Inference for new environmental contours using extreme value analysis

Emma S. Simpson[1] and Jonathan A. Tawn[2]

[1]Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT, U.K.

[2]Department of Mathematics and Statistics, Lancaster University, LA1 4YF, U.K.

January 23, 2024

Abstract

Environmental contours are often used in engineering applications to describe risky combinations of variables according to some definition of an exceedance probability. These contours can be used to both understand multivariate extreme events in environmental processes and mitigate against their effects, e.g., in the design of structures. Such ideas are also useful in other disciplines, with the types of extreme events of interest depending on the context. Despite clear connections with extreme value modelling, much of this methodology has so far not been exploited in the estimation of environmental contours; in this work, we provide a way to unify these areas. We focus on the bivariate case, introducing two new definitions of environmental contours. We develop techniques for their inference which exploit a non-standard radial and angular decomposition of the variables, building on previous work for estimating limit sets. Specifically, we model the upper tails of the radial distribution using a generalised Pareto distribution, with adaptable smoothing of the parameters of this distribution. Our methods work equally well for asymptotically independent and asymptotically dependent variables, so do not require us to distinguish between different joint tail forms. Simulations demonstrate reasonable success of the estimation procedure, and we apply our approach to an air pollution data set, which is of interest in the context of environmental impacts on health.

Keywords: bivariate extremes; environmental contours; generalised additive model; generalised Pareto distribution; structure function

1   Introduction

In environmental, engineering, financial and health contexts, it is the exposure to extreme values of prevailing conditions that cause the most risk. When these prevailing conditions relate to a univariate random variable, the extreme values correspond to observations from either its lower or upper tail, depending on the context, e.g., human health responds differently to cold weather or heatwaves, with such events being linked to respective tails of the temperature variable. When the prevailing conditions are described by a multivariate random variable $\boldsymbol{X} \in \mathbb{R}^d$, for $d \in \mathbb{N}\backslash\{1\}$, there is still a need for the identification of sets of variable combinations that are the 'riskiest', as judged for the given class of problems.

For such a set of risky combinations of $\boldsymbol{X}$ to be useful, a probability needs to be associated with the set, where this probability is small given the rarity of the events of interest. Hence, empirical methods to identify such sets are infeasible, as typically the probability will be smaller than n^{-1}, where n is the sample size. In the reverse set-up, any set in $\mathbb{R}^d$ is not uniquely defined by the probability of its occurrence alone, so additional criteria need to be imposed. These criteria could lead to sets being defined such that they have a small probability of being exceeded, or by a contour of points representing the boundaries of different regions, each with some specified shape and equal occurrence probabilities.

This paper aims to formalise these statements, in the case where $\boldsymbol{X}$ has a joint density function $f_{\boldsymbol{X}}$, to provide an inference framework for extreme sets related to combinations of $\boldsymbol{X}$, which are typically called *environmental contours*; see discussion in the engineering (Mackay and Haselsteiner, 2021) and statistical (Hafver et al., 2022) literature. In engineering, these contours are critical to the practical design of complex, multifaceted, structures that need to withstand environmental multivariate extreme events (Huseby et al., 2015). Similar ideas can be applied in environmental settings, e.g., to specify safety standards relating to combinations of pollutants, for mitigation against a range of health issues, as used as the motivating example in Heffernan and Tawn (2004); or in finance when studying risk across portfolios with different combinations of investments (Poon et al., 2004; Castro-Camilo et al., 2018).

The key complication with identifying such risky combinations is the lack of a natural ordering in multivariate problems. Barnett (1976) addressed this problem by identifying four possible ordering strategies: componentwise maxima, concomitants, convex hulls and structure variables. Subsequent extreme value theory and/or methodology has been developed for these four strategies in Tawn (1990); Ledford and Tawn (1998); Eddy and Gale (1981), Brozius and de Haan (1987) and Davis et al. (1987); and Coles and Tawn (1994), respectively. In the context of the convex hull, the asymptotic shape and number of points contributing to the convex hull are studied, and links are drawn with multivariate extreme value theory, but for environmental contours, interest lies additionally in events beyond the observed data. In contrast, the recent study of a new basis for ordering, linked to the limiting shape of a scaled sample cloud (Balkema et al., 2010; Nolde, 2014; Nolde and Wadsworth, 2022) seems to address these concerns.

For well-defined practical problems, it is the structure variable approach that is the most relevant for identifying sets of risky values of $\boldsymbol{X}$. Coles and Tawn (1994) present a statistical framework to study extremes of structure variables. For a d-dimensional vector $\boldsymbol{x} = (x_1, \ldots, x_d)$, they consider the *structure function* $\Delta(\boldsymbol{x}) : \mathbb{R}^d \to \mathbb{R}$, which is entirely determined by the context of the problem. For example, Δ

could be the stress on an offshore structure with $\boldsymbol{x}$ being wave height and wind speed at that structure. The function Δ captures information about the impact under different values of $\boldsymbol{x}$; as the value of Δ increases, the severity of the impact also increases. Similarly, in a health context Δ could be the effect of air pollutants on human lungs, with $\boldsymbol{x}$ representing levels of the different air pollutants. In these contexts, interest lies in the probability of occurrence of large structure variables, i.e.,

$$\Pr\{\Delta(\boldsymbol{X}) > v\} = \Pr(\boldsymbol{X} \in A_v) = \int_{\boldsymbol{x} \in A_v} f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x}, \tag{1}$$

where $A_v = \{\boldsymbol{x} \in \mathbb{R}^d : \Delta(\boldsymbol{x}) > v\}$ is termed the *failure region*. If v is sufficiently large to be above any observations of the univariate variable $\Delta(\boldsymbol{X})$, Bruun and Tawn (1998) show that there are substantial advantages that can be achieved in estimating $\Pr\{\Delta(\boldsymbol{X}) > v\}$ using a multivariate approach building out of the right hand side of expression (1), over analysing univariate realisations of $\Delta(\boldsymbol{X})$; this is essentially because knowledge of the structure function beyond the observations can be exploited. However, these findings are subject to having a good inference method for multivariate joint tails, which is not always straightforward. The method from Coles and Tawn (1994) only considered asymptotically dependent models for $f_{\boldsymbol{X}}$, whereas Heffernan and Tawn (2004) and Wadsworth and Tawn (2013) extended these models to also cover asymptotic independence, i.e., allowing for the possibility that only some subsets of the components of $\boldsymbol{X}$ can be simultaneously extreme.

The structure variable approach provides an extreme set of risky combinations of $\boldsymbol{X}$ via the set A_v, with an associated joint probability p of occurrence given by (1). For a single structure function Δ, then the boundary of the failure region gives the required environmental contour, $\mathcal{C}_p := \{\boldsymbol{x} \in \mathbb{R}^d : \Delta(\boldsymbol{x}) = v_p\}$, where v_p is the $(1-p)$th quantile of the structure function Δ. For structure functions of interest, $\mathcal{C}_p$ will be a connected subset of $\mathbb{R}^d$ and a continuous function of p. As an illustration, consider the sub-class of this framework when $\Delta(\boldsymbol{x}) = \min_{i=1,\ldots,d}(\phi_i x_i)$ for known constants $\phi_i > 0$ $(i = 1, \ldots, d)$, so interest lies in the joint survivor function on a given ray determined by the ϕ_i constants. Here, the risky set is $A_{v_p} = \{\boldsymbol{x} \in \mathbb{R}^d : x_i > v_p/\phi_i, i = 1, \ldots, d\}$, and $\mathcal{C}_p = \{\boldsymbol{x} \in \mathbb{R}^d : \min_{i=1,\ldots,d}(\phi_i x_i) = v_p\}$ is its lower boundary, so that $\mathcal{C}_p$ is continuous over $\mathbb{R}^d$ and changes smoothly with p.

However, the structure variable formulation is overly simplistic in practice as it assumes that Δ is entirely known. To illustrate this, consider again the structure function $\Delta(\boldsymbol{x}) = \Delta(\boldsymbol{x}; \boldsymbol{\phi}) = \min_{i=1,\ldots,d}(\phi_i x_i)$, where $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_d)$ but now with the $\phi_i > 0$ all unknown. In this case, both A_{v_p} and $\mathcal{C}_p$ vary with $\boldsymbol{\phi}$ as well as p. This means that either additional criteria are needed to optimise $\boldsymbol{\phi}$, or a broader definition of an environmental contour is required than one based on the approach of Coles and Tawn (1994) with known Δ. For this particular class of structure function, when $d = 2$ and for a fixed probability p, Murphy-Barltrop et al. (2023) estimate environmental contours corresponding to all combinations of $(x_1(p), x_2(p))$ with $\Pr\{X_1 > x_1(p), X_2 > x_2(p)\} = p$, i.e., the contour of equal joint survivor functions. The fact that Murphy-Barltrop et al. (2023) need only to find a two-dimensional solution, whereas our set up above suggested it should be three-dimensional over (ϕ_1, ϕ_2, v_p), illustrates that to get identifiability in our solution when using the Δ-approach, we need to impose that $\phi_2 = 1 - \phi_1$ and $\phi_1, \phi_2 \in [0, 1]$, say.

There is a need for techniques to estimate environmental contours in more general settings than Murphy-Barltrop et al. (2023) and which scale better with dimension, as the following examples illustrate.

First, consider the scenario where $\Delta(\boldsymbol{x}) = \Delta(\boldsymbol{x}; \boldsymbol{\phi})$, for unknown $\boldsymbol{\phi}$, occurs in an engineering design phase, e.g., for an oil-rig. It is possible that the general form of Δ is known from previous constructions, but that there are features of the design, determined by $\boldsymbol{\phi}$, which are specific to the location of the oil-rig. In this case, it is not possible to evaluate $\Delta(\boldsymbol{x}; \boldsymbol{\phi})$ for all possible $\boldsymbol{\phi}$, as numerical evaluation of $\Delta(\boldsymbol{x}; \boldsymbol{\phi})$ can be computationally demanding. Similarly, in health risk assessments for air pollutants in setting legal maximum levels, there are a range of structure functions to assess, i.e., $\Delta_i(\cdot; \boldsymbol{\phi}_i)$ for $i \in \mathcal{I}$, where the set $\mathcal{I}$ corresponds to a collection of different health risks as a consequence of the pollutants $\boldsymbol{x}$.

Ross et al. (2020) and Haselsteiner et al. (2021) provide recent reviews on environmental contours from an engineering perspective. The earliest methods in that literature go back to the IFORM (Hasofer and Lind, 1974; Winterstein et al., 1993), which is based on the implicit assumption of a Gaussian copula and a range of approximations. Since then, there have been a series of extensions to the definitions; these reduce the number and impact of various assumptions in different ways. Such extensions include the direct sampling approach of Huseby et al. (2013), the high density contour approach of Haselsteiner et al. (2017) and the ISORM (Chai and Leira, 2018), all of which are discussed in Mackay and Haselsteiner (2021). If an environmental contour cannot be defined according to a known structure function, there are a number of ways these ideas may be used and implemented in practice. Critically, these will depend on the specific application and require knowledge from domain experts, so that statisticians have only a limited role to play at such a stage. To illustrate this, we return again to our engineering and air pollution motivating examples. In engineering contexts, if environmental contours are derived at an initial design stage, then different points on the contour can be used to test a range of possible structures and narrow down the specification to a known structure function Δ; from here, the previously-described structure function approach could be used to finalise a specific design. In contrast, in health contexts, practitioners may be able to identify the combinations of pollutants on the contour of interest that are the most likely to lead to adverse health effects, possibly relating to a range of different symptoms. Contours in this setting provide vital information about the most likely dangerous combinations of air pollutants and so help guide the setting of new environmental air quality standards.

Environmental contours can be seen as an application of multivariate extremes, yet they have not been studied systematically from a methodological perspective. In multivariate settings, statistical extreme value methods have primarily focused on dependence characterisations and inference, rather than following this through to implications for practitioners. For example, for any bivariate pair $(X_i, X_j) \subseteq \boldsymbol{X}$, with $i \neq j$, much focus is given to determining the value of

$$\chi_{ij} = \lim_{u \to 1} \Pr\{F_{X_j}(X_j) > u \mid F_{X_i}(X_i) > u\} \in [0, 1], \tag{2}$$

where F_{X_i} and F_{X_j} are the marginal distribution functions of X_i and X_j, respectively. When $\chi_{ij} > 0$ ($\chi_{ij} = 0$) the variables (X_i, X_j) are said to be asymptotically dependent (asymptotically independent), respectively (Coles et al., 1999). The key difference between these two forms of extremal dependence is that under asymptotic dependence (asymptotic independence) it is possible (impossible) for both variables to take their largest values simultaneously. The asymptotically independent limit can occur for widely used dependence models, such as for all Gaussian copulas with non-perfect dependence, and distinguishing

between these extremal dependence classes can be a non-trivial task in practice.

Some previous progress has been made in deriving contours linked to multivariate extreme value methods. Specifically, Cai et al. (2011) estimate equal joint density contours in the tail of the joint distribution. However, they make restrictive assumptions about the marginal and dependence structures; they assume multivariate regular variation, which imposes identically distributed marginal distributions that must be heavy tailed, and that the variables are either asymptotically dependent or completely independent, meaning more general forms of asymptotic independence are excluded. Einmahl et al. (2013) also focus on equal joint density contours, but select the contour such that the probability of $\boldsymbol{X}$ being outside the contour is of a small target level. Although they allow different (though still heavy tailed) marginal distributions, they also assume asymptotic dependence and conditions such that the spectral measure is non-zero across all interior directional rays, for which no practical test currently exists.

The current methods have the potential to be improved using more advanced statistical techniques that can be applied in both asymptotic dependence and asymptotic independence settings, and crucially, which remove the need to determine in advance whether the variables are asymptotically dependent. Extensions of these approaches should also remove the restriction that the estimated contours must correspond to joint density contours. We aim to address all three of these gaps in current methodology in this paper.

We introduce two new environmental contour definitions and provide procedures for their estimation. We build on the methods of Simpson and Tawn (2022) for estimating the limiting boundary shape of bivariate sample clouds under some appropriate scaling. Following the strategy from copula modelling, where dependence features are considered separately to the marginal distributions, we define our contours on a standardised scale, with the variables first transformed to have common Laplace marginal distributions, and the resulting estimated contours then back-transformed to the original margins.

We end our introduction with an overview of the paper. In Section 2, we introduce two new environmental contour definitions, and explain how the approach of Simpson and Tawn (2022) is adapted to estimate these contours in Section 3. The results of a simulation study are presented in Section 4, demonstrating across a range of examples that our estimation procedure provides contours are close to the truth, as well as being reasonably successful in obtaining regions with the desired proportion (probability) of observations lying outside them for within sample (out of sample) contours, respectively. An application to modelling air pollution, considering differences in behaviour due to season, is presented in Section 5. We conclude with a discussion in Section 6.

2   An outline of new contour definitions

2.1   Overview of the strategy

We make two key decisions when defining our new environmental contours. Firstly, we work on a marginal Laplace scale, and secondly, the contours are defined in terms of radial-angular coordinates. Before providing the contour definitions in Section 2.2, we discuss the motivation for these two choices.

It is common in the literature on both environmental contours and multivariate extremes to work on

standardised margins. For instance, with the IFORM, one transforms to independent, standard Gaussian margins, before estimating a circular contour and then back-transforming to the original marginal scale. In the theoretical work of Nolde and Wadsworth (2022) and the methodological implementation of Simpson and Tawn (2022), standard exponential margins are used, but this choice comes with the drawback of only being able to capture positive dependence or independence. In order to additionally handle negative dependence, as well as lower tails on the marginal scale, we define our contours for variables (X_L, Y_L) having standard Laplace margins (Keef et al., 2013), i.e., with marginal distribution functions

$$F_{X_L}(z) = F_{Y_L}(z) = \begin{cases} \frac{1}{2}\exp(z), & z \leq 0, \\ 1 - \frac{1}{2}\exp(-z), & z > 0, \end{cases} \quad z \in \mathbb{R}. \tag{3}$$

As the Laplace marginal distribution has both upper and lower exponential tails, the probabilistic structural relationship between (X_L, Y_L) is the same as between the associated variables on exponential variables, i.e., (X_E, Y_E), in the joint upper region. However, only with Laplace margins do these relationships hold, in a similar form, when considering joint tail regions involving both upper and lower tail events. So, the benefit of using Laplace over exponential margins is that they automatically cover all joint tails in a unified way, whilst being able to exploit existing theory derived using exponential variables.

In practice, as with the IFORM, we can apply a transformation to ensure the correct margins as a preliminary step and then back-transform to obtain a contour on the original marginal scale if required; this is discussed further in Section 3. We also demonstrate in Section 4.3 that estimation on a standardised Laplace scale has benefits over an equivalent approach applied on the scale of the observed data.

The radial and angular components of (X_L, Y_L) are defined as

$$R_L = \sqrt{X_L^2 + Y_L^2} > 0, \qquad W_L = \tan^{-1}(Y_L/X_L) \in (-\pi, \pi],$$

respectively. Here, large values of R_L can be thought of as corresponding to 'extreme events' in both the Laplace and original space, with the value of R_L defining the level of extremity; we exploit this in our contour definitions and estimation procedures. In particular, both contours are defined in terms of quantiles of $R_L \mid (W_L = w)$ with $w \in (-\pi, \pi]$, and this polar coordinate setting allows us to treat W_L as a covariate when modelling R_L. This will be achieved by exploiting methods from univariate extreme value analysis, with smoothing over W_L realised using the framework of generalised additive models.

Finally, the setting of Laplace margins and polar coordinates results in contours 'centred' on the marginal median values after back-transformation to the original scale. Even so, parallels with the setting considered in Simpson and Tawn (2022) mean that much of their methodology can be applied and extended here for estimation purposes, as discussed in Section 3.

2.2 Contour definitions

We obtain contours where the probability of lying outside these is $p \in (0, 1)$; where in our extreme setting p is taken to be small. The first contour type uses the same quantile level for the radial variable across each angle, while the second allows the quantile level to vary with the density of the angular component.

Contour definition 1 If the angular density $f_{W_L}(w)$ exists, for $w \in (-\pi, \pi]$, define $r_p(w)$ such that

$$\Pr\left\{R_L > r_p(w) \mid W_L = w\right\} = \begin{cases} p, & \text{if } f_{W_L}(w) > 0, \\ 0, & \text{if } f_{W_L}(w) = 0, \end{cases} \tag{4}$$

and the corresponding contour as $\mathcal{C}_p^1 := \{r_p(w) : w \in (-\pi, \pi]\}$. It is straightforward to show that the probability of lying outside the contour $\mathcal{C}_p^1$ is p, since

$$\int_{-\pi}^{\pi} \Pr\left\{R_L > r_p(v) | W_L = v\right\} dF_{W_L}(v) = p \int_{-\pi}^{\pi} f_{W_L}(v) dv = p.$$

When only a subset of the joint tail region is of practical concern, the definition of $\mathcal{C}_p^1$ can be adapted to reflect this. Specifically, in our Laplace marginal setting, interest may lie only with extreme values of (X_L, Y_L) having angles $W_L \in \Omega \subset (-\pi, \pi]$. For example, if only the joint upper tail of both variables is of interest, we could consider $\Omega = [0, \pi/2]$. In this context, we suggest the contour $\mathcal{C}_p^1(\Omega) := \{r_p(w; \Omega) : w \in \Omega\}$, with $r_p(w; \Omega)$ such that $\Pr\{R_L > r_p(w; \Omega) \mid W_L = w\} = p/\Pr(W_L \in \Omega)$, for $w \in \Omega$. As we are interested in Ω, it follows that $\Pr(W_L \in \Omega) = \int_\Omega f_{W_L}(v) dv > 0$, and for $W_L \in \Omega$, the probability of lying outside the contour $\mathcal{C}_p^1(\Omega)$ is p.

Contour definition 2 Assuming that $f_{W_L}(w) > 0$ for all $w \in (-\pi, \pi]$, define $r_p^*(w)$ such that

$$\Pr\left\{R_L > r_p^*(w) \mid W_L = w\right\} = \frac{c_p}{\max\left\{f_{W_L}(w), p/(2\pi)\right\}}, \tag{5}$$

with the constant c_p chosen so that the probability of lying outside the contour $\mathcal{C}_p^2 := \{r_p^*(w) : w \in (-\pi, \pi]\}$ is p. This required value of c_p is the solution to

$$\int_{-\pi}^{\pi} \Pr\left\{R_L > r_p^*(v) | W_L = v\right\} dF_{W_L}(v) = c_p \int_{-\pi}^{\pi} \frac{f_{W_L}(v)}{\max\left\{f_{W_L}(v), p/(2\pi)\right\}} dv = c_p \int_{-\pi}^{\pi} \min\left\{1, \frac{2\pi f_{W_L}(v)}{p}\right\} dv = p.$$

That is,

$$c_p = p \left[\int_{-\pi}^{\pi} \min\left\{1, \frac{2\pi f_{W_L}(v)}{p}\right\} dv\right]^{-1}. \tag{6}$$

The definition of $\mathcal{C}_p^2$ is chosen so that the probability of lying outside the contour is smaller on angles where the density is larger. We cannot take $f_{W_L}(w)$ as the denominator in the right-hand side of (5), since c_p would simplify to $p/(2\pi)$, leading to the probability in (5) being greater than 1 if $f_{W_L}(w) < p/(2\pi)$, for any $w \in (-\pi, \pi]$; see Section A of the Supplementary Material. In the Supplementary Material, we also justify the choice to take the scaling function as $\max\{f_{W_L}(w), p/(2\pi)\}$. We note that as more extreme contours are considered, i.e., as $p \to 0$, scaling via the density only is recovered and $c_p \sim p/(2\pi)$.

2.3 Contour examples

Figure 1 demonstrates our two new contour definitions at varying probability levels for a range of copulas; their upper tail dependence features cover both asymptotic dependence and asymptotic independence, as
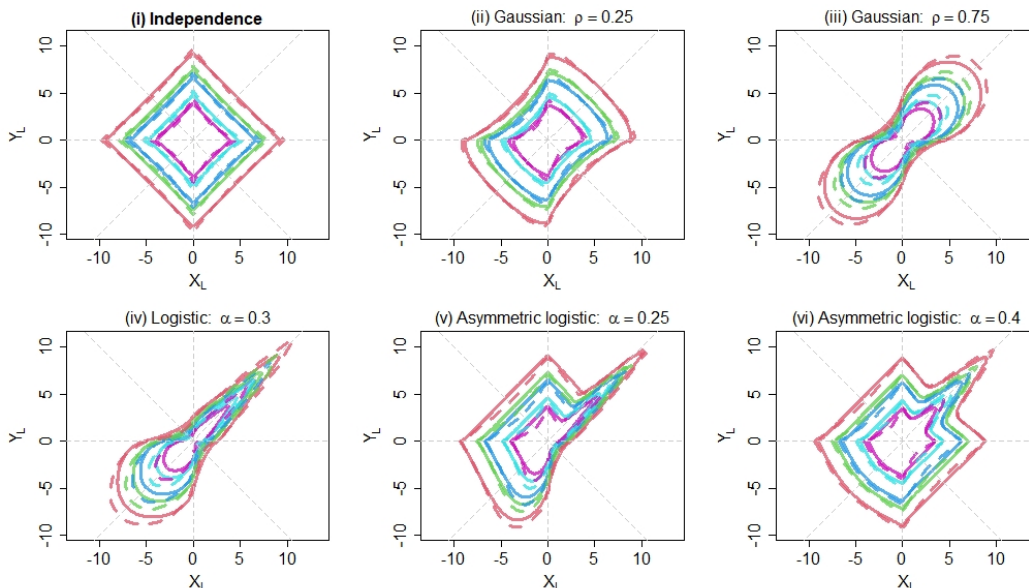
Figure 1: Examples of $\mathcal{C}_p^1$ (solid lines) and $\mathcal{C}_p^2$ (dashed) for the case of independence, two Gaussian copulas, and three bivariate extreme value copulas with logistic or asymmetric logistic models. For $p \in \{0.1, 0.05, 0.01, 0.005, 0.001\}$, the contours lie progressively further from $(0,0)$. The grey dashed lines show $X_L = 0$, $Y_L = 0$, $Y_L = X_L$ and $Y_L = -X_L$.

defined by limit (2). We first present examples for the independence setting (case (i)) and two bivariate Gaussian copulas (cases (ii) and (iii)) with correlation matrices having off-diagonal elements $\rho = 0.25, 0.75$, respectively. We then consider some common bivariate extreme value distributions having copula function

$$C(u,v) = \exp\left\{-V\left(-1/\log u, -1/\log v\right)\right\}, \qquad u, v \in [0,1],$$

with the exponent measure V satisfying certain constraints; see Coles et al. (1999). The joint distribution of the corresponding variables (X_L, Y_L) on standard Laplace scale is given by $F_{X_L, Y_L}(x,y) = C\left\{F_{X_L}(x), F_{Y_L}(y)\right\}$, with $F_{X_L}(x)$, $F_{Y_L}(y)$ defined as in (3). In Figure 1, examples (i) and (iv)-(vi) are all variants of the bivariate asymmetric logistic model of Tawn (1988), having exponent measure

$$V(x,y) = \frac{\theta_1}{x} + \frac{\theta_2}{y} + \left\{\left(\frac{1-\theta_1}{x}\right)^{1/\alpha} + \left(\frac{1-\theta_2}{y}\right)^{1/\alpha}\right\}^{\alpha},$$

with $x, y > 0$, $\alpha \in (0,1]$ and $\theta_1, \theta_2 \in [0,1]$. Setting $\theta_1 = \theta_2 = 1$ gives independence, as in case (i); case (iv) is the logistic model (Gumbel, 1960), where $\theta_1 = \theta_2 = 0$; in case (v), we set $\theta_1 = 0$ and $\theta_2 = 0.5$, leading to a mixture structure where X_L can only be large when Y_L is also large, but Y_L can be large independently of X_L; and in case (vi), we take $\theta_1 = \theta_2 = 0.5$, which has features of both the independence and logistic models. Where the value of α is relevant, i.e., in cases (iv)-(vi), its value controls the strength of dependence in the upper tails and therefore the width of the pointed section along the upper-right diagonal; the closer α is to zero, the narrower this section of the contour. Having $\alpha \in (0,1)$ corresponds to asymptotic dependence, as defined by limit (2), while $\alpha = 1$ recovers complete independence.

The independence and Gaussian examples all exhibit asymptotic independence in both the upper and lower joint tails. In contrast, while cases (iv)-(vi) have asymptotic dependence in their joint upper tails, they exhibit asymptotic independence in their joint lower tails. Another commonly considered copula in extreme value analysis is the inverted logistic model (Ledford and Tawn, 1997); we note that such examples are not necessary here, since working on Laplace margins means that the contours $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ for any inverted logistic copula are equivalent to its logistic counterpart (i.e., where both have the same dependence parameter α) up to a rotation about the origin.

For any copula and p, contours $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ differ due to the angular density $f_{W_L}(w)$. For the copulas in Figure 1, plots of the $f_{W_L}(w)$ values are shown in Figure 2. For independence and the Gaussian copula with $\rho = 0.25$, the range of values taken by $f_{W_L}(w)$ for $w \in (-\pi, \pi]$ is relatively small, meaning that $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ are quite similar. For the logistic model and the Gaussian copula with $\rho = 0.75$, $f_{W_L}(w)$ takes its largest values around $w = \pi/4$ and $w = -3\pi/4$, meaning that $\mathcal{C}_p^2$ appears to be more stretched along the diagonal $Y_L = X_L$ compared to $\mathcal{C}_p^1$; this is also true of the two asymmetric logistic cases, but to a lesser extent. The theory to derive contours such as those presented in Figure 1, for (X_L, Y_L) having some joint density $f_{X_L, Y_L}(x, y)$, is presented in Section B of the Supplementary Material. We will revisit these copula examples in Section 4, where we demonstrate the efficacy of our estimation procedure.

3 Contour estimation

3.1 Overview of the inferential approach and our new contribution

We now introduce our approach to estimating the new contours $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ by adapting the approach of Simpson and Tawn (2022), whose work focuses on limit sets, to allow for inference of sub-asymptotic contours. We begin by outlining the main ways we adapt the existing methodology for our new purpose. We then present our estimation approach for $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ in Sections 3.2 and 3.3, respectively, before discussing tuning parameter selection in Section 3.4 and marginal modelling in Section 3.5.

An important difference between this work and that of Simpson and Tawn (2022) is that the latter considers variables having standard exponential marginal distributions, while we work on standard Laplace scale. As mentioned in Section 2.1, this allows for positive *and* negative associations between the variables, as well as flexibility in modelling both tails of each marginal distribution. In exponential margins, the angular variable is restricted to $[0, \pi/2]$, whereas in Laplace margins $W_L \in (-\pi, \pi]$; since results at these endpoints should coincide, we must now carefully consider the cyclic nature of the angular distribution in our inferential procedure by updating our modelling choices to ensure this feature is preserved.

The definition of $\mathcal{C}_p^2$ in (5) depends on the value of the angular density f_{W_L}; estimation of this function was not previously required, but must be considered in the present work. The cyclic nature of $f_{W_L}(w)$ over $w \in (-\pi, \pi]$ is again a key feature that needs to be considered in our estimation approach. In addition, having the exceedance probabilities in (5) depend on f_{W_L} means that some of the radial quantiles required to estimate the contour may not be 'extreme'; we must therefore introduce a way to estimate non-extreme radial quantiles (as well as extreme ones) into our contour estimation procedure.

A final adaptation comes in the estimation of the marginal distributions. The results in Simpson and Tawn (2022) only required contours estimated on exponential scale, which was achieved via a rank transformation approach, but here we need to transform the estimated contours on Laplace margins back to the original scale. Reversing the rank transformation is not possible for extreme contours beyond the largest observed value in a one or both margins. To overcome this, we now model each marginal distribution via a combination of the empirical distribution function and generalised Pareto distributions in the upper and lower tails, as outlined in Section 3.5.

3.2 Estimation of $\mathcal{C}_p^1$

Parametric estimation of a contour $\mathcal{C}_p^1$ requires a model for $R_L \mid (W_L = w)$ for each $w \in (-\pi, \pi]$, from which we can extract high quantiles. Since we are interested in values in the upper tail of the distribution of $R_L \mid (W_L = w)$, a natural model is provided by the generalised Pareto distribution (GPD) (Davison and Smith, 1990). Following Simpson and Tawn (2022), for each $w \in (-\pi, \pi]$, and for R_L exceeding some angle-dependent threshold $u_w > 0$, we propose to have

$$\Pr\left(R_L < r \mid R_L > u_w, W_L = w\right) = 1 - \left[1 + \xi(w)\left\{\frac{r - u_w}{\sigma(w)}\right\}\right]_+^{-1/\xi(w)}, \tag{7}$$

for $x_+ = \max(x, 0)$, $r > u_w$, $\sigma(w) > 0$ and $\xi(w) \in \mathbb{R}$. Once an appropriate threshold u_w has been selected, and estimates $\hat{\sigma}(w)$ and $\hat{\xi}(w)$ of the scale and shape parameters have been obtained, the value of $r_p(w)$ in (4) can be estimated by extracting the relevant quantile of (7), namely

$$\hat{r}_p(w) = u_w + \frac{\hat{\sigma}(w)}{\hat{\xi}(w)}\left[\left\{\frac{\hat{\zeta}_u(w)}{p}\right\}^{\hat{\xi}(w)} - 1\right], \tag{8}$$

with $\hat{\zeta}_u(w)$ denoting an estimate of $\zeta_u(w) = \Pr(R > u_w \mid W = w)$, often taken to be empirical. A corresponding estimate of the contour $\mathcal{C}_p^1$ is given by $\hat{\mathcal{C}}_p^1 := \{\hat{r}_p(w) : w \in (-\pi, \pi]\}$.

Applying the GPD requires us to select thresholds u_w and derive estimates of $\sigma(w)$ and $\xi(w)$ for $w \in (-\pi, \pi]$, using observations $(r_{L,1}, w_{L,1}), \ldots, (r_{L,n}, w_{L,n})$ of (R_L, W_L). There will not be sufficiently many equal observations of any given w value to be able to fit the generalised Pareto distribution in (7) directly for $R_L \mid (W_L = w)$. Instead, following Simpson and Tawn (2022) by assuming that the parameters u_w, $\sigma(w)$ and $\xi(w)$ in (7) vary smoothly with w, we employ a generalised additive modelling (GAM) framework, which was adapted for use with the generalised Pareto distribution by Youngman (2019).

The thresholds u_w are selected via quantile regression at some quantile level $(1 - p_u)$ with $p_u > p$, i.e., such that $\hat{\zeta}_u(w) = p_u$ for all $w \in (-\pi, \pi]$. This is achieved by fitting an asymmetric Laplace (Yu and Moyeed, 2001) GAM to $\log R_L$, with cyclic P-splines (see Wood, 2017) used for the model parameters with W_L as a covariate, before back-transforming to obtain quantile estimates of R_L.

For exceedances above the threshold function, we again propose using cyclic P-splines for the parameter $\log \sigma(w)$. A spline approach could also be used to model $\xi(w)$, but taking this to be constant has shown to produce less variable estimates in our experiments, which was also found by Simpson and Tawn (2022); we therefore fix $\xi(w) = \xi$ for all $w \in (-\pi, \pi]$. Estimation for the GPD-GAM framework is carried out in
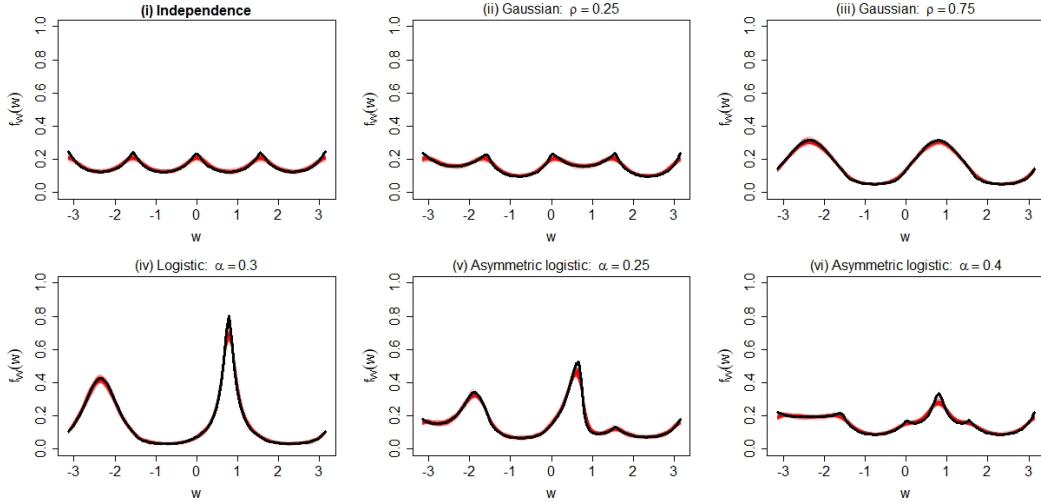
Figure 2: The angular density $f_{W_L}(w)$ for each of the models in Figure 1 (black). The red lines correspond to estimates using our cyclic kernel density approach, for 100 different samples of size $n = 10,950$.

the R package evgam (Youngman, 2020). Once all of these parameters have been calculated/estimated, equation (8) can again be used to extract radial quantiles at the required level.

The degree of the splines used within this modelling framework can impact the success of the approach. Following Simpson and Tawn (2022), we choose a spline degree $d \in \{1, 2, 3\}$, denoting the resulting radial quantile estimates by $\hat{r}_p^{(d)}(w)$, for $w \in (-\pi, \pi]$, in each case. Choice of d is discussed further in Section 3.4, along with the selection of the other necessary tuning parameters. Once the appropriate degree, denoted d^*, has been selected, our final estimate of $\mathcal{C}_p^1$ is $\hat{\mathcal{C}}_p^1 := \left\{ \hat{r}_p^{(d^*)}(w) : w \in (-\pi, \pi] \right\}$.

3.3 Estimation of $\mathcal{C}_p^2$

The main difference between the definitions of $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ is that the latter requires radial quantiles to be extracted at different probability levels, while for the former this level is constant. This requires us to estimate the angular density $f_{W_L}(w)$ and to take into account that some of the quantiles required in (5) may fall below the GPD threshold, such that we also require a model for non-extreme events.

We propose to use kernel density estimation for $f_{W_L}(w)$. We have adapted the standard approach implemented in the density base function of R to account for W_L being cyclic. To enforce this cyclicity, we first use data $\{(w_{L,1} - 2\pi), \dots, (w_{L,n} - 2\pi), w_{L,1}, \dots, w_{L,n}, (w_{L,1} + 2\pi), \dots, (w_{L,n} + 2\pi)\}$, resulting in a kernel density estimate $\hat{f}_{W_L}^*(w)$ over $(-3\pi, 3\pi]$. We then take $\hat{f}_{W_L}(w) = 3\hat{f}_{W_L}^*(w)$ for $w \in (-\pi, \pi]$. Default selections of the kernel bandwidth often rely on the standard deviation of the observations, but this will be affected by our choice to artificially replicate the data. Instead, we use the approach of Sheather and Jones (1991), which is a non-default option in R, that appears to work well in our setting. Estimation of $f_{W_L}(w)$ is demonstrated in Figure 2 for the six examples in Figure 1.

The required probability in (5) can be calculated for any angle w by substituting $\hat{f}_{W_L}(w)$ for $f_{W_L}(w)$; the required scaling constant c_p in (6) is calculated numerically. At any angle where the required ex-

ceedance probability in (5) is less than the threshold exceedance probability p_u, the appropriate radial quantile can be extracted from the GPD-GAM model, giving us an estimate $\hat{r}_p^*(w)$ of $r_p^*(w)$. However, for exceedance probabilities greater than p_u, the GPD is not appropriate. At any such levels, we adopt the same quantile regression model as for the threshold, using an asymmetric Laplace GAM for $\log R_L \mid (W_L = w)$ at this quantile level, and extract the appropriate estimate $\hat{r}_p^*(w)$ of $r_p^*(w)$. The overall estimate of $\mathcal{C}_p^2$ is $\hat{\mathcal{C}}_p^2 := \{\hat{r}_p^*(w) : w \in (-\pi, \pi]\}$.

3.4 Tuning parameter selection

The estimation procedures introduced in Sections 3.2 and 3.3 require the selection of several tuning parameters: the threshold-exceedance probability p_u; the number and location of knots in the cyclic P-splines; and the degree of the basis functions. We now discuss each of these choices in turn.

Simpson and Tawn (2022) propose to set $p_u = 0.5$, and while this does seem to provide quite a low choice for the threshold of the GPD, we have also found it to work well in our setting (see Section C of the Supplementary Material for simulation results that support this choice). We therefore adopt this as our default tuning parameter. The plots in Figure 1 demonstrate that there are important angles in particular settings, which it may be wise to consider when selecting spline knot locations. Under independence, the angles $w \in \{-\pi/2, 0, \pi/2, \pi\}$ correspond to the largest radial values on any given contour and the contours also exhibit a pointed shape here. If we are to capture such features using splines, it would be useful to have knots placed at these angles. For similar reasons, in the symmetric models exhibiting positive asymptotic dependence (cases (ii) and (iv)), the angle $w = \pi/4$ is particularly important. Angles $w \in \{3\pi/4, -\pi/4, -3\pi/4\}$ would also be important under negative asymptotic dependence or asymptotic dependence with one variable large while the other is small. All together, this suggests knots should be placed at all angles $w = -\pi + j\pi/4$, $j = 1, \ldots, 8$. However, this is generally not a sufficient number of knots, so we propose to add additional, equally-spaced knots between these angles, so the knot locations are $w = -\pi + 2j\pi/\kappa$, $j = 1, \ldots, \kappa$, with κ a multiple of eight; our default choice is $\kappa = 24$.

For $\mathcal{C}_p^1$, to choose the spline degree d^*, we again adopt the approach of Simpson and Tawn (2022) to derive local estimates of $r_p(w)$ and use these to inform the value of d^*; the procedure is summarised here. For a given angle w, a 'local estimate' is obtained by defining a neighbourhood of radial values corresponding to the m nearest points to w. Maximum likelihood estimation is then used to fit the generalised Pareto distribution in (7) using these radial values. This is repeated over k angles $w_j^* = -\pi + j\pi/100$, $j = 1, \ldots, k$, with the desired radial quantiles extracted and denoted by $\hat{r}_p^{\text{local}}(w_1^*), \ldots, \hat{r}_p^{\text{local}}(w_k^*)$. As in Simpson and Tawn (2022), our default tuning parameter choices are $m = 100$ and $k = 200$. It is clear from Figure 1 that some contours have smooth, curved shapes, while others are quite 'pointy'; this suggests that in some cases, linear basis functions may be appropriate, while in others, quadratic or cubic splines will perform better. We therefore estimate three different versions of the quantiles in (8) via the GAM-GPD approach, using degree $d = 1, 2, 3$ basis functions for both the $\log u_w$ model and $\log \sigma(w)$. We

denote the resulting estimates by $\hat{r}_p^{(d)}(w)$, for $w \in (-\pi, \pi]$, $d = 1, 2, 3$. To select the spline degree, we find

$$d^* = \underset{d=1,2,3}{\arg\min} \sum_{j=1}^{k} \left| \hat{r}_p^{(d)}(w_j^*) - \hat{r}_p^{\text{local}}(w_j^*) \right|,$$

i.e., the choice of basis function degree that allows the estimates to be as close to the local estimates as possible, but with the added benefit of smoothness in the resulting contour estimate. An equivalent approach is used to choose the spline degree in the estimation of $\mathcal{C}_p^2$, but at angles where a non-extreme radial quantile is required, an empirical quantile is used as the local estimate.

3.5   Marginal model

We end by describing the approach used to transform the original margins to Laplace scale. This is applied separately to each of the variables X and Y to obtain standard Laplace margins X_L and Y_L, from where the contour estimation procedure can be implemented. We focus on the model for X in this section, but an analogous approach is used for Y. Supposing that X has distribution function F_X, transformation to Laplace scale can be achieved through application of the probability integral transform, defining

$$X_L = \begin{cases} \log\{2F_X(X)\}, & F_X(X) \leq 1/2, \\ -\log[2\{1 - F_X(X)\}], & F_X(X) > 1/2, \end{cases} \tag{9}$$

so that the distribution function of X_L is as in (3). In practice, the distribution of X is unknown, so we use a combination of the empirical distribution function, denoted by $\tilde{F}_X$, and generalised Pareto models for both tails. That is, for v^L and v^U denoting lower and upper thresholds, respectively, we take

$$F_X(x) = \begin{cases} \lambda^L \left\{ 1 + \frac{\xi^L(v^L - x)}{\sigma^L} \right\}_+^{-1/\xi^L}, & x \leq v^L, \\ \tilde{F}_X(x), & v^L < x < v^U, \\ 1 - \lambda^U \left\{ 1 + \frac{\xi^U(x - v^U)}{\sigma^U} \right\}_+^{-1/\xi^U}, & x \geq v^U, \end{cases} \tag{10}$$

where $\sigma^L > 0, \sigma^U > 0, \xi^L \in \mathbb{R}, \xi^U \in \mathbb{R}, \lambda^L = \Pr\left(X < v^L\right)$ and $\lambda^U = \Pr\left(X > v^U\right)$. This is the two-tailed extension of the marginal modelling approach presented by Coles and Tawn (1991).

Model (10) requires the selection of two thresholds, which we choose such that $\lambda^L = \lambda^U = 0.05$. More sophisticated threshold selection techniques are available (see Scarrott and MacDonald (2012) for an overview, and Wadsworth (2016) or Northrop et al. (2017) for examples of recent developments), but we choose this approach for ease. Estimation of the scale (σ^L, σ^U) and shape (ξ^L, ξ^U) parameters is carried out via maximum likelihood estimation, as implemented in the R package ismev (Heffernan and Stephenson, 2018). Once all parameters in (10) are estimated, transformation (9) can be applied to each observation to provide the marginal observations on Laplace scale. Inversion of the transformation is also possible if the final results are required on the original scale; when inverting the empirical distribution function, we linearly interpolate where values between available X observations are required.

4 Simulation study

4.1 Overview

We now demonstrate how well the inferential procedure we introduced in Section 3 estimates the two types of contour defined in Section 2. We consider two scenarios: in Section 4.2 we work directly with data on Laplace margins; in Section 4.3 we allow the data to have more general, unknown margins. In the latter case, we compare our approach of estimating the contours on the transformed Laplace marginal scale before back transforming to the data scale, with an equivalent approach applied directly on the original scale of the data, which we demonstrate can lead to undesirable results.

Before presenting our results, we first detail the different tools and metrics that are used in our simulation study. We assess our methods based on how well the true contours are estimated, given data sampled from a range of different copula models (based on those presented in Figure 1). Throughout this section, we simulate data sets of size $n = 10,950$ (corresponding to a realistic scenario of daily observations over 30 years) and we use 100 replicates for each setting.

Our first strategy is to compare our estimated contours to the truth pictorially, allowing us to assess the estimation procedures by eye. We formalise this assessment by using two metrics for comparison. In the first, we count the number of points in the sample lying outside the estimated contour, denoting this by $\hat{n}_p$, so that the proportion of points outside the contour is $\hat{n}_p/n$. We then quantify the relative error from p, the target contour exceedance probability, to this empirical proportion by calculating the empirical relative error $\hat{n}_p/(np) - 1$. When $\hat{n}_p/(np) - 1 = 0$ we have correctly recovered the true probability.

The empirical relative error measure is suitable when the exceedance level p and the sample size n are such that a reasonable number of points is expected to lie outside the true contour. To cover more extreme cases, we also adopt the symmetric difference metric $\triangle$, used by Einmahl et al. (2013), which measures the difference between a true contour $\mathcal{C}$ and its estimate $\hat{\mathcal{C}}$ via

$$\mathcal{C} \triangle \hat{\mathcal{C}} = \int_{\{E_\mathcal{C} \setminus E_{\hat{\mathcal{C}}}\} \cup \{E_{\hat{\mathcal{C}}} \setminus E_\mathcal{C}\}} f_{X,Y}(x,y) dx dy,$$

where $E_\mathcal{C}$ denotes the set of $(x,y) \in \mathbb{R}$ that are more extreme than $\mathcal{C}$, with $E_{\hat{\mathcal{C}}}$ defined similarly, and $f_{X,Y}(x,y)$ is the joint density under the true marginal and copula specification. In our setting we calculate this in Laplace space using the polar coordinates (R_L, W_L) with associated joint density $f_{R_L,W_L}(r,w)$. Here, $E_{\hat{\mathcal{C}}}$ and $E_\mathcal{C}$ are given by $\{(r,w) : r > \hat{r}(w), -\pi < w \le \pi\}$ and $\{(r,w) : r > r_{true}(w), -\pi < w \le \pi\}$ respectively, with $\hat{r}(w)$ and $r_{true}(w)$ given implicitly by the forms of $\hat{\mathcal{C}}$ and $\mathcal{C}$, and

$$\mathcal{C} \triangle \hat{\mathcal{C}} = \int_{-\pi}^{\pi} \int_{r_{\min}(w)}^{r_{\max}(w)} f_{R_L,W_L}(r,w) dr dw \approx \sum_{i=1}^{k} (2\pi/k) \int_{r_{\min}(-\pi + \frac{2\pi i}{k})}^{r_{\max}(-\pi + \frac{2\pi i}{k})} f_{R_L,W_L}\left(r, -\pi + \frac{2\pi i}{k}\right) dr, \quad (11)$$

where $r_{\min}(w) = \min\{\hat{r}(w), r_{true}(w)\}$ and $r_{\max}(w) = \max\{\hat{r}(w), r_{true}(w)\}$, for each $w \in (-\pi, \pi]$. Symmetric difference results for contours $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ can be deceptive when comparing across different probability levels p. We instead present a scaled symmetric difference, i.e., $(\mathcal{C}_p^1 \triangle \hat{\mathcal{C}}_p^1)/p$ for contour definition 1.
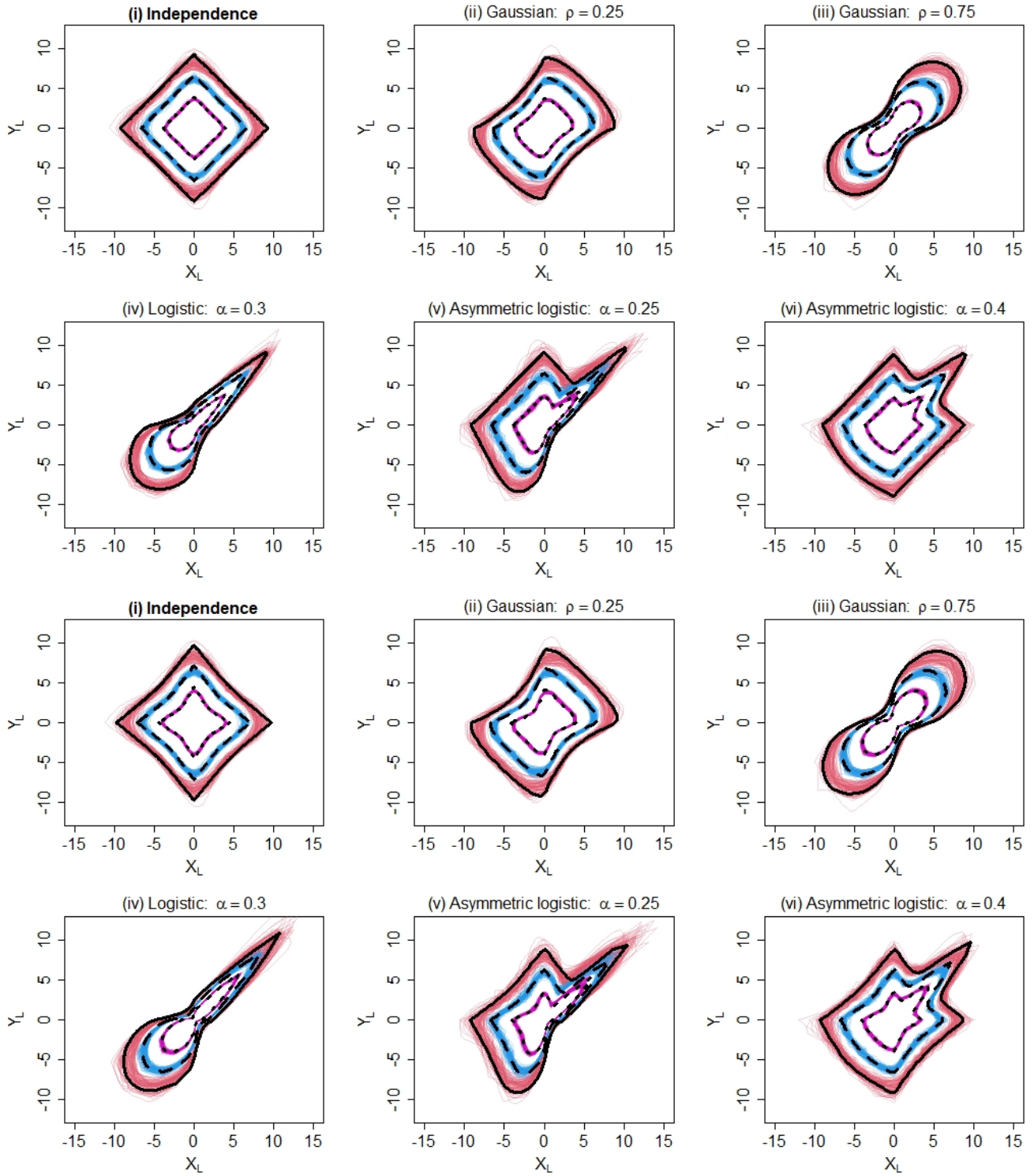
14

Figure 3: Top two rows: estimated contours $\hat{\mathcal{C}}^1_{0.1}$ (purple), $\hat{\mathcal{C}}^1_{0.01}$ (blue) and $\hat{\mathcal{C}}^1_{0.001}$ (red) for 100 data sets sampled under independence, from two Gaussian copulas, and from three bivariate extreme value copulas with logistic or asymmetric logistic models. Bottom two rows: equivalent estimated contours $\hat{\mathcal{C}}^2_p$, $p \in \{0.1, 0.01, 0.001\}$. The true contours are shown in black in all cases, for $p = 0.1$ (dotted lines), $p = 0.01$ (dashed lines) and $p = 0.001$ (solid lines).

4.2 Estimation on Laplace margins

Figure 3 shows visual assessments of our estimators where we have known Laplace marginal distributions and the dependence structure is given by the six copula models shown in Figure 1, covering both asymptotic independence and asymptotic dependence examples for the joint upper tail. Here, we produce estimates of $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ at probabilities $p \in \{0.1, 0.01, 0.001\}$. These plots show that our approach gives reasonably limited bias, but with increasing variability as p decreases, i.e., as we extrapolate further into the tails of the radial distributions at each angle. The performance appears to be equally good over all angles, for all six copulas and for both contour types. The methods pick up the contour shape whether it is pointed on the diagonal, or more rounded and flat in that region, for asymptotically dependent and asymptotically independent cases, respectively. The contours that are estimated for the asymmetric logistic models are especially impressive, given their very different levels of smoothness. The key to us achieving such flexibility comes from our selection of appropriate degrees of smoothing in the splines of the GPD parameters.

For a subset of the estimated contours in Figure 3, results for the empirical relative error metric $\hat{n}_p/(np) - 1$ are presented in Figures 4 and 5 for $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$, respectively. For clarity, we omit results for one of the Gaussian copulas (case (ii)) and one of the asymmetric logistic copulas (case (v)), but these are similar to those that are shown. Results are shown for both the estimated and true contours, with the same simulated data sets used in each case. We include these empirical relative errors for the truth to demonstrate the base level of sampling variability for this metric; we cannot expect to improve on this level of performance with the estimated contours. The empirical relative errors for the true contours naturally show no bias but exhibit increasing variability with decreasing p.

In Figures 4 and 5, there is no discernible difference between the performance of our estimators across the two different contour types, indicating that the additional complexity of estimating the angular density in $\mathcal{C}_p^2$ does not appear to have had an adverse effect. The performance, as judged by this metric, is very similar across the copula choices, with no evidence that contours for complex copulas are any more difficult to estimate than for the independence copula. Our estimation procedure is generally successful and unbiased, particularly for the $p = 0.1$ and $p = 0.01$ contours. For further extrapolation, i.e., the $p = 0.001$ contours, the empirical relative errors tend to be positive, but are less than one, suggesting the bias is small. Importantly, over all combinations of copula, p and contour type, the variability of the empirical relative error for our estimated contours is similar to that of the true contours, indicating that much of the variability in our estimators is due to sampling-variability rather than from our inference.

With a larger sample size, we would expect improved contour estimates for the $p = 0.001$ cases, but we can see from the results for the true contour that given this sample size, the empirical relative error metric has limitations at such small levels of p. In Figure 6, we therefore present a scaled symmetric difference $(\mathcal{C}_p^1 \triangle \hat{\mathcal{C}}_p^1)/p$ for $p \in \{0.1, 0.01, 0.001\}$. We only show results for the contour $\mathcal{C}_p^1$, given the similarity of results for the contour types seen in Figures 4 and 5. Although the symmetric differences naturally decrease with decreasing p, the scaled versions instead increase, in a similar pattern to the empirical relative errors. However, for the smallest p, they reveal that the asymmetric logistic with $\alpha = 0.25$ has
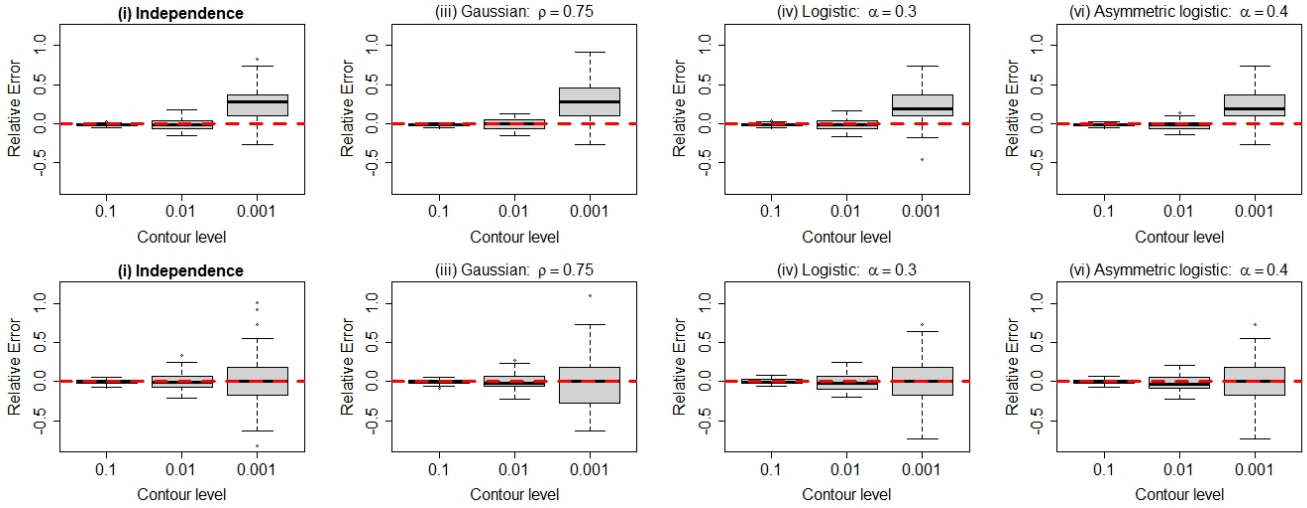
Figure 4: Boxplots of the empirical relative errors in the probability of lying outside the estimated contours $\hat{\mathcal{C}}_p^1$ (top row) and true contours $\mathcal{C}_p^1$ (bottom row) for 100 data sets sampled under independence (i), from a Gaussian copula with $\rho = 0.75$ (iii), and from two bivariate extreme value copulas with a logistic (iv) or asymmetric logistic (vi) model, with $p \in \{0.1, 0.01, 0.001\}$. Here, the numbering of the copulas is chosen to be consistent with Figures 1-3, but with cases (ii) and (v) omitted.
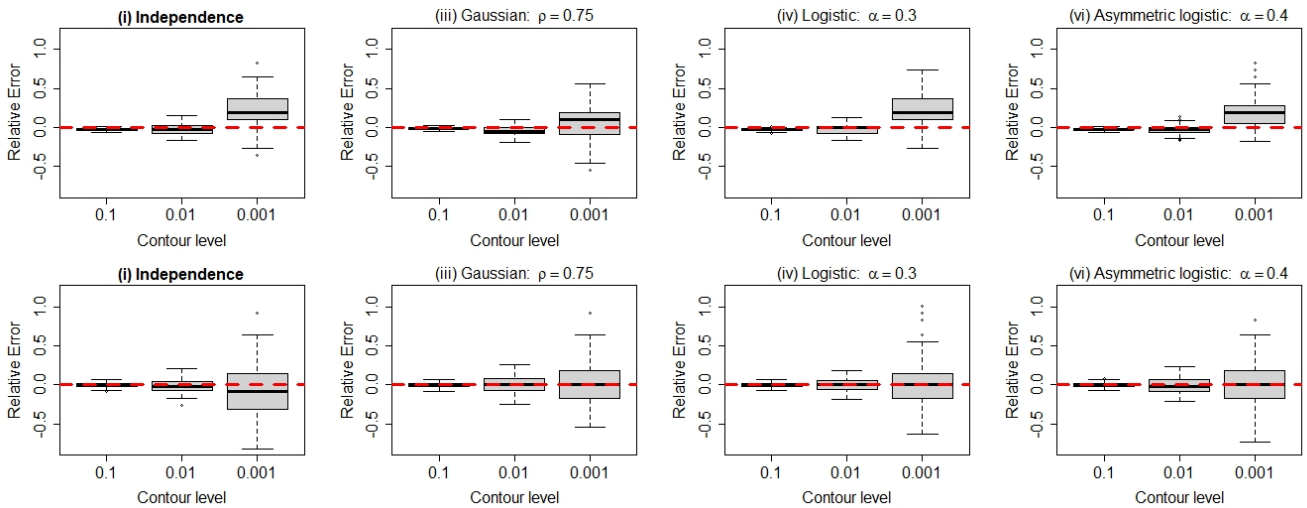

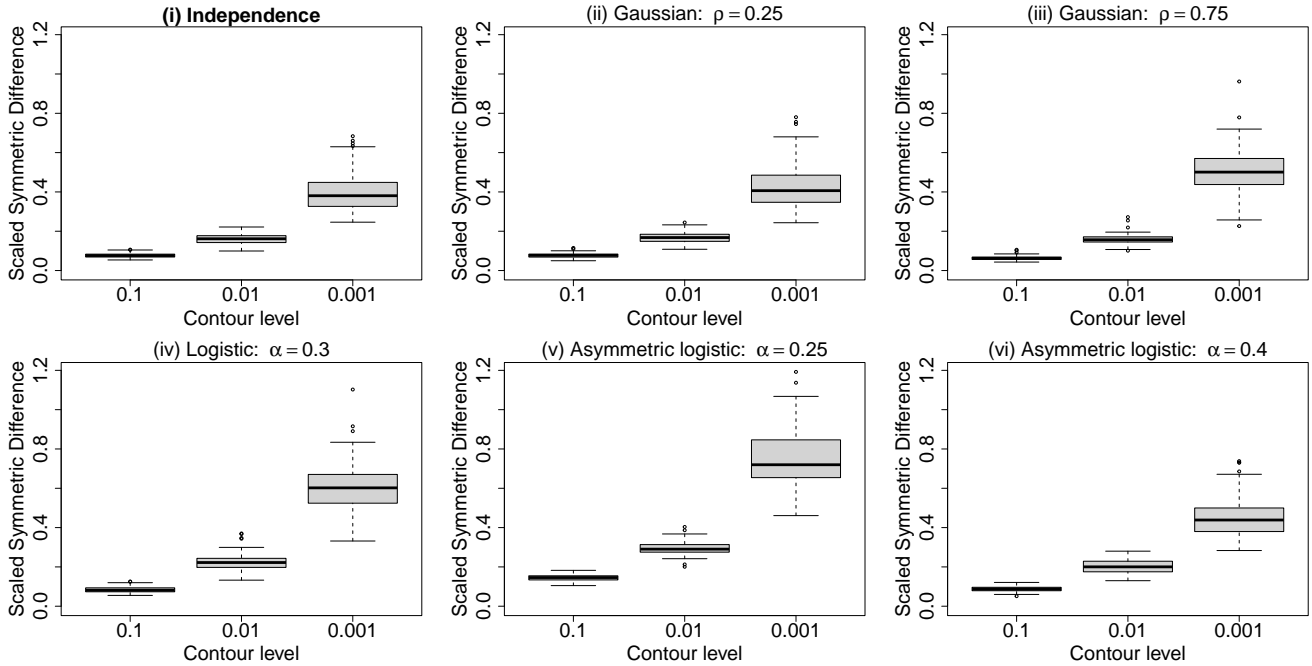Figure 5: These boxplots are equivalent to those in Figure 4, but for $\hat{\mathcal{C}}_p^2$.

Figure 6: Boxplots of the scaled symmetric difference $(\mathcal{C}_p^1 \, \triangle \, \hat{\mathcal{C}}_p^1)/p$ for 100 data sets sampled under independence (i), from two Gaussian copulas (ii-iii), and from three bivariate extreme value copulas with logistic or asymmetric logistic models (iv-vi), with $p \in \{0.1, 0.01, 0.001\}$. We used $k = 800$ in (11).

the most difficult contour to estimate of the copulas we have considered, with both the largest mean and variance in the scaled symmetric differences over replicated samples.

4.3 Estimation on original margins

A natural question is why we choose to define the contours on Laplace scale rather than considering (X, Y) as generic random variables, and defining the polar coordinates as $R = \sqrt{X^2 + Y^2}$ and $W = \tan^{-1}(Y/X)$ to construct the corresponding contours directly on the original (X, Y) scale. We considered this approach during development of this work, but found it often didn't work well, particularly when the dependence exhibited some mixture structure (e.g., in the asymmetric logistic model). We now provide additional simulations to demonstrate this issue.

One issue when working on the general (X, Y) scale is that the variables do not necessarily have a natural centre at $(0, 0)$, as they do in the Laplace case. To overcome this, we considered defining the marginal medians as m_X and m_Y for X and Y, respectively, and the polar coordinates as

$$R = \sqrt{(X - m_X)^2 + (Y - m_Y)^2}, \qquad W = \tan^{-1}\left\{\frac{Y - m_Y}{X - m_X}\right\}.$$

Contours can be defined analogously to $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$, but using these radial-angular coordinates. Estimation can be carried out in a similar way to the approach described in Section 3, but with m_X and m_Y added marginally to the resulting contour estimates as a final step, to obtain a contour centred on (m_X, m_Y).

Figure 7 presents results for the asymmetric logistic copula model from case (iv) of Figure 1 with
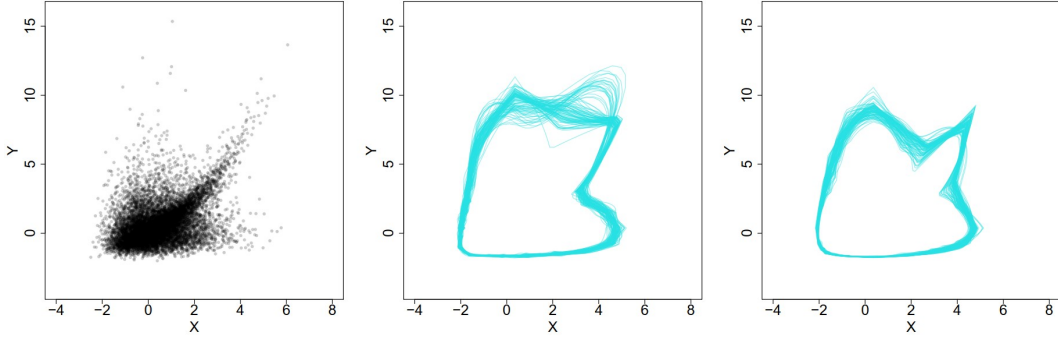
18

Figure 7: Left: example of data simulated from a bivariate extreme value copula with an asymmetric model, where X and Y have marginal GEV distributions with -0.1 and $+0.1$ shape parameters, respectively. Centre: 100 examples of contour estimates analogous to $\mathcal{C}^1_{0.01}$ but having polar coordinates defined on the original (X, Y) scale. Right: 100 examples of contour estimates of $\mathcal{C}^1_{0.01}$ calculated on Laplace margins and then transformed back to the original (X, Y) scale.

margins following generalised extreme value (GEV) distributions. For X, the GEV location, scale and shape parameters are $(0, 1, -0.1)$, and for Y they are $(0, 1, 0.1)$, so Y has the heavier upper tail. In Figure 7, we show a sample of size $n = 10,950$ from this model, as well as estimated contours at the 0.01 probability level from 100 samples of size n, using both approaches. Visually, Figure 7 clearly shows that the approach of transforming to Laplace margins before estimating the contours gives more consistent results, and there is no need to use the metrics introduced in Section 4.1 to see the distinct improvement using our proposed method. Specifically, our method is better at identifying the non-convex shape of the distribution. For these reasons, we prefer to define and estimate the contours on Laplace scale.

5 Air pollution application

We now apply our approach to air pollution data collected at a site in Bloomsbury, London, between 1 January 1993 and 31 December 2020. These data were provided by the Department for Environment Food and Rural Affairs (Defra) and are available to download from https://uk-air.defra.gov.uk[1]. We focus on daily maximum concentrations for two components of air pollution (nitric oxide NO and particulate matter PM_{10}) for the full year and also for summer (June, July, August) and winter (December, January, February) seasons. Values in the data set are rounded to the nearest whole number, so we add random noise simulated independently from a Uniform$(-0.5, 0.5)$ distribution (or Uniform$(0, 0.5)$ if the value is zero, to preserve non-negativity of concentration) to each observation, to mitigate issues with over-rounding. Approximately 9% of the NO values and 10% of the PM_{10} values are missing; we assume that they are missing at random and proceed with estimating our contours. Retaining only those instances where both pollutants are available leaves us with 8,865 complete observations.

In Figure 8, we present our estimated contours $\hat{\mathcal{C}}^1_p$ and $\hat{\mathcal{C}}^2_p$ for $p \in \{0.1, 0.01, 0.001\}$, using data from
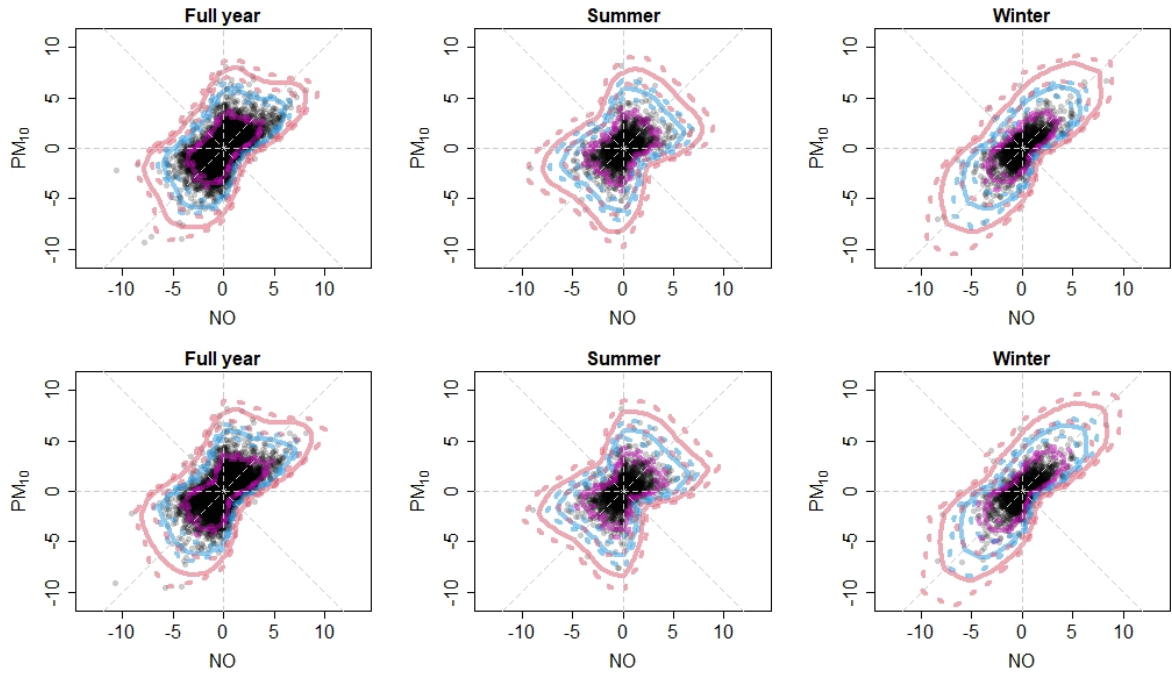

Figure 8: Estimates of $\mathcal{C}_p^1$ (top row) and $\mathcal{C}_p^2$ (bottom row) for nitric oxide and PM_{10} across the full year (left), summer only (centre) and winter only (right). Estimated contours (solid lines) are shown for $p \in \{0.1, 0.01, 0.001\}$, lying progressively further from $(0,0)$, alongside pointwise 95% bootstrapped confidence intervals (dotted lines). The grey points represent the data in each case.
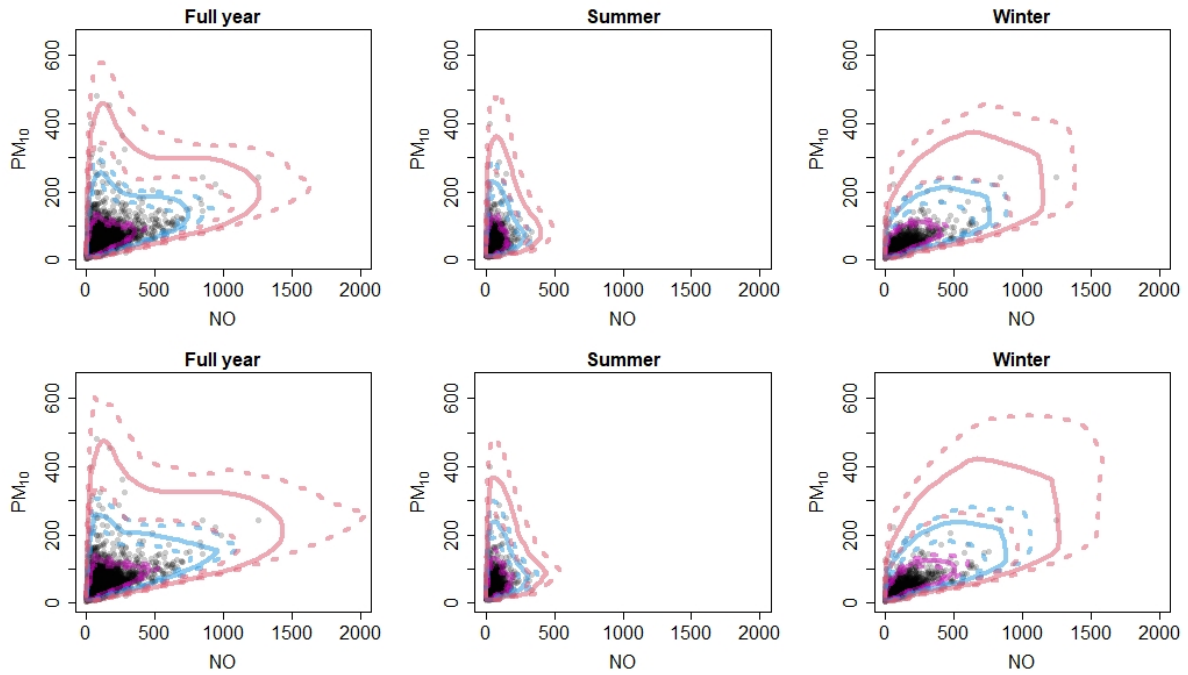

Figure 9: Data, estimated contours and confidence intervals from Figure 8 shown after back-transformation to the original margins of the data.

each of the three time periods (full year, summer, winter). The results are shown on Laplace scale, with the marginal transformations applied separately in each time period. We also provide pointwise 95% confidence intervals for the estimated radial quantiles at each angular value. This uncertainty is quantified using 200 bootstrapped samples via the block bootstrap scheme of Politis and Romano (1994); the samples are taken from the data after transformation to Laplace margins with the missing values retained. As the time series of NO and PM_{10} observations both exhibit strong auto-correlation, we fix the average block-length in the bootstrapping procedure to be 14 in all cases. Carrying out bootstrapping after marginal transformation means that only uncertainty in the dependence features is taken into account in the confidence intervals, but allows for equivalent contours and confidence intervals to be obtained on the original scale by reversing the marginal transformations; these results are given in Figure 9.

From Figures 8 and 9, we can see that the dependence properties of the data vary between season. In particular, there is much stronger extremal dependence in both the upper and lower tails of NO and PM_{10} in the winter than in the summer; this feature is visible in the data and successfully captured by our contour estimates. Further, our contours are able to capture the mixture of dependence features that is present when considering data for the full year. The maximum value across NO concentrations is 336 in summer and 1251 in winter, while PM_{10} has a smaller variation between seasons, with a maximum of 399 in summer and 243 in winter. Since the winter months dominate the upper tail of the NO values, the feature from the summer months that NO can be large while PM_{10} takes values close to its median is obscured when considering data across the full year. On the other hand, features corresponding to the upper tail of PM_{10} in both summer and winter are preserved when considering the full year, i.e., NO and PM_{10} can be simultaneously large but PM_{10} can also be large when NO takes values around its median.

Focusing only on the separate summer and winter plots in Figure 9, we now consider the implications that these estimated environmental contours have for potentially setting environmental controls. Given that large values of each pollutant have impacts on health, the key segments of the environmental contours in this case are where either of PM_{10} and NO are large. Large values of NO only occur in winter whereas very similar large levels of PM_{10} occur in both seasons, so if there are implications on health from jointly large values of these two pollutants, winter is the period of most risk. However, it is these segments of the environmental contours that are most uncertain, particularly in winter, with approximate 95% confidence intervals differing up to $\pm 25\%$ from the point estimates, even without accounting for the marginal estimation uncertainty. For these two pollutants, there are relatively small differences in the estimates of the two different environmental contours in the key segments; this is due to the angular distributions on Laplace margins being reasonably uniform here. Here, we have focused on the contour $\mathcal{C}_p^1$ but equally, given interest in large values of the pollutants, we could have focused instead on the contour $\mathcal{C}_p^1([0, \pi/2])$ proposed in Section 2.2.

Having estimated the environmental contours, contributions from domain experts are now essential. Although in this pollutant example they may not have a formally defined structure function, they should be able to identify combinations of pollutants on the contours that are the most likely to lead to adverse health effects, possibly relating to a range of different symptoms. Likewise, the pollutant regulatory

bodies can use this information to see if certain combinations of pollutants they have concerns about are likely to occur too often. If this is the case, they could look to set new control limits, or aim to remove or reduce the sources of these air pollutants so that future extreme combination levels are lowered.

6   Discussion

In some application areas, there may be only a subset of angles W_L that are of interest. If this were the case, the contour definitions could be adapted to only consider radial exceedance probabilities for angles in a subset of $(-\pi, \pi]$. The estimation procedure we present could still be used, but with radial quantiles at a different set of levels extracted from the final estimates.

An area where our approach could be improved is in the construction of the splines used in the GPD threshold model and log-scale parameter. We use basis functions of the same degree across all knots, with this degree chosen as described in Section 3, irrespective of whether the copula is asymptotically independent or asymptotically dependent. However, in some cases, it may be more appropriate to allow different types of basis function at different knots. An example of this is the asymptotically dependent logistic model, where in case (ii) of Figure 1, we can see that the contours are 'pointy' in the upper-right quadrant (so that linear basis functions may be more appropriate) and smooth in the lower-left quadrant (so that quadratic or cubic basis functions may perform better). Considering such bespoke spline constructions would add a significant level of complexity to our modelling procedure, so we did not pursue this idea here, but it does present a possible avenue for future work. Critically, our proposed method does not require us to pre-specify whether or not the copulas are asymptotically dependent, and this is a feature that it would be desirable to retain in any extension of the methodology.

Although the primary motivation for developing these novel environmental contours was for applications to engineering safety design and for setting environmental standards, an alternative potential use for these methods is in bivariate outlier or anomaly detection. Of course, multivariate outlier detection is a very widely studied topic, which goes well beyond simply using extreme value methods (see, e.g., Hubert et al., 2015), although they are highly relevant for some problems (Dupuis and Morgenthaler, 2002). Unlike the latter paper, we use univariate extreme value methods without needing to adopt a copula model, so adapting our approach to the anomaly detection setting may have the advantage of additional flexibility over existing methods.

We have focused on developing methodology for estimating bivariate contours, but the contour definitions can be extended to higher dimensions. With more than three dimensions, visualising the contours is impossible, and care would be needed to present the results in a usable format that offers easy interpretation. An option may be to project the contours onto lower dimensional subsets of interest.

Finally, the approach we have taken in this paper involves a combination of extending previous work related to limit set estimation and gauge functions (Simpson and Tawn, 2022), and exploiting an alternative radial-angular representation for a given, common marginal distribution (here, Laplace). This appears to be a profitable line of future research; specifically, while our paper has been in the review process, two independently developed papers (Mackay and Jonathan, 2023; Papastathopoulos et al., 2023) have been

uploaded to *arXiv* that consider aspects linked to these strategies, though with different objectives.

Acknowledgements

We thank the two referees for their clear, thoughtful, and rapid comments which have helped both the presentation of the manuscript and comparisons with previous approaches.

Declarations

Conflicts of interest

The authors declare that they have no conflicts of interest.

References

Balkema, A. A., Embrechts, P., and Nolde, N. (2010). Meta densities and the shape of their sample clouds. *Journal of Multivariate Analysis*, 101(7):1738–1754.

Barnett, V. (1976). The ordering of multivariate data (with discussion). *Journal of the Royal Statistical Society. Series A (General)*, 139(3):318–355.

Brozius, H. and de Haan, L. (1987). On limiting laws for the convex hull of a sample. *Journal of Applied Probability*, 24(4):852–862.

Bruun, J. T. and Tawn, J. A. (1998). Comparison of approaches for estimating the probability of coastal flooding. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 47(3):405–423.

Cai, J.-J., Einmahl, J. H. J., and de Haan, L. (2011). Estimation of extreme risk regions under multivariate regular variation. *The Annals of Statistics*, 39(3):1803–1826.

Castro-Camilo, D., de Carvalho, M., and Wadsworth, J. L. (2018). Time-varying extreme value dependence with application to leading European stock markets. *Annals of Applied Statistics*, 12(1):283–309.

Chai, W. and Leira, B. J. (2018). Environmental contours based on inverse SORM. *Marine Structures*, 60:34–51.

Coles, S. G., Heffernan, J. E., and Tawn, J. A. (1999). Dependence measures for extreme value analyses. *Extremes*, 2(4):339–365.

Coles, S. G. and Tawn, J. A. (1991). Modelling extreme multivariate events. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):377–392.

Coles, S. G. and Tawn, J. A. (1994). Statistical methods for multivariate extremes: an application to structural design (with discussion). *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43(1):1–48.

Davis, R., Mulrow, E., and Resnick, S. I. (1987). The convex hull of a random sample in $\mathbb{R}^2$. *Communications in Statistics. Stochastic Models*, 3(1):1–27.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(3):393–425.

Dupuis, D. J. and Morgenthaler, S. (2002). Robust weighted likelihood estimators with an application to bivariate extreme value problems. *Canadian Journal of Statistics*, 30(1):17–36.

Eddy, W. F. and Gale, J. D. (1981). The convex hull of a spherically symmetric sample. *Advances in Applied Probability*, 13(4):751–763.

Einmahl, J. H. J., de Haan, L., and Krajina, A. (2013). Estimating extreme bivariate quantile regions. *Extremes*, 16:121–145.

Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, 55(292):698–707.

Hafver, A., Agrell, C., and Vanem, E. (2022). Environmental contours as Voronoi cells. *Extremes*, 25:451–486.

Haselsteiner, A. F., Coe, R. G., Manuel, L., Chai, W., Leira, B., Clarindo, G., Guedes Soares, C., Hannesdóttir, A., Dimitrov, N., Sander, A., Ohlendorf, J.-H., Thoben, K.-D., de Hauteclocque, G., Mackay, E., Jonathan, P., Qiao, C., Myers, A., Rode, A., Hildebrandt, A., Schmidt, B., Vanem, E., and Huseby, A. B. (2021). A benchmarking exercise for environmental contours. *Ocean Engineering*, 236:109504.

Haselsteiner, A. F., Ohlendorf, J.-H., Wosniok, W., and Thoben, K.-D. (2017). Deriving environmental contours from highest density regions. *Coastal Engineering*, 123:42–51.

Hasofer, A. M. and Lind, N. C. (1974). Exact and invariant second-moment code format. *Journal of the Engineering Mechanics Division*, 100(1):111–121.

Heffernan, J. E. and Stephenson, A. G. (2018). *ismev: An Introduction to Statistical Modeling of Extreme Values*. R package version 1.42.

Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 66(3):497–546.

Hubert, M., Rousseeuw, P. J., and Segaert, P. (2015). Multivariate functional outlier detection. *Statistical Methods & Applications*, 24:177–202.

Huseby, A. B., Vanem, E., and Natvig, B. (2013). A new approach to environmental contours for ocean engineering applications based on direct Monte Carlo simulations. *Ocean Engineering*, 60:124–135.

Huseby, A. B., Vanem, E., and Natvig, B. (2015). Alternative environmental contours for structural reliability analysis. *Structural Safety*, 54:32–45.

Keef, C., Tawn, J. A., and Lamb, R. (2013). Estimating the probability of widespread flood events. *Environmetrics*, 24(1):13–21.

Ledford, A. W. and Tawn, J. A. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 59(2):475–499.

Ledford, A. W. and Tawn, J. A. (1998). Concomitant tail behaviour for extremes. *Advances in Applied Probability*, 30(1):197–215.

Mackay, E. and Haselsteiner, A. F. (2021). Marginal and total exceedance probabilities of environmental contours. *Marine Structures*, 75:102863.

Mackay, E. and Jonathan, P. (2023). Modelling multivariate extremes through angular-radial decomposition of the density function. *arXiv:2310.12711*.

Murphy-Barltrop, C. J. R., Wadsworth, J. L., and Eastoe, E. F. (2023). New estimation methods for extremal bivariate return curves. *Environmetrics*, 34(5):e2797.

Nolde, N. (2014). Geometric interpretation of the residual dependence coefficient. *Journal of Multivariate Analysis*, 123:85–95.

Nolde, N. and Wadsworth, J. L. (2022). Linking representations for multivariate extremes via a limit set. *Advances in Applied Probability*, 54(3):688–717.

Northrop, P. J., Attalides, N., and Jonathan, P. (2017). Cross-validatory extreme value threshold selection and uncertainty with application to ocean storm severity. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 66(1):93–120.

Papastathopoulos, I., de Monte, L., Campbell, R., and Rue, H. (2023). Statistical inference for radially-stable generalized Pareto distributions and return level-sets in geometric extremes. *arXiv:2310.06130*.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313.

Poon, S.-H., Rockinger, M., and Tawn, J. A. (2004). Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *The Review of Financial Studies*, 17(2):581–610.

Ross, E., Astrup, O. C., Bitner-Gregersen, E., Bunn, N., Feld, G., Gouldby, B., Huseby, A., Liu, Y., Randell, D., Vanem, E., and Jonathan, P. (2020). On environmental contours for marine and coastal design. *Ocean Engineering*, 195:106194.

Scarrott, C. and MacDonald, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT-Statistical Journal*, 10(1):33–60.

Sheather, S. J. and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(3):683–690.

Simpson, E. S. and Tawn, J. A. (2022). Estimating the limiting shape of bivariate scaled sample clouds for self-consistent inference of extremal dependence properties. *arXiv:2207.02626*.

Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika*, 75(3):397–415.

Tawn, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika*, 77(2):245–253.

Wadsworth, J. L. (2016). Exploiting structure of maximum likelihood estimators for extreme value threshold selection. *Technometrics*, 58(1):116–126.

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

Winterstein, S. R., Ude, T. C., Cornell, C. A., Bjerager, P., and Haver, S. (1993). Environmental parameters for extreme response: Inverse FORM with omission factors. In *Proc. 6th Int. Conf. on Structural Safety and Reliability, Innsbruck, Austria*, pages 551–557.

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, Second Edition*. Chapman and Hall/CRC, Boca Raton.

Youngman, B. D. (2019). Generalized additive models for exceedances of high thresholds with an application to return level estimation for U.S. wind gusts. *Journal of the American Statistical Association*, 114(528):1865–1879.

Youngman, B. D. (2020). *evgam: Generalised Additive Extreme Value Models*. R package version 0.1.4.

Yu, K. and Moyeed, R. A. (2001). Bayesian quantile regression. *Statistics & Probability Letters*, 54:437–447.

Supplementary Material:
Inference for new environmental contours using extreme value analysis

A   Justification of the scaling in equation (5)

For contour $\mathcal{C}_p^2$, we scale the radial exceedance probabilities by the angular density, assuming that $f_{W_L}(w) > 0$ for each $w \in (-\pi, \pi]$. If we set

$$\Pr\left\{R_L > r_p^*(w) \mid W_L = w\right\} = \frac{c_p}{f_{W_L}(w)},$$

the probability of lying outside the corresponding contour $\left\{r_p^*(w) : w \in (-\pi, \pi]\right\}$ would be

$$\int_{-\pi}^{\pi} \Pr\left\{R_L > r_p^*(v)|W_L = v\right\} dF_{W_L}(v) = \int_{-\pi}^{\pi} \frac{c_p}{f_{W_L}(v)} f_{W_L}(v) dv = \int_{-\pi}^{\pi} c_p \, dv = 2\pi c_p,$$

and equating this with p leads us to set $c_p = p/(2\pi)$, and hence

$$\Pr\left\{R_L > r_p^*(w) \mid W_L = w\right\} = \frac{p}{2\pi f_{W_L}(w)}.$$

If $f_{W_L}(w) < p/(2\pi)$, this implies $\Pr\left\{R_L > r_p^*(w) \mid W_L = w\right\} > 1$. To avoid this, we set

$$\Pr\left\{R_L > r_p^*(w) \mid W_L = w\right\} = \frac{c_p}{\max\left\{f_{W_L}(w), \epsilon\right\}}, \tag{S1}$$

for some $\epsilon \geq m_f := \min_{w^* \in (-\pi, \pi]} f_{W_L}(w^*)$, chosen to guarantee valid probabilities in (S1) for all $w \in (-\pi, \pi]$. We set $\epsilon = k m_f$ for some $k \geq 1$ to be selected. The constraint

$$\Pr\left\{R_L > r_p^*(w) \mid W_L = w\right\} \leq 1 \quad \Longleftrightarrow \quad c_p \leq \max\left\{f_{W_L}(w), k m_f\right\}$$

must hold for all $w \in (-\pi, \pi]$. This implies that

$$\begin{aligned}
c_p &\leq \min_{w \in (-\pi, \pi]} \left[\max\left\{f_{W_L}(w), k m_f\right\}\right] \\
&= \max\left[\left\{\min_{w \in (-\pi, \pi]} f_{W_L}(w)\right\}, k m_f\right] \\
&= k m_f \qquad (\text{since } k \geq 1). \tag{S2}
\end{aligned}$$

To ensure the probability of lying outside the contour is p, we also need

$$c_p \int_{-\pi}^{\pi} \max\left\{f_{W_L}(v), k m_f\right\}^{-1} f_{W_L}(v) dv = p$$

$$\Rightarrow \quad c_p = p \left[\int_{-\pi}^{\pi} \min\left\{1, \frac{f_{W_L}(v)}{k m_f}\right\} dv\right]^{-1}. \tag{S3}$$

Combining equations (S2) and (S3), we have the requirement that

$$p \left[\int_{-\pi}^{\pi} \min\left\{1, \frac{f_{W_L}(v)}{k m_f}\right\} dv\right]^{-1} \leq k m_f$$

$$\Rightarrow \quad p \leq \int_{-\pi}^{\pi} \min\left\{k m_f, f_{W_L}(v)\right\} dv. \tag{S4}$$

We can find an upper bound on the right-hand side of (S4) by noting that

$$\int_{-\pi}^{\pi} \min\left\{km_f, f_{W_L}(v)\right\} dv \le \min\left\{\int_{-\pi}^{\pi} km_f dv, \int_{-\pi}^{\pi} f_{W_L}(v)dv\right\} = \min\left\{2\pi km_f, 1\right\}.$$

Substituting this for the right-hand side of (S4), we require the smallest $k \ge 1$ such that

$$p \le \min\left\{2\pi km_f, 1\right\}. \tag{S5}$$

Setting $k = p/(2\pi m_f)$ means that equation (S5) simplifies to $p \le \min(p, 1) = p$, but does not guarantee that $k \ge 1$. Hence, we instead set $k = \max\left\{p/(2\pi m_f), 1\right\}$, which still guarantees that (S5) holds. Substituting this back into (S1), our chosen scaling for each $w \in (-\pi, \pi]$ becomes

$$\Pr\left\{R_L > r_p^*(w) \mid W_L = w\right\} = \frac{c_p}{\max\left\{f_{W_L}(w), \ p/(2\pi), \ m_f\right\}} = \frac{c_p}{\max\left\{f_{W_L}(w), \ p/(2\pi)\right\}}.$$

Additionally, since all components are non-negative, this leads to a valid probability.

B   Calculations for obtaining the contours defined in Section 2

Here, we explain how to calculate the theoretical contours $\mathcal{C}_p^1$ and $\mathcal{C}_p^2$ for a fixed value of $p \in (0, 1)$ and (X_L, Y_L) having joint density $f_{X_L, Y_L}(x, y)$.

In the definition of $\mathcal{C}_p^1$, for each $w \in (-\pi, \pi]$, we are interested in the radial quantile $r_p(w)$ defined in (4). To calculate $r_p(w)$ theoretically, we need to find the conditional density of $R_L \mid W_L$ and then use this to extract the required quantile. Inverting the transformation to polar coordinates requires $X_L = R_L \cos W_L$ and $Y_L = R_L \sin W_L$, so that the Jacobian of the transformation is given by $J = R_L$. Hence, the joint density of (R_L, W_L) is

$$f_{R_L, W_L}(r, w) = r f_{X_L, Y_L}(r \cos w, r \sin w), \qquad r > 0, \quad w \in (-\pi, \pi].$$

Therefore, for each fixed $w \in (-\pi, \pi]$ and $p \in (0, 1)$, the quantile $r_p(w)$ is the solution to

$$\frac{\int_0^{r_p(w)} t f_{X_L, Y_L}(t \cos w, t \sin w) dt}{f_{W_L}(w)} = 1 - p, \tag{S6}$$

where

$$f_{W_L}(w) = \int_0^\infty t f_{X_L, Y_L}(t \cos w, t \sin w) dt.$$

In general, both integrals in (S6) will need to be found numerically, as will the corresponding value of $r_p(w)$.

Similarly, for $\mathcal{C}_p^2$, with the scaling constant c_p defined as in (6), we find $r_p^*(w)$ by numerically solving

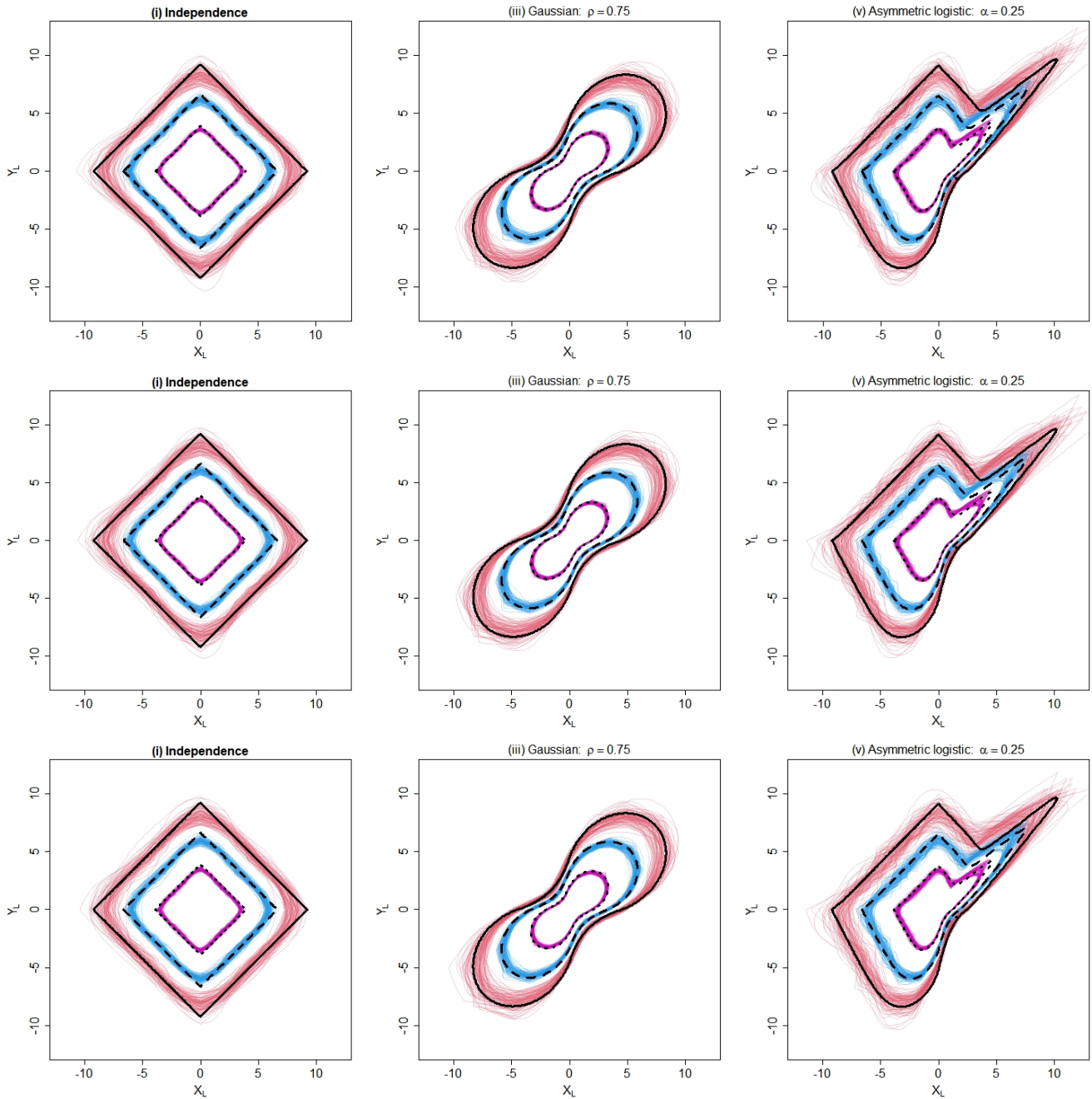$$\frac{\int_0^{r_p^*(w)} t f_{X_L, Y_L}(t \cos w, t \sin w) dt}{f_{W_L}(w)} = 1 - \frac{c_p}{\max\left\{f_{W_L}(w), p/(2\pi)\right\}}.$$

Figure A: Estimated contours $\hat{\mathcal{C}}_{0.1}^1$ (purple), $\hat{\mathcal{C}}_{0.01}^1$ (blue) and $\hat{\mathcal{C}}_{0.001}^1$ (red) for 100 data sets sampled under independence, from a Gaussian copula, and from a bivariate extreme value copulas an asymmetric logistic model. These estimates are obtained with $p_u = 0.4$ (top row), $p_u = 0.3$ (middle row) or $p_u = 0.2$ (bottom row). The true contours are shown in black in all cases, for $p = 0.1$ (dotted lines), $p = 0.01$ (dashed lines) and $p = 0.001$ (solid lines).

C   Simulation results validating our threshold choice in Section 3.4

When modelling using a generalised Pareto distribution, as in our model (7) for the tail of $R_L \mid W_L = w$, selection of an appropriate threshold is a key consideration. We choose to apply quantile regression to calculate our thresholds u_w, where the threshold exceedance probability $\Pr(R_L > u_w \mid W_L = w) = p_u$ is fixed across all angles $w \in (-\pi, \pi]$. Our default throughout the paper was to set $p_u = 0.5$. We now demonstrate that this is a reasonable choice, by providing additional simulation results for higher thresholds, with exceedance probabilities $p_u \in \{0.4, 0.3, 0.2\}$. All other tuning parameters are as described in Section 3.4 of the paper.

Figure A shows simulation results equivalent to Figure 3 of the main paper for $\mathcal{C}_p^1$ and cases of independence (i), a Gaussian copula with $\rho = 0.75$ (iii) and an asymmetric logistic copula with $\alpha = 0.25$ (v); results for $\mathcal{C}_p^2$ and the other copula examples are omitted since the results are very similar. There is clear stability across all choices of p_u, suggesting that selection of the threshold level is not especially crucial in our estimation procedure, and our choice to set $p_u = 0.5$ is reasonable.