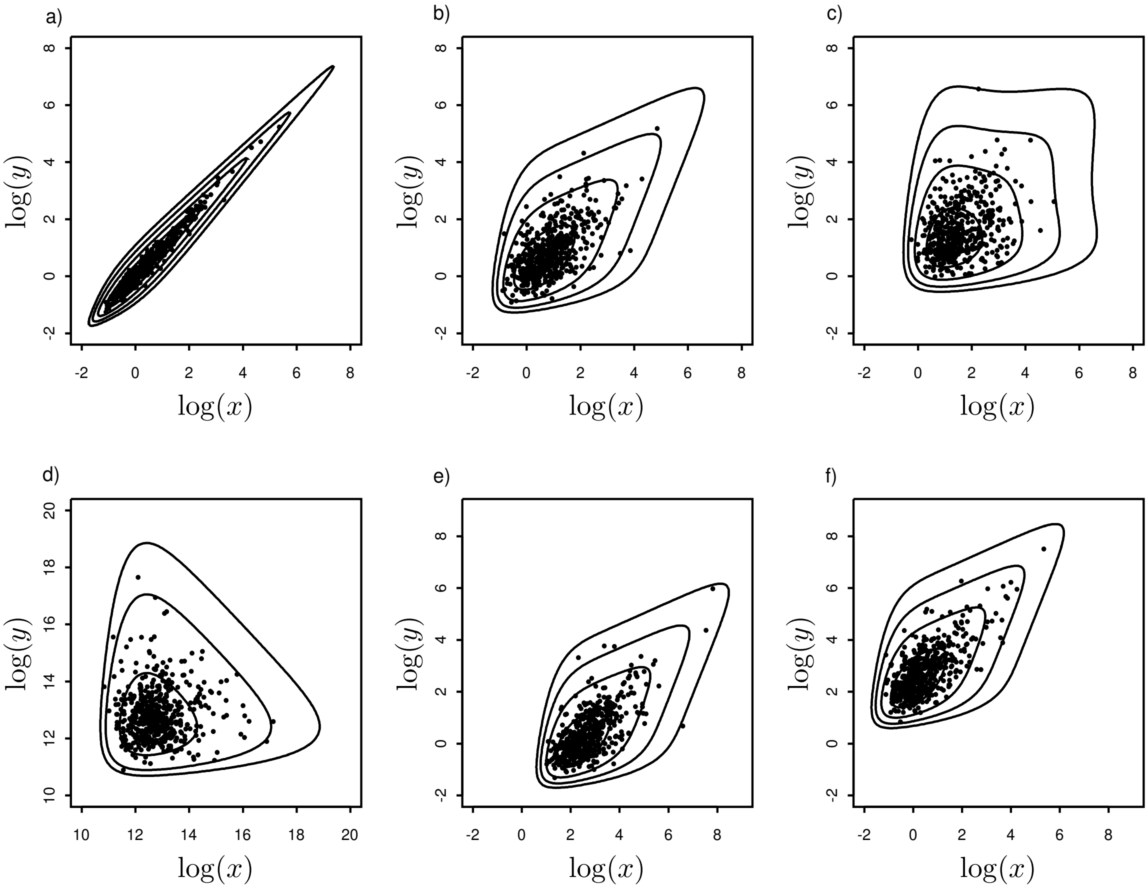

Figure 2: Simulated points from the η-asymmetric logistic componentwise maxima distribution G_η in equation (4.3), with densities superimposed. Panels a)–d) all have $\varrho = 1$ so correspond to symmetric cases and take $\eta = 0.7$ throughout, with $\alpha = \{0.1, 0.4, 0.6, 0.7 - 1e^{-8}\}$ respectively. Panels e)–f) each have $\alpha = 0.4$ and $\eta = 0.7$, and show the effect of changing asymmetry with $\varrho = \{0.3, 3\}$ respectively. Since G_η is degenerate for $\alpha \geq \eta = 0.7$, the biggest value used for α was $\eta - 1e^{-8}$.