

ASYMPTOTIC INDEPENDENCE FOR UNIMODAL DENSITIES

GUUS BALKEMA,* *University of Amsterdam*

NATALIA NOLDE,** *ETH Zürich*

Abstract

Asymptotic independence of the components of random vectors is a concept used in many applications. The standard criteria for checking asymptotic independence are given in terms of distribution functions (DFs). DFs are rarely available in an explicit form, especially in the multivariate case. Often we are given the form of the density or, via the shape of the data clouds, we can obtain a good geometric image of the asymptotic shape of the level sets of the density. In this paper we establish a simple sufficient condition for asymptotic independence for light-tailed densities in terms of this asymptotic shape. This condition extends Sibuya's classic result on asymptotic independence for Gaussian densities.

Keywords: Asymptotic independence; blunt; homothetic density; level set; skew normal; star shaped

2010 Mathematics Subject Classification: Primary 60G55; 60G70; 62E20

1. Introduction

The purpose of the present paper is to provide simple sufficient conditions that ensure asymptotic independence of the components of random vectors whose probability distribution is described by a density. Standard criteria for checking asymptotic independence are given in terms of distribution functions (DFs). However, these are not always available in an explicit form in the multivariate case, and they give little insight into what large samples from a distribution will look like. Often we are given a density in analytic form. For light-tailed densities, the data clouds give a good geometric image of the asymptotic shape of the level sets. Hence, it is of interest to have conditions for asymptotic independence in terms of the shape of the level sets of the underlying density, or in terms of a limiting shape for data clouds.

For vector-valued data, it is standard practice to plot the bivariate sample clouds for all component pairs. In our final result it is the asymptotic behaviour of the shape of these bivariate sample clouds, as the size of the data set increases, that determines asymptotic independence of the coordinates for the underlying multivariate distribution.

Unimodal densities whose level sets all have the same shape are called homothetic. The decay along any ray then is the same up to a scale constant depending on the direction, and, hence, the concept of light and heavy tails is well defined. This remains true if we only assume that the level sets have the same shape asymptotically. Our primary focus here is on light-tailed

Received 11 January 2010.

* Postal address: Department of Mathematics, University of Amsterdam, Science Park 904, 1098XH Amsterdam, The Netherlands. Email address: a.a.balkema@uva.nl

** Postal address: Department of Mathematics and RiskLab, ETH Zürich, Raemistrasse 101, 8092 Zürich, Switzerland. Email address: natalia.lysenko@math.ethz.ch

densities, but for additional insight we also include some results for the heavy-tailed case. Throughout the paper, we assume continuity of DFs and of densities.

Intuitively, for bivariate data, asymptotic independence means that large values in one coordinate are unlikely to be accompanied by large values in the other coordinate. Situations with a low chance of simultaneous extremes are often encountered in practice, for example, in applications which involve modelling environmental data (see, e.g. [16]) or network traffic data (see, e.g. [18]); see [24] for further references. It is a well-known result, dating back to 1960 (see [25]), that the components of a vector with a Gaussian density are asymptotically independent whatever the correlation. Asymptotic independence also holds for light-tailed elliptical densities (see, e.g. [14]). As a further generalization, we shall show that a vector with a continuously differentiable homothetic light-tailed density whose level sets are convex has asymptotically independent components.

Theorem 1.1. *Let D be a bounded open convex set in \mathbb{R}^d , containing the origin, with a C^1 boundary (i.e. at each boundary point there is a unique tangent plane). Let $c_n > 0$ decrease to 0 such that $c_{n+1}/c_n \rightarrow 0$, and let r_n be positive reals such that $r_{n+1}/r_n \rightarrow 1$. If X is a random vector in \mathbb{R}^d with a continuous probability density f whose level sets satisfy*

$$\{f > c_n\} = r_n D, \quad n \geq n_0, \quad (1.1)$$

then the components of X are asymptotically independent.

We shall prove more general results, namely Theorems 3.1 and 3.2, below. The generalizations we introduce are simple, but they result in theorems for which some extra terminology has to be developed.

- The condition that the level sets all be of a given shape is replaced by the condition that the level sets can be scaled to converge to a limit shape D . Theorem 1.1 remains valid if we replace (1.1) by the limit relation $\{f > c_n\}/r_n \rightarrow D$.
- Asymptotic independence holds if it holds for the bivariate marginals; in this spirit, our theorem imposes conditions on the projection D_{12} of the set D on the (x_1, x_2) -coordinate plane.
- Convex level sets are replaced by star-shaped level sets.
- The condition of a smooth (C^1) boundary is replaced by a condition which only affects the maximum: the coordinatewise supremum of the points in D_{12} should not be a boundary point of D_{12} . Such a set is called blunt.
- Our final result (Theorem 4.1) is on bivariate sample clouds. If these converge onto a blunt star-shaped set, asymptotic independence holds.

The conditions in Theorem 1.1 above on the sequences (c_n) and (r_n) ensure that the density has light tails. If $c_{n+1}/c_n \rightarrow 1$ then the density has heavy tails, and the same condition on the shape of the level sets implies asymptotic dependence.

The paper is organized as follows. In Section 2 we discuss the concept of asymptotic independence. In Section 3 we introduce the class \mathcal{H} of continuous multivariate homothetic densities with star-shaped level sets. The simple structure of these densities makes them a good starting point in our investigation of the relation between the shape of the level sets and asymptotic independence. Here we formulate our main results (Theorems 3.1 and 3.2) for light-tailed densities, and a counterpart (Theorem 3.3) for heavy-tailed densities. In Section 4

we introduce sample clouds, and state conditions for asymptotic independence in terms of their limiting shape (Theorem 4.1). Section 5 provides various examples. The appendix clears up a number of minor points in the main text, by providing supplementary results and counterexamples.

2. Asymptotic independence

In this section we discuss asymptotic independence for multivariate distributions. We begin with a heuristic approach in the bivariate setting. This will clarify the significance of the concept for risk management. We discuss the commonly used criteria for asymptotic independence, the relation with multivariate extreme value theory, and describe possible forms of asymptotic dependence.

2.1. Heuristics

In finance we are interested in the future value of stocks, say in one year's time. Let X and Y denote the future value of two stocks. Suppose that the distribution of the pair (X, Y) is given by a DF F with continuous marginals F_1 and F_2 . We are concerned with the risk that the stocks have a low value at this future date. Let x_p denote the p -quantile of the first stock, X , and y_p denote the p -quantile of the second stock. The probability that the values of both stocks lie below the p -quantile for some given small value of $p \in (0, 1)$ is $F(x_p, y_p)$. A risk-averse investor would like this probability (of simultaneous loss) to be small compared to the probability p of a loss in either of the stocks. Future extreme low values of the two stocks are said to be asymptotically independent if $F(x_p, y_p) = o(p)$ for $p \rightarrow 0$. Asymptotic dependence will increase risk for a portfolio containing these two stocks.

A different example of risk is presented by the yearly maxima for high water levels at particular points on the coast of Holland, Great Britain, and the US, say IJmuiden, Harwich, and New Orleans. Consider a data set of 200 observations stretching back to the beginning of the nineteenth century. We assume that the data have been standardized to offset tidal effects. Now, for each of these locations, pick the five largest values. This yields three subsets of five elements in the set of 200 years. We would suppose that there is considerable overlap between the five years selected for IJmuiden and for Harwich, since in both cases the cause is the same, a North Western storm in the North Atlantic forcing water into the funnel formed by the West coast of Holland and Belgium and the South coast of Great Britain, and opening into the Southern Atlantic via the Channel. High water levels in New Orleans have a different cause. So sea levels in New Orleans and in IJmuiden should be asymptotically independent, and we would expect the corresponding five point subsets to be disjoint with high probability. In general, given a positive integer k and a sequence of independent observations $\mathbf{Z}_1, \mathbf{Z}_2, \dots$ from a bivariate DF F , we could look in the sample cloud of the first n points $\mathbf{Z}_i = (X_i, Y_i)$ at the k largest observations for the coordinate X_i and for the coordinate Y_i . This yields two subsets of the index set $\{1, \dots, n\}$. Let $p_n(k)$ denote the probability that the two subsets have a point in common. In the case of asymptotic independence we would expect that, for fixed k , $p_n(k) \rightarrow 0$ as $n \rightarrow \infty$. This is indeed the case, as shown in [11, Proposition 2].

A related way to understand extremal dependence is by looking at the probability of a *record*, $p_n(1)$, the probability that the coordinatewise maximum of a sample of n points is given by one of the sample points. It was shown in [12, Theorem 2] that this probability vanishes as $n \rightarrow \infty$ exactly when the underlying vector has asymptotically independent components. For the sake of completeness, we give an alternative proof of this basic result in Appendix A; see Proposition A.1.

2.2. Sibuya’s condition

In his seminal paper [25] on multivariate extremes, Sibuya showed that the components of bivariate vectors with normal densities, whatever their correlation, are asymptotically independent. For a vector $Z = (X, Y)$ with DF F and continuous marginals F_1 and F_2 , he introduced a function P by

$$P(F_1(x), F_2(y)) = \Pr\{X > x, Y > y\} = 1 + F(x, y) - F_1(x) - F_2(y).$$

The function P is well defined and continuous on the unit square. We can now give *Sibuya’s condition* for asymptotic independence (see Theorem 2 of [25]).

Definition. Let $Z = (X, Y)$ have DF F with continuous marginals. The components X and Y are *asymptotically independent* if the function P introduced above satisfies

$$P(1 - s, 1 - s) = o(s), \quad s > 0, s \rightarrow 0. \tag{2.1}$$

By an abuse of language we also say that the vector Z or the DF F is *asymptotically independent*.

Independence of a bivariate vector is not affected by the marginal distributions, and similarly for asymptotic independence. It is preserved under coordinatewise increasing transformations. Since we assume that the marginals F_1 and F_2 are continuous, there is a unique function C on the unit square such that $F(x, y) = C(F_1(x), F_2(y))$. The function C is known as the *copula* of the DF F . It is a DF on the unit square with uniform marginals. Sibuya’s condition is a condition on the copula C since $P(u, v) = 1 + C(u, v) - u - v$ for $u, v \in [0, 1]$. If the condition holds for a vector (X, Y) with a continuous DF F , it automatically holds for any vector whose DF is continuous and has the same copula as F .

In this paper we consider asymptotic independence for maxima. For minima, we would define asymptotic independence similarly, in terms of the copula, by $C(s, s) = o(s)$ for $s \downarrow 0$.

Proposition 2.1. *Suppose that there exist $a, b > 0$ such that*

$$\frac{P(1 - as, 1 - bs)}{s} \rightarrow 0, \quad s \downarrow 0. \tag{2.2}$$

Then asymptotic independence holds, and relation (2.2) is valid for all $a, b > 0$.

Proof. We may assume that $a \leq b$ by symmetry and $a = 1$ replacing s by as in the denominator. By monotonicity, Sibuya’s condition holds. A similar argument gives (2.2) for any positive a, b .

For a vector (X, Y) with marginal DFs F_1 and F_2 , Sibuya’s condition may be formulated in terms of conditional quantile exceedances as

$$\lambda_U(X, Y) := \lim_{q \uparrow 1} \Pr\{X > F_1^{\leftarrow}(q) \mid Y > F_2^{\leftarrow}(q)\} = 0, \tag{2.3}$$

where $F^{\leftarrow}(q) := \inf\{x \in \mathbb{R} \mid F(x) \geq q\}$ for $q \in (0, 1)$ denotes the (minimal) q -quantile of F . The limit $\lambda_U \in [0, 1]$, if it exists, is known as the *upper tail dependence coefficient*. So X and Y are asymptotically independent if and only if their upper tail dependence coefficient is 0.

Sibuya’s condition is simple, but the formulation in terms of survival probabilities is inconvenient. The quantiles in (2.3) may be hard to determine since this amounts to computing the inverse of the DFs. There is a simple criterion in terms of sums.

Proposition 2.2. *Let (X_1, X_2) have a DF with continuous marginals F_1 and F_2 , and let $1 - F_i(t_{in}) = 1/n$. If $s_n := n\Pr\{X_1 + X_2 > t_{1n} + t_{2n}\} \rightarrow 0$ then X_1 and X_2 are asymptotically independent.*

Proof. Sibuya’s function P satisfies $nP(1 - 1/n, 1 - 1/n) = n\Pr\{X_1 > t_{1n}, X_2 > t_{2n}\} \leq s_n \rightarrow 0$. This gives Sibuya’s condition (2.1).

Below we give criteria in terms of continuous curves $\mathbf{x}(t) = (x_1(t), x_2(t))$, $t \geq 0$, for which $F_1(x_1(t))$ and $F_2(x_2(t))$ tend to 1 as $t \rightarrow \infty$.

Proposition 2.3. *Let (X_1, X_2) have DF F with continuous marginals F_1 and F_2 . The components X_1 and X_2 are asymptotically independent if and only if, for any $\varepsilon > 0$, there exists a continuous curve $\mathbf{x}(t)$, $t \geq 0$, such that $p_i(t) = \Pr\{X_i > x_i(t)\}$ is positive and vanishes for $t \rightarrow \infty$ for $i = 1, 2$, and such that*

$$\frac{\Pr\{X_1 > x_1(t), X_2 > x_2(t)\}}{p_i(t)} < \varepsilon, \quad t > t_\varepsilon, i = 1, 2. \tag{2.4}$$

Proof. Assume asymptotic independence. We may choose $x_1(t)$ and $x_2(t)$ continuous and increasing such that $p_1(t) = p_2(t)$ for all t . Then (2.1) gives (2.4). Now assume that (2.4) holds. Let $u \in (0, p_0]$, where $p_0 = \min\{p_1(t_0), p_2(t_0)\}$. By symmetry we may assume that $p_1(t) = u \leq p_2(t)$. Then

$$P(1 - u, 1 - u) \leq \Pr\{X_1 > x_1(t), X_2 > x_2(t)\} \leq \varepsilon p_1(t) = \varepsilon u.$$

This holds for all $u \in (0, p_0]$. So Sibuya’s condition is satisfied.

Asymptotic independence is preserved under quite severe deformations of the distribution.

Corollary 2.1. *Let C be a convex open cone in \mathbb{R}^d , and let f and g be probability densities which are positive on C , vanish off C , and for which the quotients f/g and g/f are bounded on C . Let $\mathbf{Z} = (Z_1, \dots, Z_d)$ have density f , and let $\mathbf{X} = (X_1, \dots, X_d)$ have density g . If Z_1 and Z_2 are asymptotically independent then so are X_1 and X_2 .*

Proof. The inequality $f \leq Mg$ implies by integration that the same inequalities hold for the univariate and bivariate marginal densities and tail functions. So $\Pr\{X_1 > x_1(t), X_2 > x_2(t)\} \leq M\Pr\{Z_1 > x_1(t), Z_2 > x_2(t)\}$. Similarly, the inequality $g \leq Mf$ gives an inequality with the constant M for the tail probabilities: $\Pr\{Z_i > x_i(t)\} \leq M\Pr\{X_i > x_i(t)\}$ for $i = 1, 2$. Now use (2.4) for (Z_1, Z_2) to establish the same relation for (X_1, X_2) . The extra factor M^2 has no effect in the limit.

The concept of asymptotic independence has been refined by looking at the rate at which $P(1 - s, 1 - s)$ vanishes for $s \rightarrow 0$, or, more generally, by looking at the behaviour of the survival function $P(1 - u, 1 - v)$ for $u, v \rightarrow 0$. See [22] or [24]. This second-order theory is called hidden regular variation. We shall not treat this subject in our paper. Since our interest is in multivariate densities rather than DFs, we give conditions on the densities which ensure asymptotic independence of the components. So the assumption of continuous DFs is not restrictive. We shall also assume that the densities are continuous on a convex cone and vanish outside this cone. In our context, this cone will typically be either the whole space \mathbb{R}^d or the open positive orthant $(0, \infty)^d$.

2.3. Asymptotic independence and multivariate extreme value theory

Asymptotic independence has to do with extremes, more precisely with bivariate maxima. Knowledge of multivariate extreme value theory is not indispensable for understanding asymptotic independence, but it will help to better understand asymptotic dependence.

For multivariate extreme value theory, we have to assume that each of the marginals of the multivariate DF F lies in the domain of attraction of a univariate extreme value limit law (see, e.g. Section 0.3 of [23] for a definition). As above, let $\mathbf{Z}_1, \mathbf{Z}_2, \dots$ with $\mathbf{Z}_n = (X_n, Y_n)$ be independent observations from the bivariate DF F , and write $\mathbf{Z}^{\vee n}$ for the n th coordinatewise partial maximum. For simplicity, assume that the marginals are equal with the standard Fréchet DF $F_i(t) = H(t) = e^{-1/t}$ on $(0, \infty)$. The normalized marginal maxima $U_n = X^{\vee n}/n$ and $V_n = Y^{\vee n}/n$ again have the Fréchet DF H by the scaling property $H^n(nt) = H(t)$. The scaled bivariate maximum $\mathbf{W}_n = \mathbf{Z}^{\vee n}/n$ has DF $G_n(\mathbf{w}) = F^n(n\mathbf{w})$. Suppose that G_n converges weakly to a limit distribution G , known as a multivariate extreme value distribution or a *max-stable* distribution. The limit vector $\mathbf{W} = (U, V)$ lives on $(0, \infty)^2$, and the components U and V have a Fréchet law. Asymptotic independence for the DF F is equivalent to independence of the components of the limit vector (see, e.g. Proposition 5.27 of [23] or Theorem 6.2.3 of [7]).

The convergence $F^n(n\mathbf{w}) \rightarrow G(\mathbf{w})$ becomes easier to handle if we take logarithms. Write $G(\mathbf{w}) = e^{-R(\mathbf{w})}$. Since $-\log F(n\mathbf{w})$ is asymptotic to $1 - F(n\mathbf{w})$ for $\mathbf{w} > \mathbf{0}$, we may write the limit relation as

$$n(1 - F(n\mathbf{w})) \rightarrow R(\mathbf{w}), \quad \mathbf{w} \in \mathcal{X} = [0, \infty)^2 \setminus \{(0, 0)\}.$$

The left-hand side $R_n(\mathbf{w}) = n(1 - F(n\mathbf{w}))$ is the DF of the measure $\rho_n = n\pi_n$, where π_n is the probability distribution of the vector \mathbf{Z}/n . So ρ_n is the mean measure of the sample cloud $N_n = \{\mathbf{Z}_1/n, \dots, \mathbf{Z}_n/n\}$ and $R_n(\mathbf{w}) = \rho_n([\mathbf{0}, \mathbf{w}]^c)$. The pointwise convergence $G_n \rightarrow G$ on $[0, \infty)^2$ implies the pointwise convergence $R_n \rightarrow R$ on \mathcal{X} and the vague convergence $\rho_n \rightarrow \rho$. One can prove the weak convergence $\rho_n \rightarrow \rho$ on $\mathcal{X} \setminus [\mathbf{0}, \mathbf{w}]$ for any $\mathbf{w} \in (0, \infty)^2$. It follows that the scaled sample clouds N_n converge in distribution to a Poisson point process N on \mathcal{X} weakly on the complement of centred disks. The Poisson point process N is of interest since it gives an asymptotic description of the large vectors in the sample cloud. Moreover, the coordinatewise maximum \mathbf{W}_n of the scaled sample cloud N_n converges in distribution to the coordinatewise maximum \mathbf{W} of N as the sample size goes to infinity.

Sibuya's condition holds precisely if $R(u, v) = 1/u + 1/v$. (Indeed, $n(1 - F_i(nt)) \rightarrow 1/t$ for $i = 1, 2$, and the survival function $\bar{F}(x, y) = \Pr\{X > x, Y > y\} = P(F_1(x), F_2(y))$ satisfies $n\bar{F}(n\mathbf{w}) \rightarrow 0$ by Sibuya's condition and Proposition 2.1.) In this case the measure ρ has density 0 on the open quadrant $(0, \infty)^2$ by differentiation (see also Proposition 5.24 of [23]). So ρ is the sum of two measures, one on the positive horizontal axis and one on the positive vertical axis, both with density $1/t^2$. The point process N is thus the superposition of two Poisson point processes on these half-axes, and the two point processes are independent since the half-axes are disjoint. This yields a simple description of the behaviour of large sample clouds: asymptotically, there is no relation between very large observations in the horizontal direction and very large observations in the vertical direction.

So far, we have looked at asymptotic independence for bivariate distributions. Unlike independence, for a multivariate DF, asymptotic independence holds if it holds for the bivariate marginals; see, e.g. Remark 6.2.5 of [7]. Here is the argument. For simplicity, assume that F has standard Fréchet marginals, and that the multivariate maxima converge in distribution. The scaled sample clouds $N_n = \{\mathbf{Z}_1/n, \dots, \mathbf{Z}_n/n\}$ then converge in distribution to a Poisson point process N on $\mathcal{X} = [0, \infty)^d \setminus \{\mathbf{0}\}$ with mean measure ρ weakly on the complement of

any centred ball. If Sibuya’s condition holds for the bivariate marginals then the projection ρ_{ij} of ρ on the (x_i, x_j) -plane will live on the two positive half-axes in this plane. It follows that ρ lives on the set of points in $[0, \infty)^d \setminus \{\mathbf{0}\}$ which have at most one positive coordinate, the union of the d positive half-axes. The restrictions of N to these half-axes are independent (since the d positive half-axes are disjoint). Hence, the multivariate extreme value limit vector has independent components.

2.4. Asymptotic dependence

Asymptotic dependence is an ambiguous term. Logically, it means the absence of asymptotic independence. We shall usually interpret it in a more constructive manner to mean the existence of a max-stable limit law $G = e^{-R}$. This in turn implies convergence in distribution of the normalized sample clouds to a Poisson point process N , whose mean measure ρ has DF $R = -\log G$, as sketched above. The tail dependence coefficient λ_U in (2.3) gives only a very restricted view of extremal dependence. As pointed out in Section 8.2 of [20], a positive coefficient does not imply that the underlying distribution belongs to the maximum domain of attraction of some extreme value limit law. The exponent measure ρ provides a much more informative description of asymptotic dependence. Points of N which do not lie on one of the axes denote a simultaneous occurrence of very large values in two or more coordinates in the corresponding point of the normalized sample. Thus, the point process N (or, equivalently, ρ) gives a complete description of how the extreme upper order statistics in the different coordinates are linked.

3. Densities and level sets

The aim of this section is to give conditions in terms of the density which will guarantee asymptotic independence. Two aspects of the density play an important role in our analysis: the shape of the level sets and the tail behaviour. We consider densities which are completely specified by just these two quantities—a shape for the level sets and a decreasing function governing the rate of decay of the tails. In many cases of practical interest the shape is a bounded open convex set, containing the origin. The density then is continuous if and only if the decreasing function is continuous. We shall also consider such unimodal densities which vanish outside the positive orthant.

3.1. Homothetic densities

Homothetic densities are densities whose level sets are scaled copies of a given open set D :

$$\{f > c\} := \{\mathbf{x} \in \mathbb{R}^d \mid f(\mathbf{x}) > c\} = r_c D, \quad 0 < c < c_0 := \sup f.$$

We assume that $D \subset \mathbb{R}^d$ is a bounded open star-shaped set. A *star-shaped set* has the property that, with any point \mathbf{x} , it contains all points $r\mathbf{x}$ for $0 < r < 1$. Assume that the set D contains the origin. If each ray intersects the boundary ∂D at one point, then the set D can be represented using a *gauge function* $n_D: \mathbb{R}^d \rightarrow [0, \infty)$ that satisfies (i) $n_D(t\mathbf{x}) = tn_D(\mathbf{x})$ for $t > 0, \mathbf{x} \in \mathbb{R}^d$ (homogeneity property), and (ii) $D = \{\mathbf{x} \in \mathbb{R}^d \mid n_D(\mathbf{x}) < 1\}$. The conditions on the set D and the continuity of f ensure that the gauge function is continuous. If the set D is convex then so is the gauge function. If, in addition, the set is symmetric, $-D = D$, then the gauge function is a norm on \mathbb{R}^d , and the set D is the open unit ball in this norm. For any bounded open star-shaped set D , the sets $nD, n > 0$, form an increasing family. Their union D_∞ is an open cone. It is a proper cone if the origin is a boundary point of D . We may then still define the gauge function n_D . This is now a function on the open cone D_∞ . It is continuous on the

cone if each ray in the cone intersects the boundary of D at a single point; see Proposition 3.2, below.

Gauge functions allow us to give analytic expressions for homothetic densities with the same ease with which one handles spherically symmetric densities. Continuous homothetic densities have the form $f(\mathbf{x}) = f_0(n_D(\mathbf{x}))$ for a decreasing continuous function f_0 on $(0, \infty)$. This formula holds for all \mathbf{x} for which the ray through \mathbf{x} intersects the set D .

Definition 3.1. Let \mathcal{D}_d denote the class of all bounded open star-shaped sets $D \subset \mathbb{R}^d$ for which the cone $D_\infty = \bigcup_n nD$ is convex, and for which the gauge function n_D is continuous on this cone. A density f on \mathbb{R}^d belongs to the class $\mathcal{H}(D)$ if the shape D belongs to the class \mathcal{D}_d , and if f is of the form $f(\mathbf{x}) = f_0(n_D(\mathbf{x}))$, where the density generator $f_0: [0, \infty) \rightarrow [0, \infty)$ is decreasing, positive, and continuous. We set $f \equiv 0$ outside the cone D_∞ on which the gauge function is defined.

It is apparent from the above definition that densities in \mathcal{H} are (*star*) unimodal since all the level sets $\{f > c\}$, $c \in (0, c_0)$, are star shaped; cf. Section 2.2 of [8]. Typical examples of densities in \mathcal{H} to keep in mind are the multivariate centred normal densities and, more generally, elliptically symmetric densities, discussed, for example, in [9]. See also Example A.1. In certain applications, elliptical symmetry may be too restrictive. Densities in \mathcal{H} give the flexibility to model directional irregularities present in the data clouds, and to handle distributions on the positive orthant. The regularity conditions exclude pathological sets; see Example A.2.

Before we proceed to looking at the properties of these densities, let us review some related classes of models proposed in the literature. The ℓ_p -spherical densities (see [21]) extend the class of spherical densities by allowing level sets to be balls in the ℓ_p -norm for any $p \geq 1$. A further generalization is given by so-called v -spherical densities (see [10]), where the scale function v plays the same role as the gauge function n_D defined above. In fact, our class $\mathcal{H}(D)$ is a subclass of the v -spherical densities in that we restrict level sets to be bounded and star shaped, and f to be continuous on D_∞ . In a recent paper [1], the authors advocated studying densities in terms of their contours.

For a given shape D , what conditions does the density generator f_0 have to satisfy in order that the function $f(\cdot) = f_0(n_D(\cdot))$ is a probability density on \mathbb{R}^d ? By regarding the set below the graph of f as a pile of thin D -shaped slices we obtain the following partial integration result:

$$\begin{aligned} \Pr\{X \in tD\} &= \int_{tD} f_0(n_D(\mathbf{x})) \, d\mathbf{x} \\ &= f_0(t)|tD| + |D| \int_0^t s^d |df_0(s)| \\ &= \int_0^t f_0(s) \, d|sD|, \quad t > 0. \end{aligned}$$

Observing that $|sD| = s^d|D|$ and letting t tend to ∞ , we obtain the condition (cf. Equation (5) of [10]) $1 = d|D| \int_0^\infty s^{d-1} f_0(s) \, ds$.

The class of unimodal densities introduced above is invariant under linear transformations. If the vector X has density $f \in \mathcal{H}(D)$ then the vector $Y = AX$ has density $g \in \mathcal{H}(E)$, where E is the image of the star-shaped (convex) set D under the linear transformation A , and, hence, is also star shaped (convex). A nice illustration of this invariance is the extension of spherical distributions to elliptical ones.

For densities in \mathcal{H} , the distinction between light and heavy tails is crucial for asymptotic independence. It is determined by the behaviour of the generator f_0 at ∞ . Let us recall the definitions of regular and rapid variation.

Definition 3.2. A measurable function $h: (0, \infty) \rightarrow (0, \infty)$ is *regularly varying* at ∞ with exponent θ if, for $x > 0$,

$$\lim_{t \rightarrow \infty} \frac{h(tx)}{h(t)} = x^\theta, \quad \theta \in \mathbb{R}.$$

If $\theta = 0$ then h is called *slowly varying*; h is *rapidly varying* at ∞ if

$$\lim_{t \rightarrow \infty} \frac{h(tx)}{h(t)} = \begin{cases} \infty, & 0 < x < 1, \\ 0, & x > 1. \end{cases}$$

If the generator of a density $f = f_0(n_D)$ varies rapidly, the density has *light tails*. Our concern is with light-tailed densities for which the generator f_0 is continuous, positive, and strictly decreasing on $[0, \infty)$.

Rapid variation of f_0 allows us to give strong inequalities for the measure μ with density $f = f_0(n_D)$.

Proposition 3.1. Let μ have density $f \in \mathcal{H}(D)$ with a rapidly varying density generator f_0 .

(i) For any $\varepsilon > 0$,

$$\mu(rD^c) \ll \mu(rD \setminus e^{-\varepsilon}rD), \quad r \rightarrow \infty. \tag{3.1}$$

(ii) For any nonempty open set $U \subset D$,

$$\mu(rD^c) \ll \mu(rU), \quad r \rightarrow \infty. \tag{3.2}$$

Proof. Rapid variation implies that $f_0(e^\delta t) < f_0(t)/M$ eventually for $t \rightarrow \infty$, and, hence, the rings $R_n = e^{n\delta}tD \setminus e^{(n-1)\delta}tD$, $n \geq 0$, have measure $\mu(R_{n+2}) \leq e^{2d\delta} \mu(R_n)/M$. The rings R_n are disjoint, and their union is the complement of tD . On summing the odd and the even terms, we find that, with $\eta = e^{2d\delta}/(M - e^{2d\delta})$,

$$\mu(e^{2\delta}tD^c) = \sum_{n=1}^{\infty} (\mu(R_{2n}) + \mu(R_{2n+1})) \leq \eta(\mu(R_0) + \mu(R_1)) = \eta\mu(e^{2\delta}tD \setminus tD).$$

This gives (3.1). The integral over a thin ring is much larger than the integral over the set outside the ring. Formally, take $\delta = \varepsilon/2$ and $r = t - \varepsilon$ to obtain (3.1). To prove (3.2), take a nonempty open subset $U_0 \subset U$ whose closure lies in D . Then $e^{2\varepsilon}U_0 \subset D$ if ε is small, and, hence, by rapid variation of f_0 as above, the infimum of f over rU_0 is much larger than the supremum of f over the ring $R = rD \setminus e^{-\varepsilon}rD$. Since $|R|/|rU_0|$ is a constant, we conclude that $\mu(R) \ll \mu(rU_0)$, and (3.2) follows from (3.1).

We now give some extra details on star-shaped sets.

Proposition 3.2. Let D be a bounded open star-shaped set. Suppose that, for each nonzero vector $x \in D$, there is one positive real r_0 such that r_0x lies on the boundary of D . Then the gauge function n_D is continuous on the cone $D_\infty = \bigcup_n nD$.

Proof. Let B denote an open unit ball. If D contains the origin, it contains a ball εB , and, by homogeneity, $n_D < \delta$ on $\delta\varepsilon B$. So the gauge function is then continuous at the origin. Now suppose that it is discontinuous at a point \mathbf{p} outside the origin. We may assume that $n_D(\mathbf{p}) = 1$ and that there is a sequence $\mathbf{p}_n \in D_\infty$ such that $\mathbf{p}_n \rightarrow \mathbf{p}$ and $n_D(\mathbf{p}_n) \rightarrow c \neq 1$. Let $r \in (0, 1)$ be close to 1. Then $n_D(r\mathbf{p}_n) \rightarrow rc$, and $r\mathbf{p} \in D$ implies that $r\mathbf{p}_n \in D$ eventually; hence, $rc \leq 1$ and so $c < 1$ since r is arbitrary. This implies that $\mathbf{p}_n \in D$ eventually, and this also holds for $s\mathbf{p}_n$ for $1 < s < 1/c$. Hence, $s\mathbf{p}$ is a boundary point. But so is \mathbf{p} . This contradicts our assumption on the boundary of D .

In general, densities in $\mathcal{H}(D)$ are not closed under projection, even if D is convex. If a random vector (X_1, \dots, X_d) has density $f \in \mathcal{H}(D)$, then the density of (X_1, \dots, X_{d-1}) need not be homothetic, and the univariate marginals need not even be unimodal; see Example A.3. There are some exceptions. Projections of spherical densities are spherical, and if the level sets are balls in ℓ_p for some $p \in [1, \infty]$ then this also holds for projections along the coordinate axes. The class \mathcal{D}_d is itself closed under projection. If D is a bounded open convex set containing the origin then so is the vertical projection E of D onto the horizontal hyperplane; if the origin is a boundary point of D , it may be an interior point of E .

Proposition 3.3. *Suppose that $D \in \mathcal{D}_d$. Let E be the vertical projection of D onto the horizontal hyperplane. Then $E \in \mathcal{D}_{d-1}$.*

Proof. Write $\mathbf{z} = (\mathbf{x}, y)$ to distinguish the horizontal and vertical parts of the vector \mathbf{z} . It is clear that E is a bounded open star-shaped set. Moreover E_∞ is the projection of the cone D_∞ , and, hence, an open convex cone. We have to prove continuity of the gauge function n_E on E_∞ . If n_E is not continuous, there exists a vector \mathbf{x} and a sequence $\mathbf{x}_n \rightarrow \mathbf{x}$ such that $n_E(\mathbf{x}) = 1$ and $n_E(\mathbf{x}_n) < c_0 < 1$ because E is open. Hence, $n_D(\mathbf{x}, y) \geq 1$ for all y and there exist y_n such that $n_D(\mathbf{x}_n, y_n) < c_0$. Since D is bounded, the sequence (y_n) is bounded, and we may assume that it converges to some element y_0 . Then $(\mathbf{x}_n, y_n) \rightarrow (\mathbf{x}, y_0)$, and continuity of n_D implies that $n_D(\mathbf{x}, y_0) \leq c_0 < 1$. This is a contradiction.

3.2. Densities whose level sets are asymptotically star shaped

In this section we relax the condition that all level sets have the same shape $D \in \mathcal{D}_d$ to the condition that the level sets, properly scaled, converge to a set $D \in \mathcal{D}_d$. We restrict attention to the light-tailed setting.

Definition 3.3. Let $D \in \mathcal{D}_d$. A positive probability density f on \mathbb{R}^d belongs to the set $\mathcal{A}(D)$ if there exist sequences $c_n > 0$ and $r_n \rightarrow \infty$ with $c_{n+1}/c_n \rightarrow 0$ and $r_{n+1} \sim r_n$ such that, for any $\varepsilon > 0$, eventually

$$e^{-\varepsilon} r_n D \subset \{f > c_n\} \subset e^\varepsilon r_n D, \quad n \geq n_0. \tag{3.3}$$

We write $\{f > c_n\}/r_n \rightarrow D$. A continuous positive function \tilde{f} is *shape equivalent* to f if its level sets satisfy (3.3).

The sequences r_n and c_n determine a set of continuous decreasing functions η which satisfy $\eta(c_n) \sim r_n$. All these functions $\eta(c)$ vary slowly for $c \rightarrow 0^+$ by the assumption that $c_{n+1}/c_n \rightarrow 0$ and $r_{n+1}/r_n \rightarrow 1$. It is this set of slowly varying functions rather than the particular sequences c_n and r_n which is of interest.

Proposition 3.4. *If the slowly varying function η above is strictly decreasing, defined on $(0, c_\eta]$ for some $c_\eta > 0$, and vanishes in c_η , then the inverse function $g_0 = \eta^\leftarrow$ is a continuous, positive,*

strictly decreasing function on $[0, \infty)$ which varies rapidly at ∞ , and $g = g_0(n_D)$ is shape equivalent to f , and, for all $\varepsilon > 0$,

$$e^{-\varepsilon}\{g > c\} \subset \{f > c\} \subset e^{\varepsilon}\{g > c\}, \quad 0 < c < c_\varepsilon. \tag{3.4}$$

Proof. Rapid variation of the inverse function g_0 follows from Theorem 2.4.7(i) of [6]. Shape equivalence holds since $\{g > c\} = \eta(c)D$ and $\eta(c_n) \sim r_n$ by assumption.

Since $r_n \rightarrow \infty$, we may take a strictly increasing subsequence such that the asymptotic equality $r_{n+1} \sim r_n$ remains valid. Take $c_0 > c_1$. Any continuous strictly decreasing function η on $(0, c_0]$ which vanishes at c_0 and has the value r_n at c_n satisfies the conditions of the proposition above. So there are many continuous, strictly decreasing functions g_0 on $[0, \infty)$ which vary rapidly at ∞ such that $g_0(n_D)$ is shape equivalent to f .

For functions in $\mathcal{A}(D)$, the inequalities in Proposition 3.1 also hold. Moreover, they have the nice property that the d marginals g_i of $g \in \mathcal{A}(D)$ will lie in $\mathcal{A}(D_i)$, where D_i is the projection of D on the i th coordinate.

Lemma 3.1. *Let $g \in \mathcal{A}(D)$. Let ρ_r for $r \geq 1$ denote the measure with density $\mathbf{u} \mapsto g(r\mathbf{u})/g(r\mathbf{q})$ for a fixed nonzero vector \mathbf{q} . Then, for any open set U which intersects D ,*

$$\rho_r(D^c) \ll \rho_r(U). \tag{3.5}$$

Proof. Let $f = f_0(n_D)$ be shape equivalent to g , and define μ_r to have density $f_r(\mathbf{u}) = f(r\mathbf{u})/f_0(r)$. It suffices to prove (3.5) for the measures ρ_r with density $g_r(\mathbf{u}) = g(r\mathbf{u})/f_0(r)$. Let V be a nonempty open subset of U which lies in the complement of $\delta_0 D$ for some $\delta_0 > 0$, and whose closure lies in D . For $r > r_0/\delta$, we may apply the pointwise inequality (3.6), below, and conclude that $\rho_r(U) \geq e^{-\varepsilon d} \mu_r(e^\varepsilon V)$, where we choose $\varepsilon > 0$ so small that $e^{3\varepsilon} V \subset D$. Then, using (3.2), $\mu_r(e^{-\varepsilon} D^c) \ll \mu_r(e^\varepsilon V)$ and another application of the pointwise inequality gives (3.5).

Proposition 3.5. *Suppose that X has density $f \in \mathcal{A}(D)$. Let $Y = \xi(X)$ be a nonzero linear combination of the components of X . Then Y has a density $g \in \mathcal{A}(J)$, where J is the open interval $\xi(D)$.*

Proof. We may assume that ξ is the vertical coordinate. The condition $f(e^\varepsilon \mathbf{x}) < f(\mathbf{x})/M$ for $\|\mathbf{x}\| \geq r_0$ implies, by integration over horizontal hyperplanes, that $g(e^\varepsilon y) < e^{(d-1)\varepsilon} g(y)/M$ for $|y| \geq r_0$. This gives rapid variation. The thin tails of f ensure that g is continuous. Let $\xi(D) = (a, b)$. We claim that the average of g over $r(a, a + \varepsilon a)$ and over $r(b - \varepsilon b, b)$, say a_1 and b_1 , is much larger than over both $r(a - \varepsilon a, a)$ and $r(b, b + \varepsilon b)$, say a_2 and b_2 , as $r \rightarrow \infty$. It suffices to show that $\Pr\{ra < Y < ra + \varepsilon ra\}$ and $\Pr\{rb - \varepsilon rb < Y < rb\}$ are much larger than $\Pr\{X \in rD^c\}$. This follows from (3.5) since the horizontal slice $\{ra < x_d < ra + \varepsilon ra\}$ contains the open set $U = rD \cap \{x_d < ra + \varepsilon ra\}$, and $\Pr\{X \in U\}$ is much larger than $\Pr\{X \in rD^c\}$. A similar argument holds for the strip $\{rb - \varepsilon rb < x_d < rb\}$. Let c lie between $a_1 \wedge a_2$ and $b_1 \vee b_2$. Then $g > c$ on $[ra - 2\varepsilon ra, rb + 2\varepsilon rb]$ and $g < c$ holds off $[ra - 2\varepsilon ra, rb + 2\varepsilon rb]$. Hence, $e^{-3\varepsilon} rJ \subset \{g > c\} \subset e^{3\varepsilon} rJ$ holds for small ε and sufficiently large r .

It is not known whether a similar result holds for the bivariate marginals of f .

We can also define shape equivalence for functions with heavy tails. In this case shape equivalence is the same as asymptotic equality at ∞ ; see Proposition A.2. For light tails, the behaviour of the quotient of two shape-equivalent functions may be very erratic. Example A.4 in Appendix A lists some functions which are shape equivalent to the bivariate Gaussian density.

The level set inequalities in (3.4) imply the pointwise inequalities:

$$f(e^\varepsilon \mathbf{x}) \leq g(\mathbf{x}) \leq f(e^{-\varepsilon} \mathbf{x}), \quad \mathbf{x} \in r_\varepsilon D^c. \tag{3.6}$$

This shows that a function $g = qf$ is shape equivalent to $f = f_0(n_D)$ for $q = e^\psi$ and $f_0 = e^{-\varphi}$ if, for every $\varepsilon > 0$,

$$\varphi(r) - \varphi(e^\varepsilon r) \leq \psi(\mathbf{x}) \leq \varphi(r) - \varphi(e^{-\varepsilon} r), \quad \mathbf{x} \in \partial r D, r \geq r_\varepsilon. \tag{3.7}$$

Rapid variation implies that the left-hand side goes to $-\infty$ and the right-hand side goes to ∞ . Hence we obtain the following proposition.

Proposition 3.6. *If q is positive and $\log q$ is bounded, then $g \in \mathcal{A}(D)$ is shape equivalent to $\tilde{g} = qg$.*

For Weibull-like functions of the form $f = f_0(n_D)$, where $f_0 = e^{-\varphi}$ for a continuous, strictly increasing function φ which varies regularly with positive exponent, there is a simple alternative description of shape equivalence.

Lemma 3.2. *Suppose that $f_0 = e^{-\varphi}$, where φ is continuous, strictly increasing, and varies regularly at ∞ with exponent $\theta > 0$. Then f_0 varies rapidly at ∞ and $g = e^{-\gamma} \in \mathcal{A}(D)$ is shape equivalent to $f_0(n_D)$ if and only if $\gamma(\mathbf{x}_n) \sim \varphi(n_D(\mathbf{x}_n))$ holds whenever $\|\mathbf{x}_n\| \rightarrow \infty$.*

Proof. Regular variation of φ gives $\varphi(e^\varepsilon r) - \varphi(r) \sim (e^{\theta\varepsilon} - 1)\varphi(r)$ and $\varphi(r) - \varphi(e^{-\varepsilon} r) \sim (1 - e^{-\theta\varepsilon})\varphi(r)$ for $r \rightarrow \infty$. The claim then follows from (3.7) with $\gamma = \varphi - \psi$.

Proposition 3.7. *Suppose that $g = e^{-\gamma}$ is a continuous positive density on \mathbb{R}^d . Let there exist a function v on \mathbb{R}^d which is positive outside a bounded set, and a nonzero vector \mathbf{q} such that*

$$\frac{\gamma(t_n \mathbf{u}_n)}{\gamma(t_n \mathbf{q})} \rightarrow v(\mathbf{u}), \quad t_n \rightarrow \infty, \mathbf{u}_n \rightarrow \mathbf{u}, \mathbf{u} \in \mathbb{R}^d. \tag{3.8}$$

Then there exist a set $D \in \mathcal{D}_d$ containing the origin, a positive constant θ such that $v = n_D^\theta$, and a continuous, strictly increasing function φ on $[0, \infty)$ which varies regularly with exponent θ at ∞ such that γ is asymptotic to $\varphi(n_D)$ at ∞ , and g is shape equivalent to $f_0(n_D)$, where $f_0 = e^{-\varphi}$.

Proof. Set $a(t) = \gamma(t\mathbf{q})$. Then a varies regularly at ∞ with exponent $\theta \geq 0$. For any unit vector ω , the function $t \mapsto \gamma(t\omega)$ varies regularly at ∞ since $\gamma(t\omega_n)/a(t)$ has a positive limit $v(\omega)$ for $\omega_n \rightarrow \omega$ and $t \rightarrow \infty$. Uniform convergence on compact sets implies that v is continuous on the unit sphere, and $a(st)/a(t) \rightarrow s^\theta$ implies that $\gamma(st\omega)/a(t) \rightarrow v(\omega)s^\theta$. This proves that $v(r\omega) = r^\theta v(\omega)$, and $\gamma(r\omega) \sim a(r)v(\omega) \asymp a(r)$ for $\|r\omega\| \rightarrow \infty$ since (3.8) implies uniform convergence on compact sets, and v is continuous and, hence, bounded on compact sets. Hence, $a(r) \rightarrow \infty$ and (3.8) with $\mathbf{u}_n = \mathbf{0}$ gives $v(\mathbf{0}) = 0$, and continuity of v implies that $\theta > 0$. Set $D = \{v < 1\}$. Then $v = n_D^\theta$. Let $\mathbf{p} \in \partial D$. Then $t \mapsto \gamma(t\mathbf{p})$ varies regularly with exponent θ , and we may choose φ positive, continuous, and strictly increasing, and asymptotic to this function for $t \rightarrow \infty$. Then $\gamma(r\omega) \sim a(r)v(\omega)$ gives $\gamma(\mathbf{w}) \sim \varphi(n_D(\mathbf{w}))$ for $\|\mathbf{w}\| \rightarrow \infty$. Lemma 3.2 then shows that g is shape equivalent to $e^{-\varphi(n_D)}$, and, since $e^{-\varphi}$ varies rapidly, if φ varies regularly with exponent $\theta > 0$, it follows that $g \in \mathcal{A}(D)$.

Remark 3.1. Condition (3.8) is related to multivariate regular variation; see, e.g. Section 5.4.2 of [23].

3.3. Criteria for asymptotic independence

A random vector X with a spherically symmetric density $f(\mathbf{x}) = f_0(\|\mathbf{x}\|_2)$ will have asymptotically independent components if the generator f_0 varies rapidly, and it will have asymptotically dependent components if f_0 varies regularly (see, e.g. Theorem 4.3 of [14], or Proposition 3.2 of [13]). These results remain valid for $f \in \mathcal{A}(B)$, where B denotes the open unit ball. We may replace B by certain bounded open star-shaped sets D , as will be shown in Theorems 3.2 and 3.3, below.

With any bounded open set D we may associate the open intervals $D_i = (a_i, b_i)$, $i = 1, \dots, d$, obtained by projecting D onto the i th coordinate. Then (\mathbf{a}, \mathbf{b}) is the smallest open box containing D , and $\mathbf{b} = (b_1, \dots, b_d) = \sup D$ is the coordinatewise supremum of all points in D (and $\mathbf{a} = (a_1, \dots, a_d) = \inf D$).

Asymptotic independence depends on the bivariate marginal distributions. Hence, we introduce the projections of D on the (x_i, x_j) coordinates. For $1 \leq i < j \leq d$, we denote by $D_{ij} \subset \mathbb{R}^2$ the projection of D onto the two-dimensional space spanned by the unit base vectors \mathbf{e}_i and \mathbf{e}_j . The sets D_{ij} lie in \mathcal{D}_2 by Proposition 3.3, and D_{ij} fits exactly into the rectangle $(a_i, b_i) \times (a_j, b_j)$. The cone generated by D_{ij} is the projection of the cone D_∞ generated by D .

Definition 3.4. The set $D \in \mathcal{D}_2$ is *blunt* if the point $(b_1, b_2) = \sup D$ does not lie in the closure of D .

A bounded open convex set D in \mathbb{R}^d is *smooth* at the boundary point \mathbf{p} if there is a unique hyperplane which contains \mathbf{p} , but which does not intersect D , the tangent plane to D at \mathbf{p} . For a planar set D , this means that \mathbf{p} is not a vertex. If the convex hull of D is smooth at all points then all bivariate projections D_{ij} are blunt.

We can now state our main results for asymptotic independence in terms of densities.

Theorem 3.1. *If X has a light-tailed homothetic density $f \in \mathcal{H}(D)$, and D is convex with a smooth boundary, then, for any two distinct unit vectors \mathbf{a} and \mathbf{b} , the random variables $\mathbf{a}^\top X$ and $\mathbf{b}^\top X$ are asymptotically independent. The result remains valid if the density of X is in $\mathcal{A}(D)$.*

Proof. First assume that \mathbf{a} and \mathbf{b} are linearly independent. Introduce new coordinates such that \mathbf{a} and \mathbf{b} become the first two base vectors \mathbf{e}_1 and \mathbf{e}_2 . It suffices to check that the assumption holds for vertical tangent planes, hyperplanes which project onto a line in the two-dimensional (x_1, x_2) -plane. The characterization of D is geometrical and so it is preserved under linear transformations. The projection D_{12} in the new coordinates is also convex and smooth. Hence, D_{12} is blunt, and we may apply Theorem 3.2, below. If $\mathbf{b} = -\mathbf{a}$, the bivariate distribution lies on the counterdiagonal, $y = -x$, and asymptotic independence is trivial by applying Proposition 2.3 with $\mathbf{x}(t) = (t, t)$ for $t \geq 0$.

Theorem 3.2. *Suppose that X has density $g \in \mathcal{A}(D)$. If the bivariate projection D_{12} is blunt then X_1 and X_2 are asymptotically independent.*

Proof. There is a simple analytic argument. Let $\sup D_{12} = (b_1, b_2)$. The sum $Y = X_1 + X_2$ has density $g \in \mathcal{A}(J)$ by Proposition 3.5, where J has upper endpoint $\mathbf{b} = \sup\{x + y \mid (x, y) \in D_{12}\}$ and $b < b_1 + b_2$ by bluntness. Now use Sibuya’s condition for sums in Proposition 2.2. We give a more probabilistic proof in Section 4.

3.4. Criteria for asymptotic dependence

We now give the counterpart to Theorem 3.2 for heavy-tailed densities. We are interested in the case where the partial maxima go to ∞ for all coordinates. If the shape D lies in the negative orthant then the coordinatewise maxima converge to $\mathbf{0}$; if it lies in a negative coordinate half-space $\{x_i < 0\}$, all partial maxima will lie in this half-space too. We exclude these cases in the theorem below.

Theorem 3.3. *Suppose that X has density $f \sim f_0(n_D)$ with $D \in \mathcal{D}_d$ and that f_0 is a continuous, strictly decreasing positive function on $[0, \infty)$ which varies regularly at ∞ with exponent $-(\lambda + d)$ for some $\lambda > 0$. Assume that, for each coordinate $i \in \{1, \dots, d\}$, the set D contains a point whose i th component is positive. The components of X are asymptotically dependent unless D is contained in the set S of points with at most one positive coordinate, a union of $d + 1$ orthants. The partial maxima $X^{\vee n}$ may be scaled to converge in law to a vector W whose components have DF $\Pr\{W_i \leq t\} = e^{-(a_i/t)^\lambda}$ for positive constants a_1, \dots, a_d . The exponent measure ρ^+ of W is the image under the map $\mathbf{x} \mapsto \mathbf{x}^+ = (x_1 \vee 0, \dots, x_d \vee 0)$ of the excess measure ρ with intensity $c/n_D^{\lambda+d}$. We may take $c = 1$ by a suitable choice of the scaling constants for the maxima.*

Proof. Choose a point w_0 in the cone D_∞ on the boundary of D , and, for $r \geq 1$, set $h_r(w) = f(rw)/f(rw_0) = f_0(rn_D(w))/f_0(r)$. By regular variation for $w \in D_\infty$, $w \neq \mathbf{0}$,

$$h_{r_n}(w_n) \rightarrow h(w) = n_D(w)^{-(\lambda+d)}, \quad w_n \rightarrow w, r_n \rightarrow \infty.$$

The convergence $h_r \rightarrow h$ holds uniformly on the intersection of the cone D_∞ with any ring $rB \setminus \varepsilon B$, where B denotes the open unit ball. (The function $1/n_D$ is bounded on such sets.) By Potter’s theorem (see Theorem 1.5.6 of [6]), for any $\varepsilon > 0$, there exists r_ε such that

$$\frac{f_0(rs)}{f_0(r)} \leq \frac{2s^\varepsilon}{s^{\lambda+d}}, \quad r \geq r_\varepsilon, s \geq 1.$$

This yields an integrable majorant for the convergence $h_r \rightarrow h$ on $D_\infty \setminus B$. Lebesgue’s dominated convergence theorem implies that $h_r \rightarrow h$ in L^1 on $D_\infty \setminus B$ and, because of the uniform convergence above, also on $D_\infty \setminus \varepsilon B$ for any $\varepsilon > 0$.

Let $\rho(r)$ be the finite measure with density h_r , and choose r_n so that $\rho(r_n)$ has mass n . Then $\rho(r_n)$ is the mean measure of the scaled sample cloud $N_n = \{X_1/r_n, \dots, X_n/r_n\}$, and $\rho(r_n) \rightarrow \rho$ weakly on $D_\infty \setminus \varepsilon B$ implies that $N_n \Rightarrow N$ weakly on $D_\infty \setminus \varepsilon B$, where N is the Poisson point process on D_∞ with intensity h . This tells us that the maxima converge. The measure ρ is an excess measure on $\mathbb{R}^d \setminus \{\mathbf{0}\}$:

$$\rho(rA) = \frac{\rho(A)}{r^\lambda}, \quad r > 0, A \text{ a Borel set in } \mathbb{R}^d \setminus \{\mathbf{0}\}.$$

For half-spaces $A = \{x_i \geq 1\}$ or $A = \{x_i \leq -1\}$, this relation also holds and implies that the marginals ρ_i of ρ satisfy the same relation, and, hence, there exist nonnegative constants c_i^\pm such that

$$\rho_i[r, \infty) = \frac{c_i^+}{r^\lambda}, \quad \rho_i(-\infty, -r] = \frac{c_i^-}{r^\lambda}, \quad r > 0. \tag{3.9}$$

For the coordinatewise maxima of heavy-tailed distributions, it is convenient to work on the nonnegative orthant, and replace the vector X by X^+ , where we use the continuous map $(x_1, \dots, x_d) = \mathbf{x} \mapsto \mathbf{x}^+ = (x_1 \vee 0, \dots, x_d \vee 0)$. We shall write N_n^+ and N^+ for the images of N_n and N under this map, and denote the mean measures by ρ_n^+ and ρ^+ . By assumption, each

component has a positive probability of being positive. Hence, $W_n := \max N_n^+ = \max N_n \Rightarrow \max N = \max N^+ =: W$. Write $R(\mathbf{w}) = \rho((-\infty, \mathbf{w}]^c) = \rho^+([\mathbf{0}, \mathbf{w}]^c)$, $\mathbf{0} \neq \mathbf{w} \geq \mathbf{0}$. Then the limit distribution H of the coordinatewise maxima is

$$\Pr\{W \leq \mathbf{w}\} = \Pr\{N([\mathbf{0}, \mathbf{w}]^c) = 0\} = e^{-\rho^+([\mathbf{0}, \mathbf{w}]^c)} = e^{-R(\mathbf{w})}.$$

Using (3.9), the same argument gives the marginals $H_i(t) = e^{-\rho_i[t, \infty)} = e^{-c_i^+/t^\lambda}$ with $a_i^\lambda = c_i^+$. Since \mathbf{x}^+ lies on a coordinate axis precisely if $\mathbf{x} \in S$, the exponent measure ρ^+ lives on the positive half-axes if and only if $D \subset S$.

4. Sample clouds

In this section we look at the asymptotic behaviour of clouds of independent observations from a given light-tailed distribution as the number of data points in the sample approaches infinity. In particular, we are interested in the limiting shape of these sample clouds under suitable scaling. We remark that sample clouds can be viewed as finite point processes with a fixed number of points. The motivation for looking at sample clouds is threefold. First of all, there is a relation between the asymptotic shape of the level sets of the underlying light-tailed density and the limit set onto which corresponding scaled sample clouds converge; see Proposition 4.2, below. Secondly, the point process approach will yield an intuitive proof of our main results. Finally, for sample clouds, projection onto lower-dimensional coordinate planes is simple: just delete the redundant coordinates in each sample point.

As before, we consider a sequence of independent and identically distributed random vectors $\{X_1, X_2, \dots\}$ from a continuous distribution on \mathbb{R}^d with density f . Let

$$N_n = \{X_1/s_n, \dots, X_n/s_n\}$$

denote a scaled n -point random sample (or *sample cloud*) with the scaling constant $s_n > 0$, $s_n \rightarrow \infty$ for $n \rightarrow \infty$. Alternatively, for any Borel set $A \subset \mathbb{R}^d$,

$$N_n(A) = \sum_{i=1}^n \mathbf{1}_A\left(\frac{X_i}{s_n}\right).$$

The mean measure of N_n is given by $\rho_n = n\pi_n$, where π_n is the distribution of the scaled vector X_1/s_n . The intensity of the n -point sample cloud N_n is $h_n(\mathbf{u}) = ns_n^d f(s_n \mathbf{u})$ for $\mathbf{u} \in \mathbb{R}^d$.

Under a suitable choice of s_n , the scaled observations X_i/s_n from a density $f \in \mathcal{A}(D)$ with high probability will fill out the closure of the shape set D in the sense of the following definition.

Definition 4.1. Let E be a compact set in \mathbb{R}^d , and let the μ_n be finite measures. We say that the measures μ_n converge onto E if $\mu_n(\mathbf{p} + \varepsilon B) \rightarrow \infty$ for any ε -ball centred at a point $\mathbf{p} \in E$, and if $\mu_n(U^c) \rightarrow 0$ for all open sets U containing E . The finite point processes N_n converge onto E if $\Pr\{N_n(U^c) > 0\} \rightarrow 0$ for open sets U containing E , and if $\Pr\{N_n(\mathbf{p} + \varepsilon B) > m\} \rightarrow 1$, $m \geq 1$, $\varepsilon > 0$, $\mathbf{p} \in E$.

We call the set E in the definition above the *limit set*. In fact, the limit set, if it exists, is always star shaped (see Proposition 4.1 of [15]). The following simple criterion is useful for checking convergence (in probability) of scaled sample clouds (see [5] for a proof).

Proposition 4.1. If N_n is an n -point sample cloud from a probability distribution π_n on \mathbb{R}^d , then N_n converges onto E if the mean measures $\mu_n = n\pi_n$ converge onto E .

The next theorem gives a sufficient condition for asymptotic independence of a distribution in terms of the limit set of the associated sample clouds. We first prove a lemma.

Lemma 4.1. *Let $W_n = (U_n, V_n)$ be the componentwise maximum of the sample*

$$N_n = \{Z_{n1}, \dots, Z_{nn}\}$$

from the distribution π_n on \mathbb{R}^2 . Suppose that $n\pi_n\{(0, \infty)^2\} \rightarrow 0$, $n\pi_n\{u > 0\} \rightarrow \infty$, and $n\pi_n\{v > 0\} \rightarrow \infty$. Then the probability $p_n = \Pr\{W_n \in N_n\}$ that the coordinatewise maximum is a sample point vanishes for $n \rightarrow \infty$.

Proof. Since W_n lies in the positive quadrant or $U_n \leq 0$ or $V_n \leq 0$, we find that

$$\Pr\{W_n \in N_n\} \leq \Pr\{N_n((0, \infty)^2) > 0\} + \Pr\{N_n(\{u > 0\}) = 0\} + \Pr\{N_n(\{v > 0\}) = 0\}.$$

These three binomial probabilities all vanish for $n \rightarrow \infty$.

Theorem 4.1. *Let X_1, X_2, \dots be independent observations from a continuous DF F on \mathbb{R}^2 . Let D be an open bounded star-shaped set which belongs to \mathcal{D}_2 . Suppose that there exist scaling constants s_n such that the scaled sample clouds $N_n = \{X_1/s_n, \dots, X_n/s_n\}$ converge onto the closure of D . If D is blunt, the two coordinates of the vector X_1 are asymptotically independent.*

Proof. Let $b = (b_1, b_2)$ denote the coordinatewise supremum of D . Since D is blunt, there exists a $\delta > 0$ such that the shifted quadrant $(e^{-\delta}b, \infty)$ and the set $e^\delta D$ are disjoint (see Figure 1). Let $n\pi_n$ be the mean measure of N_n . Then $n\pi_n(e^{-\delta}b, \infty) \leq n\pi_n(e^\delta D)^c \rightarrow 0$ and both $n\pi_n(D \cap \{v > e^{-\delta}b_1\})$ and $n\pi_n(D \cap \{u > e^{-\delta}b_2\})$ go to ∞ by Definition 4.1. The lemma

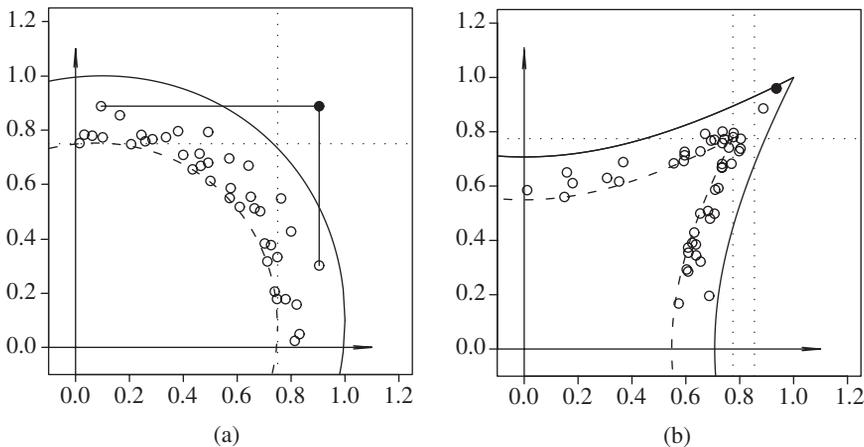


FIGURE 1: Points at the edge of simulated sample clouds of size $n = 10^5$ from (a) a bivariate normal distribution with mean 0 and correlation $\rho = 0.1$, and (b) a bivariate meta-Cauchy distribution with standard normal marginals, restricted to $(0, \infty)^2$ in both cases. Sample points are scaled by a factor $s_n = \sqrt{2 \log n}$. The filled circles indicate the coordinatewise maxima for these samples. The boundaries of the corresponding limit sets $E = \{x^2 - 2\rho xy + y^2 \leq 1 - \rho^2\}$ for (a) and $E = \{|x|^2 + |y|^2 + 1 \geq 3\|(x, y)\|_\infty^2\}$ for (b) are depicted by solid curves. The sample points inside the dashed curves are not displayed.

above applied to the shifted clouds shows that the probability p_n of a record in the sample cloud N_n vanishes for $n \rightarrow \infty$. This gives asymptotic independence; see [12] or Proposition A.1.

In order to complete the proof of Theorem 3.2, we now establish a link between the asymptotic shape of the scaled level sets of a light-tailed density and the shape of the limit set of the associated sample clouds.

Proposition 4.2. *Let X_1, X_2, \dots denote independent and identically distributed random vectors from a density $g \in \mathcal{A}(D)$. Then the sequence of scaled sample clouds $N_n = \{X_1/s_n, \dots, X_n/s_n\}$ converges onto the closure of the set D as $n \rightarrow \infty$ if the scaling constants s_n are chosen appropriately.*

Proof. The density g is shape equivalent to $f = f_0(n_D)$; see Proposition 3.4. The function $f_r(\mathbf{u}) = f(r\mathbf{u})/f_0(r)$ also lies in $\mathcal{H}(D)$, and by rapid variation and monotonicity of f_0 , for any $M > 1$ and $\varepsilon > 0$, eventually $f_r > M$ on $e^{-\varepsilon}D$ and $f_r < 1/M$ off $e^\varepsilon D$. By the pointwise inequality (3.6), the functions $g_r(\mathbf{u}) = g(r\mathbf{u})/f_0(r)$ satisfy the same inequalities eventually if we replace ε by 2ε . The measure ρ_r with density g_r satisfies $\rho_r(D^c) \ll \rho_r(U)$, $r \rightarrow \infty$, for any open set U which intersects D by Lemma 3.1. Choose s_n such that $ns_n^d f_0(s_n) = 1$ for $n \geq n_0$. Then ρ_{s_n} is the mean measure of the sample cloud N_n , and by Proposition 4.1, the sample clouds N_n converge onto D since their mean measures do.

5. Examples

In this section we illustrate the applicability of Theorems 3.1, 3.2, and 4.1 in determining whether a given distribution has asymptotically independent components. We also wish to see whether the conditions are sharp.

For any open bounded convex set D in the plane whose closure contains the origin, the function e^{-n_D} is integrable and, hence, $f(\mathbf{x}) = c_0 e^{-n_D(\mathbf{x})}$ for suitable $c_0 > 0$ is a probability density. Along rays it is an exponential function. If U is uniformly distributed on D then it is simple to decide when the coordinates U_1 and U_2 are asymptotically independent, but asymptotic independence of the coordinates of the vector X with density f is a different matter, even in the simple example where D is the intersection of a disk of radius $r = 2$ centred at $(1, -1)$ and the open set above the diagonal. The light-tailed examples below are of a more general nature.

Example 5.1. (*Rotund-exponential densities.*) Let X have a continuous homothetic density f with convex shape D and generating function f_0 . If D has a C^1 boundary and f_0 varies rapidly, the coordinates are asymptotically independent, and the sample clouds, properly scaled, converge onto the closure of the set D by Theorem 3.1 and Proposition 3.7. Now assume that D is rotund, i.e. the boundary ∂D is C^2 with positive definite curvature at every point. Also, assume that the generating function f_0 is asymptotic to a von Mises function $e^{-\psi}$. Then f is a so-called rotund-exponential density; see [3, Sections 9 and 10]. If we zoom in onto a boundary point of D so as to distinguish individual sample points, the sample clouds converge to a Gauss-exponential point process: $N_n \Rightarrow N$ vaguely on \mathbb{R}^d . The limit N is a Poisson point process with intensity

$$g(\mathbf{u}) = \frac{e^{-u_d} e^{-(u_1^2 + \dots + u_{d-1}^2)/2}}{(2\pi)^{(d-1)/2}}$$

if we choose the normalization appropriately. Weak convergence holds on all half-spaces

$$\{u_d \geq c_0 + c_1 u_1 + \dots + c_{d-1} u_{d-1}\}.$$

The restriction of g to the upper half-space $\{u_d \geq 0\}$ is a probability density. The corresponding vector has independent components. This vector is the limit of the high risk scenarios X^H , properly normalized, where X^H is the vector X conditioned to lie in the half-space H , and H moves off to infinity in the sense that $\Pr\{X \in H\} \rightarrow 0$. These results remain valid under certain perturbations (if the density $f = e^{-\psi(n_D)}$ is multiplied by a flat function L , see Section 11 of [3]). Such a perturbation does not affect the asymptotic behaviour of the exponent ψ , but may affect the limit shape. (Take $\psi = \log^2(1 + r)$ with $r^2 = x^2 + y^2$ and $L = e^\lambda$ with $\lambda = \log(1 + 2r + cx)$, $c \in [-1, 1]$. If we choose $c = c(r) = \sin(\log \log r)$ then L is still flat, but the shape of the level sets no longer converges.) See Proposition 14.1 of [3]. So we see that, under the extra conditions on the homothetic density f , there are three alternative asymptotic descriptions of large sample clouds from this density. Convergence onto the closure of the set D describes the global behaviour of the sample clouds; weak convergence in the space $\mathcal{X} = [-\infty, \infty]^d \setminus \{-\infty\}$ to a Poisson point process whose mean measure is the exponent measure of an extreme value limit law (Gumbel with independent components); weak convergence to the Gauss-exponential Poisson point process N on certain half-spaces describes the local behaviour at boundary points of D .

Example 5.2. (*Skew-normal densities.*) A symmetric density g satisfies $g(-x) = g(x)$. It may be transformed into an asymmetric density by multiplication with a positive continuous asymmetric function θ which satisfies $\theta(x) + \theta(-x) = 2$. The skew-normal distributions $SN(\Omega, \alpha)$ introduced in [2] have a density f which is the product of a centred Gaussian density with covariance Ω and the function $x \mapsto 2\Phi(\alpha^\top x)$, where Φ is the standard normal DF and α is a nonzero linear functional. These densities are log-concave (since $(-\log \Phi)'$ is positive) and, hence, have convex level sets.

We claim that $f \in \mathcal{A}(D)$ for a convex set D with C^1 boundary. Write $f = e^{-h}$. The function $-\log \Phi(t)$ is asymptotic to $t^2/2$ for $t \rightarrow -\infty$ and vanishes for $t \rightarrow \infty$. Let u_0 satisfy $\alpha^\top u_0 > 0$ and $u_0^\top \Omega^{-1} u_0 = 1$. Then $h(tu_0)/t^2 \rightarrow \frac{1}{2}$ and

$$\frac{h(tu)}{h(tu_0)} \rightarrow v(u) = \begin{cases} u^\top \Omega^{-1} u & \text{for } \alpha^\top u \geq 0, \\ u^\top \Omega^{-1} u + (\alpha^\top u)^2 & \text{for } \alpha^\top u < 0, \end{cases} \quad t \rightarrow \infty.$$

It follows from Proposition 3.7 that there is a limit shape, $D = \{v < 1\}$, which is the covariance ellipsoid on the half-space where $\alpha^\top u$ is positive, and a flattened version of this ellipsoid on the complementary half-space; see Figure 2. The set D is convex. To see that the boundary is C^1 , choose coordinates such that the underlying Gaussian density is standard, with spherical level sets, and then choose the vertical coordinate in the direction of α . At these coordinates D agrees with the unit ball B on the upper half-space, and with the cylinder symmetric ellipsoid $\{x_1^2 + \dots + x_{d-1}^2 + (1 + c^2)x_d^2 < 1\}$ for some $c > 0$ on the lower half-space $\{x_d \leq 0\}$. For boundary points p in the horizontal coordinate plane, the tangent plane is vertical: $p^\top x = 1$. Theorem 3.1 applies. If X has a skew-normal distribution then the random variables $\xi_1(X), \dots, \xi_m(X)$ are asymptotically independent whenever the linear functionals ξ_1, \dots, ξ_m are linearly independent. Asymptotic independence of the skew-normal distribution has been partially proven in [17] using a direct analytic approach based on Sibuya’s condition.

Example 5.3. Densities of the form $f(x) = f_0(\|x\|_p)$ for $p > 0$ have level sets which are balls in ℓ_p . Lower-dimensional marginals have the same form but with a different generator, since the projection of the d -dimensional unit ball on the space spanned by the first m coordinates is the m -dimensional unit ball. The two-dimensional unit ball is blunt for all $p \in [1, \infty)$, and, hence,

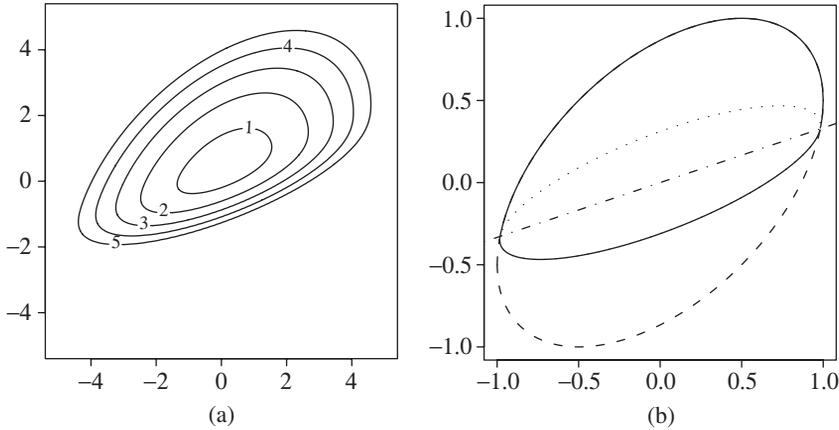


FIGURE 2: (a) Contours of the density of the $SN_2(\alpha; \omega)$ distribution with $\omega = 0.5$ and $\alpha = (-1, 3)$; levels are indicated as powers of 10^{-1} . (b) Asymptotic shape of the scaled level sets (*solid curve*); the dash-dotted straight line is given by the equation $\alpha^T \mathbf{u} = 0$.

vectors \mathbf{X} with light-tailed densities f as above have asymptotically independent coordinates. However, for $p = \infty$, the supnorm, the unit ball is a cube $(-1, 1)^d$, and the square is not blunt. The components of \mathbf{X} are still asymptotically independent, but we need extra work to prove this. The bivariate marginals of $f(\mathbf{x}) = f_0(\|\mathbf{x}\|_\infty)$ have the same form with a different generator f_0 which is still continuous, strictly decreasing, and rapidly varying at ∞ . (The cubic slices are replaced by square slices.) It suffices to consider bivariate densities $f(x, y) = f_0(|x| \vee |y|)$ for continuous, strictly decreasing positive functions f_0 on $[0, \infty)$ which vary rapidly at ∞ . The marginals f_i of f , $i = 1, 2$, are equal by symmetry, and $f_2(y) = 2yf_0(y) + 2R(y)$, where $R(y) = \int_y^\infty f_0(t) dt \ll yf_0(y)$ by rapid variation of f_0 . Hence, $2R(t)/f_1(t) \rightarrow 0$ for $t \rightarrow \infty$, and, by l'Hôpital's rule, also the quotient $\Pr\{X > t, Y > t\}/\Pr\{X > t\}$. Thus, Sibuya's condition holds. However, asymptotic independence need not hold if the level sets are only asymptotically cubic; see [4, Example 2].

Another example showing that results for $\mathcal{H}(D)$ do not need to carry over to $\mathcal{A}(D)$ is given in Example A.5, below. Our last example illustrates an asymptotically dependent distribution with a density in $\mathcal{A}(D)$ where bivariate projections of D are nonblunt.

Example 5.4. Let \mathbf{Z} have a bivariate t density with $\lambda > 0$ degrees of freedom. Transform the marginals to obtain a vector \mathbf{X} with standard Gaussian components. The new density g is called a meta- t density; see [19, p. 193]. It has normal marginals, but the copula of the elliptic t distribution. The shape of the level sets of the density g converges to the symmetric subset $D = \{u_1^2 + u_2^2 + \lambda > (\lambda + 2)\|(u_1, u_2)\|_\infty^2\}$ of the square $(-1, 1)^2$; see [5]. Figure 1(b) shows a detail of the limit set D . The set D is not blunt, and the components of \mathbf{X} are asymptotically dependent since those of \mathbf{Z} are.

6. Conclusion

We have explored conditions for asymptotic independence of the components of a multivariate random vector expressed in terms of the limiting shape of the level sets of the underlying

density. A distinction had to be made between light and heavy tails. For light-tailed densities, the limiting shape of level sets is essential in determining whether asymptotic independence holds. In contrast, for heavy-tailed densities, the (limiting) shape of level sets is irrelevant as long as the shape intersects the positive orthant. In the light-tailed case there is a simple sufficient condition for asymptotic independence of two components in terms of the corresponding bivariate projection of the shape. This subset of the plane has to be blunt. The more delicate question of the relation between shape and asymptotic independence when the bivariate projection is not blunt will be treated in a future publication.

Asymptotic dependence is a basic concern in multivariate risk analysis. The light-tailed densities studied in this paper have the property that sample clouds will have the same shape as the level sets of the density asymptotically. For sample clouds, persistence of the shape, as the number of sample points increases, opens the possibility of using the shape to construct densities over the whole space. This makes it possible to estimate probabilities of regions far out which contain only a few or no sample points.

Appendix A

A.1. Supplementary results

Proposition A.1. *Let X_1, X_2, \dots be independent observations from the continuous DF F on \mathbb{R}^2 . The probability of a record amongst the first n observations goes to 0 if and only if F is asymptotically independent.*

Proof. We may assume that F is a copula. Set $c_n = nP(1 - 1/n, 1 - 1/n)$. Then the Poisson approximation gives a probability $p_n = c_n e^{-c_n} e^{-(1-c_n)} = c_n e^{c_n - 2}$ to the event: among the first n observations there is exactly one in the complement of $[0, 1 - 1/n]^2$, and that observation lies in the square $(1 - 1/n, 1]^2$. In the case of asymptotic dependence, $c_{k_n} \rightarrow c > 0$ for some subsequence, and, hence, the probability of a record in a sample of size k_n will exceed $ce^{c-2}/2$ eventually. Conversely, asymptotic independence implies that $nP(1 - M/n, 1 - M/n) \rightarrow 0$ for any $M > 1$, whereas the marginals satisfy $nP(0, 1 - M/n) = M$. As in Lemma 4.1, the probability of a record vanishes.

Proposition A.2. *Suppose that $f \in \mathcal{H}(D)$ and that the generator f_0 varies regularly or satisfies O -variation:*

$$r_{n+1} \sim r_n \rightarrow \infty \Rightarrow f_0(r_{n+1}) \sim f_0(r_n). \tag{A.1}$$

Then g is shape equivalent to f if and only if $g \sim f$.

Proof. Condition (A.1) implies that, for every $\varepsilon > 0$, there exists $\delta > 0$ such that $f_0(e^\varepsilon r) > e^{-\delta} f_0(r)$ for all $r > 1$. Suppose that g is shape equivalent to f . Let $n_D(\mathbf{x}_n) \rightarrow \infty$. Suppose that $g(\mathbf{x}_n) = f_0(r_n)$. Then, for any $\varepsilon > 0$, the point \mathbf{x}_n eventually lies in the ring $e^\delta r_n D \setminus e^{-\delta} r_n D$ on which f fluctuates by a factor at most e^ε . To show the converse, suppose that $g \sim f$. Let $g(\mathbf{x}_n) \sim f(\mathbf{x}_n) = f_0(r_n) \sim f_0(r_{n+1})$ for $r_{n+1} \sim r_n \rightarrow \infty$. Then, for any $\varepsilon, \varepsilon_1 > 0$, eventually

$$\begin{aligned} (1 - \varepsilon_1) f_0(e^\varepsilon r_n) &\leq (1 - \varepsilon_1) f_0(r_{n+1}) \\ &\leq g(\mathbf{x}_n) \leq (1 + \varepsilon_1) f_0(r_{n+1}) \leq (1 + \varepsilon_1) f_0(e^{-\varepsilon} r_n), \quad n \geq n_0. \end{aligned}$$

Since ε_1 is arbitrary, we have $f(e^\varepsilon \mathbf{x}_n) \leq g(\mathbf{x}_n) \leq f(e^{-\varepsilon} \mathbf{x}_n)$ for $n \geq n_0$, as required for shape equivalence by (3.3).

A.2. Counterexamples

This section contains counterexamples mentioned in the main text.

Example A.1. A density $f \in \mathcal{H}(D)$ may have spherical level sets without exhibiting spherical symmetry. This will be the case if D is an off-centre ball with $n_D(x) = \sqrt{\|x\|_2^2 + \beta^\top x - \beta^\top x}$ for some $\beta \in \mathbb{R}^d$. To be star shaped, the origin has to lie in D , or be a boundary point. In the latter case the set $\{f > 0\}$ is a ball (if $\{f_0 > 0\}$ is a bounded interval) or an open half-space.

Example A.2. For any $\varepsilon > 0$, there exists a bounded open star-shaped set D which contains the origin, whose closure is the cube $K = [-1, 1]^d$, and such that the volume of D is small, $|D| < \varepsilon$. To construct such a star-shaped set, take a dense sequence x_n on the boundary ∂K , and define $U \subset \partial K$ as the union of open disks with centre x_n and radius ε_n , where $\varepsilon_n \rightarrow 0$ so fast that the area of U is $\varepsilon/2$. Now let D be the union of an open centred ball with volume $\varepsilon/2$ and the set of all points ru with $0 < r < 1$ and $u \in U$.

Example A.3. Consider a continuous, strictly positive density f on \mathbb{R}^3 whose level sets $\{f > c\} = r_c D$ all have the same shape. The set D is convex, even rotund, and the function $c \mapsto r_c$ is continuous and strictly increasing. The marginal densities are not necessarily all unimodal.

Start with the uniform distribution on the tetrahedron T_0 with two vertices in the horizontal plane, say e_1 and e_2 . The other two vertices are e_3/m and $-e_3$. Here m is a positive integer to be chosen later. The marginal along the vertical axis has a continuous density g_0 on $(-1, 1/m)$ which is parabolic on the interval $(-1, 0)$ and on $(0, 1/m)$, vanishing at the endpoints of the interval and with a maximum at the origin. The shifted tetrahedron $T = T_0 + e_3/2 - (e_1 + e_2)/8$ contains the origin as an interior point. Its vertical marginal density g is g_0 shifted upwards over $\frac{1}{2}$ and has its maximum at $\frac{1}{2}$. The vertical marginal density \tilde{g} of the uniform distribution on the tetrahedron $T/2$ has its maximum at $\frac{1}{4}$. The fair mixture of the uniform distribution on T and $T/2$ has a density \hat{f} whose marginal $\hat{g} = (g + \tilde{g})/2$ is not convex if m is large (since the left derivative of \tilde{g} at $\frac{1}{4}$ is large). Now choose rotund sets D_n converging to T and densities f_n converging to \hat{f} which satisfy the conditions of the example. If infinitely many of the vertical marginals g_n were unimodal then the limit \hat{g} would be. We conclude that eventually g_n is not unimodal.

Example A.4. Here are some examples of functions $h = qg$ which are shape equivalent to the standard normal density g on the plane. In view of Lemma 3.2, we see that qg is shape equivalent to g for a continuous function $q = e^\psi$ precisely if $|\psi(x)| \leq \chi(\|x\|_2)$ for a function $\chi(r) \ll r^2$. We may take q to be one of the following functions: $1 + |x|$, $(1 + r^2)^m$ with $m \geq 1$, e^r , $e^{x-|y|^{3/2}}$ where $r^2 = x^2 + y^2$. These functions may be multiplied with a function like $\exp(\sin \pi e^{x^2} \sin \pi e^{y^6})$ which fluctuates rapidly but is weakly asymptotic to a constant. The level sets $\{h > c\}$ will then look locally like a shore with many small islands, and lakes, even though the sets are asymptotic to disks $\{x^2 + y^2 < r^2\}$ with $r = \sqrt{2 \log(1/c)}$.

Example A.5. Let D be the open triangle with vertices $(1, 1)$, $(-1, 0)$, and $(0, -1)$. It contains the origin. Let $f \in \mathcal{A}(D)$ have convex level sets and be shape equivalent to $g = e^{-nD}$. It is possible that f is asymptotically independent. The function g has triangular level sets $\{g > e^{-t}\} = tD$. Let f have level sets $\{f > e^{-t}\} = D_t = tD \setminus \{x + y \geq 2t - \sqrt{t}\}$. Then D_t/t agrees with the triangle D except that the extreme top has been sliced off. If we choose $p(t) = (t, t) - (\sqrt{t}, \sqrt{t})/2$, then

$$h_t(u) = \frac{f(p(t) + u)}{f(p(t))} \rightarrow h(u), \quad t \rightarrow \infty,$$

where $\{h > e^{-t}\} = C + (t, t)$ for the half-space $C = \{u + v < 0\}$. Let ρ_t have density h_t , and let ρ have density h . Then $\rho_t \rightarrow \rho$ weakly on $[0, \infty)^2$. Since $\rho[0, \infty)^2$ is finite and ρ gives infinite weight to the half-spaces $\{y \geq 0\}$ and $\{x \geq 0\}$, Sibuya's condition holds by Proposition 2.3 with curve $p(t)$, $t \geq 1$.

Acknowledgements

Paul Embrechts suggested that the relation between the shape of level sets and asymptotic independence should be written up. We thank him for several stimulating discussions. The financial support of the Institute for Mathematical Research (FIM) and RiskLab, Switzerland, is gratefully acknowledged.

References

- [1] ARNOLD, B. C., CASTILLO, E. AND SARABIA, J. M. (2008). Multivariate distributions defined in terms of contours. *J. Statist. Planning Infer.* **138**, 4158–4171.
- [2] AZZALINI, A. AND DALLA VALLE, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83**, 715–726.
- [3] BALKEMA, A. A. AND EMBRECHTS, P. (2007). *High Risk Scenarios and Extremes*. European Mathematical Society, Zürich.
- [4] BALKEMA, A. A., EMBRECHTS, P. AND NOLDE, N. (2009). Sensitivity of the asymptotic behaviour of meta distributions. Preprint. Available at <http://arxiv.org/abs/0912.5337v1>.
- [5] BALKEMA, A. A., EMBRECHTS, P. AND NOLDE, N. (2010). Meta densities and the shape of their sample clouds. *J. Multivariate Anal.* **101**, 1738–1754.
- [6] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press.
- [7] DE HAAN, L. AND FERREIRA, A. (2006). *Extreme Value Theory. An Introduction*. Springer, New York.
- [8] DHARMADHIKARI, S. AND JOAG-DEV, K. (1988). *Unimodality, Convexity, and Applications*. Academic Press, Boston, MA.
- [9] FANG, K. T., KOTZ, S. AND NG, K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.
- [10] FERNÁNDEZ, C., OSIEWALSKI, J. AND STEEL, M. F. J. (1995). Modeling and inference with v -spherical distributions. *J. Amer. Statist. Assoc.* **90**, 1331–1340.
- [11] GNEDIN, A. V. (1993). On multivariate extremal processes. *J. Multivariate Anal.* **46**, 207–213.
- [12] GNEDIN, A. V. (1994). On a best-choice problem with dependent criteria. *J. Appl. Prob.* **31**, 221–234.
- [13] HASHORVA, E. (2005). Extremes of asymptotically spherical and elliptical random vectors. *Insurance Math. Econom.* **36**, 285–302.
- [14] HULT, H. AND LINDSKOG, F. (2002). Multivariate extremes, aggregation and dependence in elliptical distributions. *Adv. Appl. Prob.* **34**, 587–608.
- [15] KINOSHITA, K. AND RESNICK, S. I. (1991). Convergence of scaled random samples in \mathbb{R}^d . *Ann. Prob.* **19**, 1640–1663.
- [16] LEDFORD, A. W. AND TAWN, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169–187.
- [17] LYSENKO, N., ROY, P. AND WAEBER, R. (2009). Multivariate extremes of generalized skew-normal distributions. *Statist. Prob. Lett.* **79**, 525–533.
- [18] MAULIK, K., RESNICK, S. AND ROOTZÉN, H. (2002). Asymptotic independence and a network traffic model. *J. Appl. Prob.* **39**, 671–699.
- [19] McNEIL, A. J., FREY, R. AND EMBRECHTS, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.
- [20] MIKOSCH, T. (2006). Copulas: tales and facts. *Extremes* **9**, 3–20.
- [21] OSIEWALSKI, J. AND STEEL, M. F. J. (1993). Robust Bayesian inference in ℓ_q -spherical models. *Biometrika* **80**, 456–460.
- [22] RAMOS, A. AND LEDFORD, A. (2009). A new class of models for bivariate joint tails. *J. R. Statist. Soc. B* **71**, 219–241.
- [23] RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.
- [24] RESNICK, S. I. (2002). Hidden regular variation, second order regular valuation and asymptotic independence. *Extremes* **5**, 303–336.
- [25] SIBUYA, M. (1960). Bivariate extreme statistics. *Ann. Inst. Statist. Math.* **11**, 195–210.