Monge-Kantorovich superquantiles and expected shortfalls with applications to multivariate risk measurements

Bernard Bercu, Jérémie Bigot and Gauthier Thurin

Université de Bordeaux,
Institut de Mathématiques de Bordeaux et CNRS UMR 5251,
351 Cours de la libération,
33400 Talence cedex, France,
e-mail: bernard.bercu@math.u-bordeaux.fr*;* jeremie.bigot@math.u-bordeaux.fr*,*
gauthier-louis.thurin@math.u-bordeaux.fr

Abstract: We propose center-outward superquantile and expected shortfall functions, with applications to multivariate risk measurements, extending the standard notion of value at risk and conditional value at risk from the real line to $\mathbb{R}^d$. Our new concepts are built upon the recent definition of Monge-Kantorovich quantiles based on the theory of optimal transport, and they provide a natural way to characterize multivariate tail probabilities and central areas of point clouds. They preserve the univariate interpretation of a typical observation that lies beyond or ahead a quantile, but in a meaningful multivariate way. We show that they characterize random vectors and their convergence in distribution, which underlines their importance. Our new concepts are illustrated on both simulated and real datasets.

MSC2020 subject classifications: Primary 62H05, 62P99, 60F05.
Keywords and phrases: Center-outward quantiles, tails of a multivariate distribution, convergence in distribution, multivariate risk analysis.

1. Introduction

1.1. Superquantile, expected shortfall

Modeling the dependency between the components of a random vector is at the core of multivariate statistics. To that end, one way to proceed is to characterize the multivariate probability tails. For distributions supported on the real line, this is often tackled with the use of superquantiles or expected shortfalls, that complement the information given by the quantiles. Let X be an integrable absolutely continuous random variable with cumulative distribution function F. For all $\alpha \in]0, 1[$, the quantile $Q(\alpha)$ of level α is given by
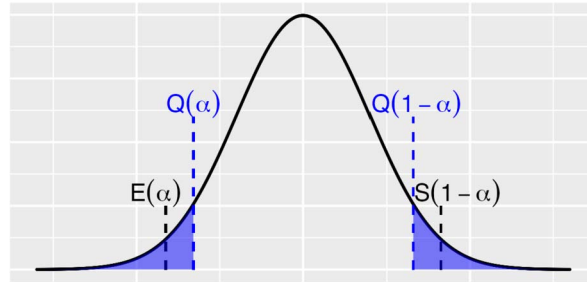
$$Q(\alpha) = \inf\{x : F(x) \geq \alpha\},$$

FIG 1. *Illustration of the notions of superquantile $S(\alpha)$ and expected shortfall $E(\alpha)$ for a univariate Gaussian distribution.*

whereas the superquantile $S(\alpha)$ and expected shortfall $E(\alpha)$ are defined by

$$S(\alpha) = \mathbb{E}[X|X \geq Q(\alpha)] = \frac{\mathbb{E}[X\mathbb{1}_{X \geq Q(\alpha)}]}{\mathbb{P}(X \geq Q(\alpha))} = \frac{1}{1-\alpha}\mathbb{E}[X\mathbb{1}_{X \geq Q(\alpha)}], \qquad (1.1)$$

and

$$E(\alpha) = \mathbb{E}[X|X \leq Q(\alpha)] = \frac{\mathbb{E}[X\mathbb{1}_{X \leq Q(\alpha)}]}{\mathbb{P}(X \leq Q(\alpha))} = \frac{1}{\alpha}\mathbb{E}[X\mathbb{1}_{X \leq Q(\alpha)}]. \qquad (1.2)$$

As illustrated in Figure 1, $S(\alpha)$ focuses on the upper-tail while $E(\alpha)$ targets the lower-tail. We emphasize that using the terms of superquantile and expected shortfall is a subjective consideration taken from [1, 54]. Most of the time, one does not consider the upper and the lower tails together, so that a single name is required, up to considering the distribution of $-X$. In this vein, depending on the application, the expected shortfall may refer to the same as the Conditional-Value-at-Risk, Conditional-Tail-Expectation, see *e.g.* [2], or even the superquantile, that aims to be a neutral alternative name in statistics [54].

The main contribution of the present paper is to extend (1.1) and (1.2) towards a notion of multivariate superquantile and expected shortfall. As part of the difficulty, both the mathematical meanings of "ahead", "beyond" and "typical" do not adapt canonically in $\mathbb{R}^d$. We argue that sufficient notions are provided by the Monge-Kantorovich (MK) quantiles, ranks and signs, introduced in [13, 37]. In particular, the traditional left-to-right ordering is replaced in our approach by a center-outward one that is more intuitive for a point cloud [13]. Hence, the two subsets of observations that we are interested in are located at the outward or near the mean value, which requires to adapt the concepts of (1.1) and (1.2) in $\mathbb{R}^d$. It is well-known from a simple change of variables in $\mathbb{R}$ that S and E average observations beyond and ahead the quantile of level α, in the sense that

$$S(\alpha) = \frac{1}{1-\alpha}\int_\alpha^1 Q(t)dt \qquad \text{and} \qquad E(\alpha) = \frac{1}{\alpha}\int_0^\alpha Q(t)dt. \qquad (1.3)$$

In Section 2, our definitions generalize the formulation (1.3). If Q stands for the multivariate MK quantile function instead of the classical univariate one,

center-outward superquantile and expected shortfall functions are defined, for any u in the unit ball $\mathbb{B}(0,1)\backslash\{0\}$, by

$$S(u) = \frac{1}{1-\|u\|} \int_{\|u\|}^{1} Q\Big(t\frac{u}{\|u\|}\Big)dt \qquad \text{and} \qquad E(u) = \frac{1}{\|u\|} \int_{0}^{\|u\|} Q\Big(t\frac{u}{\|u\|}\Big)dt.$$

MK quantiles have already led to many applications, among which lie statistical tests [31, 39, 47, 57, 58], regression [12, 23], risk measurement [5], or Lorenz maps [27, 38]. We also refer to the recent review [36] on the concept of MK quantiles. Importantly, the univariate quantile and related functions are deeply rooted in risk analysis. On the one hand, risk measures which are both coherent and regular can be characterized by integrated quantile functions [35]. On the other hand, given a level α, fundamental risk measures are given by $Q(\alpha)$ and $S(\alpha)$, called Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR), respectively. As a matter of fact, a natural main contribution of the present paper is to provide meaningful multivariate extensions of VaR and CVaR.

1.2. Background on multivariate risk measurement

MK quantiles and the related concepts have already been applied to multivariate risk measurement in [5], and, in some sense, in the previous works [26, 29]. The theory in [26] states ideal theoretical properties for coherent regular risk measures, while the maximal correlation risk measure of [5] furnishes a real-valued risk measure with these properties. This constitutes, to the best of our knowledge, the short literature on risk measurement based on the MK quantile function. In this work, we argue that an adequate procedure of multivariate risk measurement shall account for all the information on the tails, both in terms of direction and spreadness. To answer this issue, vector-valued risk measures are natural candidates. There also exist several extensions of VaR or CVaR to the multivariate setting, including [3, 11, 14, 15, 32, 40, 50, 59, 60], but none of them is based on the theory of optimal transportation and its associated potential benefits. In particular, our concepts do not require any assumption on the tail behavior of the data, nor any statistical model, because MK quantiles adapt naturally to the shape of a point cloud. On the real line, the VaR and the CVaR have a clear interpretation: for a level $\alpha \in [0,1]$, the VaR is the worst observation encountered with probability $1 - \alpha$ whereas CVaR is the average value beyond this worst observation. Such a meaningful definition is surely part of the reason for their wide use in practice. Under the name of Conditional-Tail-Expectation, the idea proposed in [15, 24] preserves this interpretation, but relies on level sets defined from the theory of copulas. Specifically, the obtained quantile levels do not adapt automatically to the shape of the data. Still, this notion averages over a certain quantile level, and it returns a tail observation of the same dimension as the data. Our work is inspired by this approach, as we aim to give the same information about multivariate tails, but we use the MK quantile function, which yields, to our opinion, concepts with even better interpretability.

1.3. Main contributions

Our main contributions are the definitions of a center-outward superquantile function and related risk measures, both real-valued and vector-valued. Doing so, we provide an extension to the multivariate case of the fundamental Value-at-Risk and Conditional-Value-at-Risk. Furthermore, we provide a center-outward expected shortfall function, that describes the central areas of a given point cloud. Our center-outward expected shortfalls and superquantiles are uniquely defined and characterize convergence in distribution, and they are closely related to the potential whose gradient gives the MK quantile function. A result of independent interest is also provided, giving a new family of Monge maps between known probability distributions. This may motivate changing the reference distribution under generalized gamma models.

1.4. Outline of the paper

Section 2 details definitions and properties of our new concepts of center-outward superquantiles and expected shortfalls. It includes our main results and the crucial relation between these functions and Kantorovich potentials. An alternative class of reference measures, which differ from the standard spherical uniform distribution on the unit ball, is provided in Section 3. The multivariate definitions of VaR and CVaR are given in Section 4. Finally, we present in Section 5 a regularized version of our superquantile and expected shortfall functions, using entropically regularized optimal transport that has fundamental computional benefits to estimate the center-outward quantile function. Numerical experiments are also provided to shed some light on the benefits of our new concepts of MK superquantile, expected shortfall and multivariate VaR and CVaR for multivariate data analysis. A conclusion and some perspectives are given in Section 6.

2. Center-outward superquantiles and expected shortfalls

2.1. Main definitions

On the real line, the notion of superquantile and expected shortfall relies heavily on the one of quantile. It is then natural to make use of the Monge-Kantorovich (MK) quantile function and its appealing properties in order to define associated superquantile and expected shortfall functions. By simplicity, we shall restrict ourselves to the set of integrable probability measures over $\mathbb{R}^d$, that is

$$\mathcal{P}_1(\mathbb{R}^d) = \big\{\nu : \mathbb{E}_{X \sim \nu}[\|X\|] < +\infty\big\}.$$

Recall that a probability measure $\nu \in \mathcal{P}_1(\mathbb{R}^d)$ is the push-forward of $\mu \in \mathcal{P}_1(\mathbb{R}^d)$ by $T : \mathbb{R}^d \to \mathbb{R}^d$ if $T(U)$ has distribution ν as soon as U is distributed according to μ. This is denoted by $T_{\#}\mu = \nu$. The following has been introduced in [37].

Definition 2.1. *The MK quantile function of a multivariate distribution ν, with respect to a reference distribution μ, is a push-forward map $Q_{\#}\mu = \nu$ such that there exists a convex potential $\psi : \mathbb{R}^d \to \mathbb{R}$ satisfying $\nabla\psi = Q$ μ-almost everywhere.*

It follows from McCann's theorem, [44] that, as soon as μ is absolutely continuous, such a Monge map Q exists and is unique. Moreover, if μ and ν have finite moments of order two, by the result known in the literature as Brenier's theorem [10, 16], Q is characterized as the solution of the following Monge problem of optimal transport,

$$Q = \operatorname*{argmin}_{T:T_{\#}\mu=\nu} \int_{\mathcal{X}} \|u - T(u)\|^2 d\mu(u). \tag{2.1}$$

Intuitively, the reference distribution μ must be chosen so that a relevant notion of quantiles can be derived from it, whereas being the gradient of a convex function ψ is a generalization of monotonicity. For instance, one can choose the *spherical uniform* distribution, denoted by $\mu = U_d$. It is given by the product $R\Phi$ between two independent random variables R and Φ, being drawn respectively from a uniform distribution on $[0,1]$ and on the unit sphere. Samples from U_d are distributed from the origin to the outward within the unit ball, so that the balls of radius $\alpha \in [0,1]$ have probability α while being nested, as α grows. With this in mind, the hyperspheres of radius α are relevant quantile contours with respect to μ. Being the gradient of a convex function, Q adequately transports this *center-outward* ordering towards the distribution ν. Thus, when $\mu = U_d$, we shall refer to Q as the *center-outward* quantile function of ν. This property is illustrated in Figure 2, where radius, in red, and circles, in blue, are transported from the unit ball to a banana-shaped distribution thanks to the mapping Q obtained with the computational approach described in Section 5.

This relevant ordering clearly catches the geometry of the support of the target distribution ν, and it comes with quantile regions indexed by a probability level $\alpha \in [0,1]$, by use of the change of variables formula for push-forward maps. More details are given in [13, 37]. Hereafter, we use the spherical uniform as the reference distribution, but other distributions could be chosen, depending on the applications, [18, 27, 31]. Note that if ν is an empirical measure based on random observations, it has a finite second order moment, and the estimation of the quantile map amounts to solve the OT problem (2.1). Following [37], we assume, without loss of generality, that ψ in Definition 2.1 satisfies $\psi(0) = 0$ and, for $u \in \mathbb{R}^d$ such that $\|u\| = 1$,

$$\psi(u) = \liminf_{\substack{v \to u \\ \|v\| < 1}} \psi(v). \tag{2.2}$$

Moreover, we impose, for all $u \in \mathbb{R}^d$ such that $\|u\| > 1$,
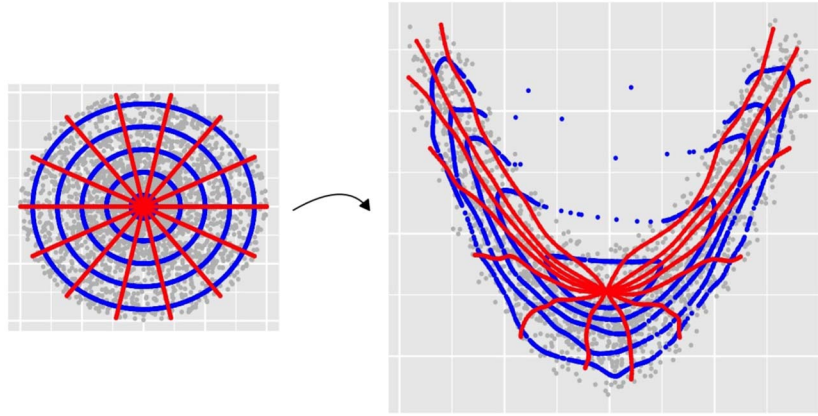
$$\psi(u) = +\infty. \tag{2.3}$$

FIG 2. *(Left) center-outward quantiles of the spherical uniform distribution* $\mu = U_d$, *and (right) center-outward quantiles of a discrete distribution* ν *obtained by* $Q_{\#}\mu = \nu$.

This being said, the convex potential ψ is uniquely defined over its domain $\mathrm{Dom}(\psi) = \{u | \psi(u) < +\infty\}$, that verifies $\mathbb{B}(0,1) \subset \mathrm{Dom}(\psi) \subset \overline{\mathbb{B}}(0,1)$. Although ψ is continuous over $\mathbb{B}(0,1)$, [52][Theorem 10.3], the gradient $\nabla\psi$ is only defined almost everywhere. At every u where ψ is not differentiable, one can still define the subdifferential

$$\partial\psi(u) = \{z \in \mathbb{R}^d : \forall x \in \mathbb{R}^d, \psi(x) - \psi(u) \geq \langle z, x - u \rangle\}.$$

Following the suitable suggestion of an anonymous referee, for all $u \in \mathbb{B}(0,1)$, we can define $Q(u)$ as the average of $\partial\psi(u)$ so that Q is defined everywhere. In fact, from [28], as soon as ν is a continuous probability measure with non vanishing density on $\mathbb{R}^d$, ψ can be shown to be differentiable everywhere on $\mathbb{B}(0,1)\backslash\{0\}$. The same result is showed under milder assumptions in [20], that is presented hereafter.

Assumption 2.1. *Let ν be an absolutely continuous measure with probability density function p defined on its support $\mathcal{X}$. For every $R > 0$, there exist two constants $0 < \lambda_R < \Lambda_R$ such that, for all $x \in \mathcal{X} \cap \mathbb{B}(0,R)$,*

$$\lambda_R \leq p(x) \leq \Lambda_R.$$

Assumption 2.2. *The support $\mathcal{X} \subset \mathbb{R}^d$ of ν is convex.*

Under these assumptions, the next theorem is given in [20].

Theorem 2.1 (Regularity of the center-outward quantile function, [20])**.** *Under Assumptions 2.1 and 2.2, there exists a compact convex set K with Lebesgue measure 0 such that the center-outward quantile function Q is a homeomorphism from $\mathbb{B}(0,1)\backslash\{0\}$ to $\mathcal{X}\backslash K$, with inverse Q^{-1} the center-outward distribution function. Moreover, $Q^{-1} = \nabla\psi^*$ where ψ^* is the Fenchel-Legendre transform*

of ψ such that $Q = \nabla\psi$,

$$\psi^*(x) = \sup_{u \in \mathbb{B}(0,1)} \{\langle x, u \rangle - \psi(u)\}.$$

Hereafter, this allows us to properly define quantile contours, ranks and signs. For the sake of simplicity, we define $\mathcal{P}_*(\mathbb{R}^d)$ as the set of integrable probability measures for which Q is a homeomorphism from $\mathbb{B}(0,1)\backslash\{0\}$ to its image, see Theorem 2.1. This includes any ν satisfying assumptions 2.1 and 2.2, that is also integrable, to ensure finiteness of the center-outward superquantiles. We emphasize that, for any $\nu \in \mathcal{P}_*(\mathbb{R}^d)$, Q is defined everywhere on $\mathbb{B}(0,1)\backslash\{0\}$, as opposed to almost everywhere on $\mathbb{B}(0,1)$. The next definitions, taken from [13, 37], gather the main concepts that we use in the following.

Definition 2.2 (Quantile contours, ranks and signs)**.** *Let $\nu \in \mathcal{P}_*(\mathbb{R}^d)$ with center-outward quantile function Q and supported on $\mathcal{X} \subset \mathbb{R}^d$. Then, for the distribution ν,*
(i) the quantile region $\mathbb{C}_\alpha$ of order $\alpha \in [0,1]$ is the image by Q of the ball $\mathbb{B}(0,\alpha)$.
(ii) the quantile contour $\mathcal{C}_\alpha$ of order $\alpha \in [0,1]$ is the boundary of $\mathbb{C}_\alpha$.
(iii) the rank function $\mathcal{R}_\nu : \mathcal{X} \to [0,1]$ is defined by $\mathcal{R}_\nu(x) = \|Q^{-1}(x)\|$.
(iv) the sign function $\mathcal{D}_\nu : \mathcal{X} \to \mathbb{B}(0,1)$ is defined by $\mathcal{D}_\nu(x) = Q^{-1}(x)/\|Q^{-1}(x)\|$.

A few remarks follow from Definition 2.2. The ν-probability of $\mathbb{C}_\alpha$ is α, by the change of variables formula for push-forward maps, which is a first requirement for quantile regions. In addition, one may note that the rank and sign functions require the invertibility of Q. From [37][Section 2], the continuity and invertibility of Q outside the origin ensures the crucial fact that the quantile contours are closed and nested. It is the notion of center-outward ranks that allows to order points in $\mathcal{X}$ relatively to ν, consistently with Tukey's halfspace depth, as highlighted in [13]. It induces the following weak order.

Definition 2.3. *For $\nu \in \mathcal{P}_*(\mathbb{R}^d)$ and $x, y \in \mathcal{X}$, we denote $x \geq_\mathcal{R} y$ and say that y is deeper than x if*

$$\|Q^{-1}(x)\| \geq \|Q^{-1}(y)\|.$$

The deeper a point is in $\mathcal{X}$, the less extremal it is with respect to ν. For a fixed $u \in \mathbb{B}(0,1)\backslash\{0\}$, we can then write $x \geq_\mathcal{R} Q(u)$ which enables to consider the observations "beyond" $Q(u)$ in some sense. In a context where the focus is on the *central* observations, $u \mapsto \mathbb{E}[X|X \leq_\mathcal{R} Q(u)]$ has been introduced in [38] as a center-outward Lorenz function for X, see Section 6.1. The reverse conditional expectation $\mathbb{E}[X|X \geq_\mathcal{R} Q(u)]$ might be thought of as a natural candidate for a superquantile concept. Nevertheless, it cannot be understood as a "typical" observation within *outward* areas. Indeed, consider the following example.

Example 2.1. *Suppose that $\nu = U_d$, so that Q is the identity. For every $u \in \mathbb{B}(0,1)\backslash\{0\}$, $\mathbb{E}[X|X \geq_\mathcal{R} Q(u)] = \mathbb{E}\left[X\big|\|X\| \geq \|u\|\right]$ is the expectation of a symmetric distribution over an annulus centered at the origin. Therefore, one has that $\mathbb{E}[X|X \geq_\mathcal{R} Q(u)] = 0$ and it cannot be thought of as a typical extreme observation.*

This is caused by a lack of information in $\mathbb{E}[X|X \geq_{\mathcal{R}} Q(u)]$. In fact, the rank function neglects the directional information of X, which lies in the sign function. In order to overcome this issue, we have to introduce a few notation. For any $u \in \mathbb{B}(0,1)\backslash\{0\}$, let

$$L_u = \left\{t\frac{u}{\|u\|} : t \in [0,1]\right\} \tag{2.4}$$

be a parametrization of the radius of $\mathbb{B}(0,1)\backslash\{0\}$ whose direction is $u/\|u\|$. The *sign curve* C_u associated to u is the image of L_u by the center-outward quantile function, namely

$$C_u = Q(L_u).$$

Several sign curves are represented in red at the right-hand side of Figure 2. Averaging observations less deep that $Q(u)$ along the sign curve C_u induces a "typical" value "beyond" $Q(u)$ in a meaningful multivariate way.

Definition 2.4. *Let $\nu \in \mathcal{P}_1(\mathbb{R}^d)$ with center-outward quantile function Q. The center-outward superquantile function of ν is the function S defined, for any $u \in \mathbb{B}(0,1)\backslash\{0\}$, by*

$$S(u) = \frac{1}{1 - \|u\|} \int_{\|u\|}^1 Q\left(t\frac{u}{\|u\|}\right)dt,$$

where the above integral is to be understood component-wise.

Remark 2.1 (Consistency with the univariate case)**.** *On the real line, there is only one sign curve $C_1 = \{Q(t); t \in [0,1]\}$. Thus, definition (1.3) can be seen as averaging observations less deep than $Q(\alpha)$, w.r.t. ν, along the sign curve C_1. Nonetheless, note that center-outward quantiles slightly differ from classical quantiles in dimension $d = 1$, because $U([-1,1]) \neq U([0,1])$, even if they carry the same information, [37][Appendix B].*

By the same token, we can define the center-outward expected shortfall function.

Definition 2.5. *Let $\nu \in \mathcal{P}_1(\mathbb{R}^d)$ with center-outward quantile function Q. The center-outward expected shortfall function of ν is the function E defined, for any $u \in \mathbb{B}(0,1)\backslash\{0\}$, by*

$$E(u) = \frac{1}{\|u\|} \int_0^{\|u\|} Q\left(t\frac{u}{\|u\|}\right)dt,$$

where the above integral is to be understood component-wise.

Remark 2.2. *For $\nu \in \mathcal{P}_*(\mathbb{R}^d)$, Q is neither defined at the origin nor at the boundary of the unit ball, unless the support of ν is compact. Hence the integrals in $S(u)$ and $E(u)$ are improper and they shall be understood respectively as*

$$\lim_{r \to 1^-} \int_{\|u\|}^r Q\left(t\frac{u}{\|u\|}\right)dt \qquad and \qquad \lim_{r \to 0^+} \int_r^{\|u\|} Q\left(t\frac{u}{\|u\|}\right)dt.$$

However, note that they are convergent as soon as $\nu \in \mathcal{P}_1(\mathbb{R}^d)$. In fact, the necessary assumption is rather the following integrability along sign curves,

$$\int_0^1 \left\| Q(t\frac{u}{\|u\|}) \right\| dt < \infty. \tag{2.5}$$

By the change of variables formula for push-forwards maps and by definition of U_d, $\mathbb{E}_\nu[\|X\|] = \mathbb{E}_{U_d}[\|Q(U)\|] = \mathbb{E}_{(R,\Phi)}[\|Q(R\Phi)\|]$. In other words, denoting by $\mathbb{P}_\mathbb{S}$ the uniform probability measure on the sphere $\mathbb{S}^{d-1} = \{\varphi \in \mathbb{R}^d : \|\varphi\| = 1\}$,

$$\mathbb{E}_\nu[\|X\|] = \int_{\mathbb{S}^{d-1}} \int_0^1 \|Q(t\varphi)\| dt \, d\mathbb{P}_\mathbb{S}(\varphi).$$

Thus, one can see by contradiction that, as soon as $\mathbb{E}_\nu[\|X\|] < \infty$, (2.5) holds for almost all $\varphi \in \mathbb{S}^{d-1}$, thus S and E must be finite almost everywhere.

The following naturally extends Definition 2.2.

Definition 2.6 (Superquantile and expected shortfall regions and contours)**.** *Let $\nu \in \mathcal{P}_1(\mathbb{R}^d)$ with center-outward superquantile function S and expected shortfall E. Then,*
(i) the superquantile (resp. expected shortfall) region $\mathbb{C}_\alpha^s$ (resp. $\mathbb{C}_\alpha^e$) of order $\alpha \in [0, 1]$ is the image by S (resp. E) of the ball $\mathbb{B}(0, \alpha)$.
(ii) the superquantile (resp. expected shortfall) contour $\mathcal{C}_\alpha^s$ (resp. $\mathcal{C}_\alpha^e$) of order $\alpha \in [0, 1]$ is the boundary of $\mathbb{C}_\alpha^s$ (resp. $\mathbb{C}_\alpha^e$).
(iii) averaged sign curves by E or S are respectively defined by $E(L_u)$ and $S(L_u)$ for any $u \in \mathbb{B}(0, 1)\backslash\{0\}$.

These concepts describe point clouds through periphery or central areas, and are illustrated in the numerical experiments carried out in Section 5.

2.2. Invariance properties

The two following lemmas are immediate consequences of [31, Lemmas A.7,A.8].

Lemma 2.1. *Assume that $X \in \mathbb{R}^d$ is an integrable random vector. Suppose that $a > 0$, $b \in \mathbb{R}^d$ and $Y = aX + b$. Denote by S_X, S_Y and E_X, E_Y their center-outward superquantile and expected shortfall functions. Then, for $u \in \mathbb{B}(0, 1)\backslash\{0\}$,*

$$S_Y(u) = aS_X(u) + b \quad and \quad E_Y(u) = aE_X(u) + b.$$

Proof. We only detail S_Y, as one can deduce E_Y identically. From [31][Lemma A.7],

$$\begin{aligned}
S_Y(u) &= \frac{1}{1 - \|u\|} \int_{\|u\|}^1 Q_Y(t\frac{u}{\|u\|})dt, \\
&= \frac{1}{1 - \|u\|} \int_{\|u\|}^1 \left(aQ_X(t\frac{u}{\|u\|}) + b\right)dt = aS_X(u) + b. \qquad \square
\end{aligned}$$

Lemma 2.2. *Assume that $X \in \mathbb{R}^d$ is an integrable random vector. Let S_X, S_Y and E_X, E_Y be the center-outward superquantile and expected shortfall functions associated respectively with X and $Y = AX$, for A an orthonormal matrix. Then, for $u \in \mathbb{B}(0,1)\backslash\{0\}$,*

$$S_Y(u) = AS_X(A^T u) \quad and \quad E_Y(u) = AE_X(A^T u).$$

Proof. Again, we only detail S_Y, as the proof for E_Y is identical. Combining [31][Lemma A.8] with the fact that $\|A^T u\| = \|u\|$ for an orthogonal matrix,

$$S_Y(u) = \frac{1}{1 - \|u\|} \int_{\|u\|}^1 Q_Y(t\frac{u}{\|u\|})dt,$$
$$= \frac{1}{1 - \|u\|} \int_{\|u\|}^1 AQ_X(tA^T\frac{u}{\|u\|})dt = AS_X(A^T u). \qquad \square$$

2.3. Main results

Quoting Rockafellar & Royset in [54],

> "the superquantile function [...] is as fundamental to a random variable as the distribution and quantile functions".

This assertion is partially motivated by the fact that the distribution, quantile and superquantile functions are uniquely determined one to another, and by the fact that pointwise convergence of these functions metrizes convergence in distribution. Such properties hold for our integrated concepts and are stated hereafter. First of all, we shall make repeated use of the fact that the center-outward superquantile and expected shortfall functions are two sides of the same coin, that is, for $u \in \mathbb{B}(0,1)\backslash\{0\}$,

$$\int_0^1 Q\Big(t\frac{u}{\|u\|}\Big)dt = \|u\|E(u) + (1 - \|u\|)S(u). \tag{2.6}$$

This is a generalization of the immediate property that, for the univariate setting given in the introduction with (1.1) and (1.2), we have for all $\alpha \in]0,1[$,

$$\mathbb{E}[X] = \alpha E(\alpha) + (1 - \alpha)S(\alpha).$$

Our first main result is as follows.

Theorem 2.2. *Let $\nu_1, \nu_2 \in \mathcal{P}_1(\mathbb{R}^d)$ with respective center-outward quantile, superquantile and expected shortfall functions denoted by Q_1, S_1, E_1 and Q_2, S_2, E_2. Then, the following are equivalent.*

(i) $\nu_1 = \nu_2$
(ii) $Q_1 = Q_2$ U_d-a.e.
(iii) $S_1 = S_2$ U_d-a.e.
(iv) $E_1 = E_2$ U_d-a.e.

Proof. First of all, it is already known that $(i) \Leftrightarrow (ii)$. Indeed, with the choice of U_d as the reference distribution, McCann's theorem [44] ensures that $(i) \Rightarrow (ii)$. Obviously, $(ii) \Rightarrow (i)$, from the very definition of push-forward measures that is $\nu_1(A) = U_d(Q_1^{-1}(A))$ and $\nu_2(A) = U_d(Q_2^{-1}(A))$. Hereafter, we proceed through the implication loop $(ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (ii)$. Let $\mathbb{P}_\mathbb{S}$ be the uniform probability measure on $\mathbb{S}^{d-1} = \{\varphi \in \mathbb{R}^d : \|\varphi\|_2 = 1\}$. Let R and φ be drawn uniformly on $[0,1]$ and $\mathbb{S}^{d-1}$, respectively. Suppose that they are independent, so that their joint distribution is given by $d\mathbb{P}_{(R,\varphi)}(r,\varphi) = d\mathbb{P}_R(r)d\mathbb{P}_\mathbb{S}(\varphi)$. There, by definition, if $f : (r,\varphi) \mapsto r\varphi$ on $[0,1] \times \mathbb{S}^{d-1}$, then $U_d = f_\# \mathbb{P}_{(R,\varphi)}$. Hence, from the change of variables formula for push-forwards, for any Borel set $A = f(B \times C)$, for $B \subset [0,1]$, $C \subset \mathbb{S}^{d-1}$,

$$\int_A (S_1 - S_2)(u)dU_d(u) = \int_B \int_C (S_1 - S_2)(r\varphi)d\mathbb{P}_\mathbb{S}(\varphi)dr,$$
$$= \int_B \int_C \Big(\frac{1}{1-r}\int_r^1 (Q_1 - Q_2)(t\varphi)dt\Big)d\mathbb{P}_\mathbb{S}(\varphi)dr. \quad (2.7)$$

Nonetheless, for any $r \in B$, by the same change of variables,

$$\int_C \int_r^1 (Q_1 - Q_2)(t\varphi)dt\,d\mathbb{P}_\mathbb{S}(\varphi) = \int_{f(C \times [r,1])} (Q_1 - Q_2)(u)dU_d(u). \quad (2.8)$$

Note that $U_d(f(C \times [r,1])) = (1-r)\mathbb{P}_\mathbb{S}(C)$, which is positive since $r < 1$. Obviously, (2.8) vanishes as soon as $Q_1 = Q_2$ U_d-a.e. It implies that (2.7) also vanishes, which justifies that $(ii) \Rightarrow (iii)$. Furthermore, we claim that $(iii) \Rightarrow (iv)$. Indeed, we have

$$\int_0^1 Q_1(t\frac{u}{\|u\|})dt = \lim_{r \to 0^+} S_1(r\frac{u}{\|u\|}) \quad \text{and} \quad \int_0^1 Q_2(t\frac{u}{\|u\|})dt = \lim_{r \to 0^+} S_2(r\frac{u}{\|u\|}).$$

Consequently, if we assume that $S_1 = S_2$ U_d-a.e., we obtain that

$$\int_0^1 (Q_1 - Q_2)(t\frac{u}{\|u\|})dt = 0. \quad (2.9)$$

Using (2.6), the desired result follows, $(iii) \Rightarrow (iv)$. Finally, assume that $E_1 = E_2$ U_d-a.e. Consequently, for all $r \in]0,1[$ and for all $\varphi \in \mathbb{S}^{d-1}$,

$$\int_0^r Q_1(t\varphi)dt = \int_0^r Q_2(t\varphi)dt.$$

Using that $\int_a^b = \int_0^b - \int_0^a$, for any $0 \le a \le b \le 1$,

$$\int_a^b Q_1(t\varphi)dt = \int_a^b Q_2(t\varphi)dt. \quad (2.10)$$

For any measurable $B \subset \mathbb{S}^{d-1}$, by integrating (2.10) w.r.t. $\mathbb{P}_\mathbb{S}$,

$$\int_B \int_a^b Q_1(t\varphi)d\mathbb{P}_R(t)d\mathbb{P}_\mathbb{S}(\varphi) = \int_B \int_a^b Q_2(t\varphi)d\mathbb{P}_R(t)d\mathbb{P}_\mathbb{S}(\varphi).$$

By use of $U_d = f_\# \mathbb{P}_{(R,\varphi)}$, a change of variables above yields, for any $A \subset \mathbb{B}(0,1)\backslash\{0\}$,

$$\int_A Q_1(u)dU_d(u) = \int_A Q_2(u)dU_d(u),$$

and the result follows. $\qquad\square$

Interestingly enough, our proposed integrated quantile functions are both simply related to the Kantorovich potential. Somehow, this development generalizes the work of [55] where the distribution function is related to the univariate superquantile by the way of the surexpectation function, which is nothing more than a particular primitive of the distribution function.

Proposition 2.1. *The center-outward expected shortfall function of $\nu \in \mathcal{P}_1(\mathbb{R}^d)$ satisfies, for any $u \in \mathbb{B}(0,1)\backslash\{0\}$,*

$$\langle E(u), u \rangle = \psi(u). \tag{2.11}$$

Moreover, the center-outward superquantile function of a compactly supported probability measure $\nu \in \mathcal{P}_*(\mathbb{R}^d)$ verifies, for every $u \in \mathbb{B}(0,1)\backslash\{0\}$,

$$\langle S(u), u \rangle = \frac{\|u\|}{1 - \|u\|}\Big(\psi(\frac{u}{\|u\|}) - \psi(u)\Big). \tag{2.12}$$

Proof. Fix $u \in \mathbb{B}(0,1)\backslash\{0\}$ and let $f(t) = \psi(tu)$. Then f is a finite and convex function on $]-a, a[$ for some $a > 1$, so that one can apply [52][Corollary 24.2.1]. Combined with *a.e.* differentiability, [52][Theorem 25.3], this yields

$$\psi(u) = \int_0^1 \langle \nabla\psi(tu), u \rangle dt, \tag{2.13}$$

which can be rewritten as

$$\psi(u) = \langle \int_0^1 Q(tu)dt, u \rangle = \langle E(u), u \rangle,$$

from a simple change of variables, leading to our first point (2.11). Moreover, denote for any $u \in \mathbb{B}(0,1)\backslash\{0\}$,

$$f(t) = \psi(t\frac{u}{\|u\|}).$$

This function is finite on $[0, 1]$ but not on a larger interval anymore. But, because $\nu \in \mathcal{P}_*(\mathbb{R}^d)$, f is differentiable everywhere on $]0, 1[$, and we have by the chain rule formula,

$$f'(t) = \langle Q(t\frac{u}{\|u\|}), \frac{u}{\|u\|} \rangle.$$

Since $f'(t)$ is a non-decreasing function from $\mathbb{R}$ to $[-\infty, +\infty]$ finite at $t = \|u\|$, [52][Theorem 24.2] ensures that the function F, defined for all $x \in \mathbb{R}$, by

$$F(x) = \int_{\|u\|}^x f'(t)dt,$$

is a well-defined closed proper convex function on $\mathbb{R}$. We emphasize that $F(x)$ takes infinite values as soon as $x > 1$, while, for $x = 1$, F is well-defined as a Lebesgue integral, or as a limit of Riemann integrals, as explained in the proof of [52][Theorem 24.2], and it may take finite or infinite values. In addition, the latter theorem tells us that $F(x) = f(x) + \alpha$ for some $\alpha \in \mathbb{R}$ everywhere. But $F(\|u\|) = 0$ and $f(\|u\|) = \psi(u)$, so $F(x) = f(x) - \psi(u)$. Assuming that the support $\mathcal{X}$ of ν is compact, ψ must be lipschitz continuous, [45][Lemma 2], a fortiori bounded and the subdifferential $\partial\psi(K)$ must also be bounded, see *e.g.* [53]. As a byproduct, $\psi(u/\|u\|)$ and $Q(u/\|u\|)$ are finite. In this case, $F(1) = f(1) - \psi(u)$ is finite and can be rewritten as

$$F(1) = \int_{\|u\|}^{1} \langle Q(t\frac{u}{\|u\|}), \frac{u}{\|u\|}\rangle dt = \psi(\frac{u}{\|u\|}) - \psi(u),$$

which implies (2.12), completing the proof of Proposition 2.1. $\qquad\square$

Thanks to this relation between the potential ψ whose gradient gives Q and the center-outward expected shortfall function, we are now able to characterize the convergence in distribution for a sequence of random vectors through superquantiles and expected shortfalls. For that purpose, we rely on existing results on the relation between convergence in distribution and center-outward quantile functions.

Lemma 2.3. *Let* (X, X_n) *be a sequence of random vectors with distributions* ν *and* ν_n *respectively, in* $\mathcal{P}_*(\mathbb{R}^d)$*. Then,*

$$X_n \overset{\mathcal{L}}{\to} X \quad \Leftrightarrow \quad \forall u \in \mathbb{B}(0,1), \ \lim_{n\to\infty} \psi_n(u) = \psi(u), \tag{2.14}$$

$$\Leftrightarrow \quad \lim_{n\to\infty} \sup_{u\in K} |\psi_n(u) - \psi(u)| = 0, \tag{2.15}$$

$$\Leftrightarrow \quad \lim_{n\to\infty} \sup_{u\in K} \|Q_n(u) - Q(u)\| = 0, \tag{2.16}$$

for every compact $K \subset \mathbb{B}(0,1)\backslash\{0\}$*. In fact, uniform convergence of* ψ_n *towards* ψ *even holds on every compact* $K \subset \mathbb{B}(0,1)$*.*

Proof. On the one hand, assume that (X_n) converges in distribution to X. Then, the right-hand side of (2.16) is the main result of [31][Theorem 4.1] when Q_n and Q are homeomorphisms between convex sets, with uniform convergence on any compact subset. But, with reference distribution U_d, center-outward quantile maps are not, because of the discontinuity at the origin, see [37][Remark 3.4]. Nonetheless, as highlighted in the proof of [37][Proposition 3.3], the assumption that $X_n \overset{\mathcal{L}}{\to} X$ for $\nu_n, \nu \in \mathcal{P}_*(\mathbb{R}^d)$ is sufficient to apply Theorem 2.8 in [22]. As a consequence, $\psi_n(u)$ converges towards $\psi(u)$ for every $u \in \mathbb{B}(0,1)$, that is the right-hand side of (2.14). Being finite and convex functions, [52][Theorem 10.8] ensures that such pointwise convergence implies the right-hand side of (2.15). Since ψ_n and ψ are differentiable on $\mathbb{B}(0,1)\backslash\{0\}$ as soon as $\nu_n, \nu \in \mathcal{P}_*(\mathbb{R}^d)$, one can apply [52][Theorem 25.7] on any open and convex set $K' \subset \mathbb{B}(0,1)\backslash\{0\}$. This gives us the uniform convergence of Q_n towards Q for any compact K included

in an open and convex set $K' \subset \mathbb{B}(0,1)\backslash\{0\}$. To extend this to any compact K, we use that, by compacity, one can extract from the open cover $\{\mathbb{B}(x,\delta); x \in K\}$ a finite cover. Consequently, it exists $N \in \mathbb{N}$ and $x_1, \cdots, x_N \in K$ such that

$$\sup_{u \in K} \|Q_n(u) - Q(u)\| \leq \sum_{k=1}^{N} \sup_{u \in \mathbb{B}(x_k,\delta)} \|Q_n(u) - Q(u)\|. \tag{2.17}$$

Here, the closure of each ball $\mathbb{B}(x_k, \delta)$ is a compact set, which, because of the choice of δ, is a subset of some open and convex set $K' \subset \mathbb{B}(0,1)\backslash\{0\}$. But we already know that Q_n uniformly converges towards Q on such a set, thus the right-hand side of (2.17) vanishes when $n \to 0$, which yields the right-hand side of (2.16). It only remains to prove that $X_n \xrightarrow{\mathcal{L}} X$ as a direct consequence. This last claim relies on the Portmanteau theorem which says that (X_n) converges in distribution to X iff for any bounded and continuous function f,

$$\lim_{n \to +\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]. \tag{2.18}$$

However, we clearly have for any bounded and continuous function f,

$$\mathbb{E}[f(X_n)] = \int_{\mathbb{R}^d} f(x)d\nu_n(x) = \int_{\mathbb{B}(0,1)} f(Q_n(u))dU_d(u).$$

using the change of variables $\nu_n = Q_{n\#}U_d$. Hence, as $f \circ Q_n$ is uniformly bounded, the dominated convergence theorem leads to (2.18). $\square$

Theorem 2.3. *Let (X, X_n) be a sequence of random vectors with distributions ν and ν_n respectively, in $\mathcal{P}_*(\mathbb{R}^d)$. Then,*

$$X_n \xrightarrow{\mathcal{L}} X \quad \Leftrightarrow \quad \forall u \in \mathbb{B}(0,1)\backslash\{0\}, \lim_{n \to +\infty} E_n(u) = E(u), \tag{2.19}$$

$$\Leftrightarrow \quad \lim_{n \to \infty} \sup_{u \in K} \|E_n(u) - E(u)\| = 0, \tag{2.20}$$

for every compact $K \subset \mathbb{B}(0,1)\backslash\{0\}$.

Proof. On the one hand, assume that (X_n) converges in distribution to X. Then, it follows from (2.16) that (Q_n) converges uniformly to Q on any compact $K \subset \mathbb{B}(0,1)\backslash\{0\}$. For any $u \in \mathbb{B}(0,1)\backslash\{0\}$, we have from Definition 2.5 that

$$\|E_n(u) - E(u)\| \leq \frac{1}{\|u\|} \int_0^{\|u\|} R_n(t\frac{u}{\|u\|})dt,$$

where $R_n(v) = \|Q_n(v) - Q(v)\|$. Fix a compact $K \subset \mathbb{B}(0,1)\backslash\{0\}$ such that for every $u \in K$, $C < \|u\| < D$ for some positive constants C, D. Then, we clearly have

$$\sup_{u \in K} \|E_n(u) - E(u)\| \leq \frac{1}{C} \int_0^D \sup_{u \in K} R_n(t\frac{u}{\|u\|})dt.$$

Consequently, for any $\xi \in \mathbb{R}$ such that $0 \leq \xi < D$,

$$\sup_{u \in K} \|E_n(u) - E(u)\| \leq \frac{1}{C} \left(\int_0^\xi \sup_{u \in K} R_n(t\frac{u}{\|u\|})dt + \int_\xi^D \sup_{u \in K} R_n(t\frac{u}{\|u\|})dt \right). \quad (2.21)$$

On the one hand, for all $t \in [\xi, D]$ and $u \in K$, $tu/\|u\|$ lies inside a compact $K' \subset \mathbb{B}(0,1)\backslash\{0\}$, so that the second term in (2.21) satisfies

$$\int_\xi^D \sup_{u \in K} R_n(t\frac{u}{\|u\|})dt \leq (D - \xi) \sup_{v \in K'} R_n(v) \leq \sup_{v \in K'} R_n(v), \quad (2.22)$$

which vanishes when $n \to +\infty$. On the other hand,

$$\int_0^\xi \sup_{u \in K} R_n(t\frac{u}{\|u\|})dt \leq \xi \sup_{v \in \overline{\mathbb{B}}(0,D)} R_n(v). \quad (2.23)$$

We now claim that $\sup_{v \in \overline{\mathbb{B}}(0,D)} R_n(v)$ is bounded by a constant. Recall that, for any $v \in \overline{\mathbb{B}}(0,D)$, $Q_n(v)$ and $Q(v)$ belong to the subdifferentials $\partial \psi_n(v)$ and $\partial \psi(v)$, by definition. Because $S = \overline{\mathbb{B}}(0,D)$ is a compact subset of the open unit ball, and because, for any $u \in S$, $\psi_n(u)$ is convergent and a fortiori bounded, see Lemma 2.3, we are in position to apply [52][Theorem 10.6]. It directly implies that the sequence (ψ_n) is uniformly bounded and equi-Lipschitz on S. However, it is well-known that any L-lipschitz convex function ϕ on S must have a bounded subdifferential $\partial \phi(S)$, see *e.g.* [53]. Indeed, for any $u \in S$ and any $\xi \in \partial \phi(S)$, consider $\delta > 0$ such that $y = u - \delta\xi$ belongs to S. By definition of $\partial \phi(S)$ and the lipschitz constant L,

$$L\delta\|\xi\| \geq \phi(y) - \phi(u) \geq \langle \xi, y - u \rangle = \delta\|\xi\|^2.$$

This immediately gives us the existence of a uniform bound on the family of subdifferentials $(\partial \psi_n(S))_n$. Because $\partial \psi(S)$ is also bounded, by lipschitz regularity of ψ on S,

$$\sup_{v \in \overline{\mathbb{B}}(0,D)} R_n(v) \leq \sup_{v \in \overline{\mathbb{B}}(0,D)} \|Q_n(v)\| + \|Q(v)\| \leq M < +\infty. \quad (2.24)$$

Therefore the right-hand side of (2.23) can be bounded by ξM for M given in (2.24). Then it follows from (2.21) and (2.22) that, for any $\xi \in [0, D[$, it exists $n_1 \in \mathbb{N}$ such that, for all $n \geq n_1$,

$$\sup_{u \in K} \|E_n(u) - E(u)\| \leq \frac{1}{C}\Big(\xi M + \sup_{v \in K'} R_n(v)\Big).$$

Thus, for any $\xi \in [0, D[$,

$$\lim_{n \to +\infty} \sup_{u \in K} \|E_n(u) - E(u)\| \leq \frac{\xi M}{C}.$$

which leads to the right-hand side of (2.20) as ξ goes to zero. In addition, the right-hand side of (2.20) immediately implies the right-hand side of (2.19). On the other hand, assume that the right-hand side of (2.19) holds. Then, we obtain from (2.11) that $\forall u \in \mathbb{B}(0, 1)$,

$$\lim_{n \to +\infty} \psi_n(u) = \psi(u).$$

Thus, the desired result follows from Lemma 2.3, which completes the proof of Theorem 2.3. $\qquad\square$

Our last result requires some uniform integrability assumption on the sequence (X_n).

Theorem 2.4. *Let (X, X_n) be a sequence of random vectors with distributions ν and ν_n respectively, in $\mathcal{P}_*(\mathbb{R}^d)$. In addition, suppose that there exists a random variable Z greater than 1 such that $\mathbb{E}[Z \ln(Z)] < +\infty$ and for all $n \in \mathbb{N}$,*

$$\|X_n\| \leq Z \qquad a.s. \tag{2.25}$$

Then,

$$X_n \overset{\mathcal{L}}{\to} X \Leftrightarrow \lim_{n \to \infty} S_n(u) = S(u), \tag{2.26}$$

for almost all $u \in \mathbb{B}(0, 1)$.

Remark 2.3. *One can observe that if the support of every ν_n is bounded by some constant M, (2.25) is no longer needed. Under this restrictive assumption, there is no mass going out to infinity in any direction and one can easily check that*

$$X_n \overset{\mathcal{L}}{\to} X \Leftrightarrow \lim_{n \to \infty} \sup_{u \in K} \|S_n(u) - S(u)\| = 0,$$

for every compact $K \subset \mathbb{B}(0, 1)\backslash\{0\}$.

Proof. On the one hand, assume that (X_n) converges in distribution to X. We clearly have from (2.25) that, if $\Phi(x) = x \ln(x)$, then

$$\mathbb{E}[\sup_{n \in \mathbb{N}} \Phi(X_n)] \leq \mathbb{E}[Z \ln(Z)] < +\infty.$$

Hence, using the same change of variables as in the proof of Theorem 2.2, we obtain that

$$\mathbb{E}[\sup_{n \in \mathbb{N}} \Phi(X_n)] = \int_{\mathbb{S}^{d-1}} \int_0^1 \sup_{n \in \mathbb{N}} \|Q_n(t\varphi)\| \ln(\|Q_n(t\varphi)\|) dt d\mathbb{P}_{\mathbb{S}}(\varphi) < +\infty, \tag{2.27}$$

where we recall that $\mathbb{S}^{d-1} = \{\varphi \in \mathbb{R}^d : \|\varphi\|_2 = 1\}$. Consequently, we deduce from (2.27) that for almost all $\varphi \in \mathbb{S}^{d-1}$,

$$\sup_{n \in \mathbb{N}} \int_0^1 \|Q_n(t\varphi)\| \ln(\|Q_n(t\varphi)\|) dt \leq \int_0^1 \sup_{n \in \mathbb{N}} \|Q_n(t\varphi)\| \ln(\|Q_n(t\varphi)\|) dt < +\infty. \tag{2.28}$$

It clearly leads to the uniform integrability of (Q_n) along sign curves, see [9][Theorem 4.5.9]. However, we already saw from Lemma 2.3 that the convergence in distribution of (X_n) to X implies the pointwise convergence of Q_n to Q on $\mathbb{B}(0,1)\backslash\{0\}$. Therefore, it follows from the Lebesgue-Vitali theorem [9][Theorem 4.5.4] that for almost all $\varphi \in \mathbb{S}^{d-1}$,

$$\lim_{n\to+\infty} \int_0^1 Q_n(t\varphi)dt = \int_0^1 Q(t\varphi)dt,$$

which means that for almost all $u \in \mathbb{B}(0,1)\backslash\{0\}$,

$$\lim_{n\to+\infty} \int_0^1 Q_n\Big(t\frac{u}{\|u\|}\Big)dt = \int_0^1 Q\Big(t\frac{u}{\|u\|}\Big)dt. \tag{2.29}$$

Hereafter, we already saw from (2.6) that

$$S_n(u) - S(u) = \frac{1}{1-\|u\|}\left(\int_0^1 (Q_n - Q)\Big(t\frac{u}{\|u\|}\Big)dt - \|u\|\Big(E_n(u) - E(u)\Big)\right). \tag{2.30}$$

Finally, the right-hand side of (2.26) follows from (2.19) together with (2.29) and (2.30). On the other hand, assume that the right-hand side of (2.26) holds. We have for almost all $u \in \mathbb{B}(0,1)\backslash\{0\}$,

$$\lim_{n\to+\infty} \int_{\|u\|}^1 Q_n\Big(t\frac{u}{\|u\|}\Big)dt = \int_{\|u\|}^1 Q\Big(t\frac{u}{\|u\|}\Big)dt. \tag{2.31}$$

Hence, we obtain from (2.31) that for almost every $r \in]0,1[$ and $\varphi \in \mathbb{S}^{d-1}$,

$$\lim_{n\to+\infty} \int_r^1 Q_n(t\varphi)dt = \int_r^1 Q(t\varphi)dt. \tag{2.32}$$

Obviously, for almost all $r \in]0,1[$ and almost all $\varphi \in \mathbb{S}^{d-1}$,

$$\lim_{n\to+\infty} \int_0^1 (Q_n - Q)(t\varphi)dt = \lim_{n\to+\infty}\left(\int_0^r (Q_n - Q)(t\varphi)dt + \int_r^1 (Q_n - Q)(t\varphi)dt\right),$$
$$= \lim_{n\to+\infty} \int_0^r (Q_n - Q)(t\varphi)dt.$$

In addition, this integral is always finite since (2.5) holds for integrable probability measures. As the left-hand side does not depend on r,

$$\lim_{n\to+\infty} \int_0^1 (Q_n - Q)(t\varphi)dt = \lim_{r\to 0}\lim_{n\to+\infty} \int_0^r (Q_n - Q)(t\varphi)dt. \tag{2.33}$$

From the uniform integrability of (Q_n) along sign curves, we obtain that

$$\lim_{r\to 0}\lim_{n\to+\infty} \int_0^r \|Q_n(t\varphi) - Q(t\varphi)\|dt = 0.$$

Hence, by use of (2.33), we find that

$$\lim_{n \to +\infty} \int_0^1 Q_n(t\varphi)dt = \int_0^1 Q(t\varphi)dt.$$

It ensures that for almost all $u \in \mathbb{B}(0,1)\backslash\{0\}$,

$$\lim_{n \to +\infty} \int_0^1 Q_n\Big(t\frac{u}{\|u\|}\Big)dt = \int_0^1 Q\Big(t\frac{u}{\|u\|}\Big)dt. \tag{2.34}$$

Consequently, we deduce from (2.30) and (2.34) that for a.e. $u \in \mathbb{B}(0,1)\backslash\{0\}$,

$$\lim_{n \to +\infty} E_n(u) = E(u) \tag{2.35}$$

Finally, it follows from (2.35) together with (2.19) that (X_n) converges in distribution to X, which completes the proof of Theorem 2.4. □

3. A class of reference measures

Until now, the concepts of MK quantiles, superquantiles and expected shortfalls strongly rely on the choice of the reference distribution, and one may question the choice of U_d. This issue is discussed hereafter, with explicit MK quantile functions for generalized gamma models, by calibrating the reference distribution. Ultimately, one obtains nested regions indexed by their probability content. The choice of the reference distribution μ is a major tool to adapt MK quantiles, ranks and signs to any task encountered. Notably, the uniform distribution over the unit hypercube has advantages concerning marginal independence, [19, 31], whereas orthogonal invariance only holds with a spherical reference distribution, [31, 38]. U_d plays a specific role for the interpretation of quantile regions indexed by a probability content level. This section deals with an alternative class of reference measures defined by α-level sets on the p-unit ball, that preserve the interpretability of U_d. We also consider the restriction of U_d to $\mathbb{R}_+^d$, with a *left-to-right* ordering that might be preferred in the subsequent risk applications to the *center-outward* one. Denote by $\mathbb{S}^{d,p}$ the unit sphere $\mathbb{S}^{d,p} = \{\varphi \in \mathbb{R}^d : \|\varphi\|_p = 1\}$ with

$$\|\varphi\|_p^p = \sum_{k=1}^d |\varphi_k|^p,$$

and $\mathbb{S}_+^{d,p} = \mathbb{S}^{d,p} \cap \mathbb{R}_+^d$. Moreover, let q be the Hölder conjugate of p, given by

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Definition 3.1. *The q-spherical conjugate distribution $U_{d,q}$ is defined as the product $R\Phi$ between two independent random variables R and Φ where*

- *R has uniform distribution on $[0,1]$,*

- *the distribution of $\Phi \in \mathbb{S}^{d,q}$ is given by $\Phi = \Psi^{\otimes(p-1)}$ for Ψ uniformly drawn on $\mathbb{S}^{d,p}$, and $\otimes$ the component-wise exponent.*

The restriction of $U_{d,q}$ to the convex cone $\mathbb{R}^d_+$ is denoted by $U^+_{d,q}$.

With reference measure $U_{d,q}$, the MK quantile function Q_q can always be defined almost everywhere from a convex potential ψ by Definition 2.1, for which we can always assume that $\psi(0) = 0$. At any point u of non differentiability of ψ, one can still define $Q_q(u)$ as the average of the subdifferential $\partial\psi(u)$, so that Q_q is defined everywhere on the q-unit ball. The interpretability of the quantile concepts remains the same among this class, as the $U_{d,q}$-probability of the q-ball of radius $\alpha \in [0,1]$ is α,

$$\mathbb{P}(\|R\Phi\|_q \leq \alpha) = \mathbb{P}(R \leq \alpha) = \alpha.$$

Crucially for us, our definitions and properties of superquantiles and expected shortfalls naturally adapt. Definition 2.4 as well as Definition 2.5 extend, for all $u \in \mathbb{B}(0,1)$, to

$$S_q(u) = \frac{1}{1 - \|u\|_q} \int_{\|u\|_q}^1 Q_q(t\frac{u}{\|u\|_q})dt$$

and

$$E_q(u) = \frac{1}{\|u\|_q} \int_0^{\|u\|_q} Q_q(t\frac{u}{\|u\|_q})dt.$$

We now develop on how to sample uniformly on p-spheres $\mathbb{S}^{d,p}$. Let Γ stand for the Euler Gamma function and denote by $\mathcal{L}_p$ the probability distribution with density function

$$f_p(x) = \frac{p}{\Gamma(p^{-1})} \exp(-x^p) I_{\mathbb{R}_+}(x).$$

A random variable X on $\mathbb{R}_+$ is drawn from $\mathcal{L}_p$ if and only if X^p follows a Gamma distribution with shape and scale parameters $1/p$ and 1. The following lemma indicates how to sample uniformly on p-spheres when $p > 0$.

Lemma 3.1 ([4, 56]). *For any real $p > 0$, the components of a random vector $X \in \mathbb{R}^d_+$ are independent with distribution $\mathcal{L}_p$ if and only if the random variables $\|X\|_p$ and $\|X\|_p^{-1}X$ are independent where*

- *$\|X\|_p^{-1}X$ is uniformly distributed on $\mathbb{S}^{d,p}_+$,*

- *$\|X\|_p^p$ has Gamma distribution with shape and scale parameters d/p and 1.*

For $p = 2$, $\mathcal{L}_p$ corresponds to $|Z|$ where Z is drawn from a $\mathcal{N}(0,1)$ distribution. The MK distribution function with reference U_d (or $U_{d,2}$) has been obtained in [13][Section 2.4]. Here, we use similar arguments to find gradient-of-convex maps for our alternative class of reference measures. We emphasize that such maps are invariant to shifts, see [31][Lemma A.7], whereas invariance to rotations requires a spherically symmetric reference measure, [31][Lemma A.8].

Proposition 3.1. *Suppose that the components of a random vector $X \in \mathbb{R}_+^d$ are independent with $\mathcal{L}_p$ distribution for $p > 1$, and let $q = p/(p-1)$. Then, the MK distribution function with respect to $U_{d,q}^+$ has the explicit formulation*

$$F_q(x) = (\|x\|_p^{-1} x)^{\otimes(p-1)} G(\|x\|_p), \tag{3.1}$$

where $\otimes$ stands for the component-wise exponent and G is the univariate distribution function of $\|X\|_p$. Its inverse, the MK quantile function, is given, for all $u \in \mathbb{R}_+^d$ such that $\|u\|_q \leq 1$, by

$$Q_q(u) = (\|u\|_q^{-1} u)^{\otimes(q-1)} G^{-1}(\|u\|_q). \tag{3.2}$$

The MK superquantile and expected shortfall functions are respectively given, for all $u \in \mathbb{R}_+^d$ such that $\|u\|_q \leq 1$, by

$$S_q(u) = (\|u\|_q^{-1} u)^{\otimes(q-1)} \overline{S}(\|u\|_q), \tag{3.3}$$

$$E_q(u) = (\|u\|_q^{-1} u)^{\otimes(q-1)} \overline{E}(\|u\|_q), \tag{3.4}$$

where $\overline{S}$ and $\overline{E}$ are the univariate superquantile and expected shortfall functions associated with the distribution of $\|X\|_p$.

Proof. Denote by G the probability distribution function of $\|X\|_p$. Moreover, for $z \in \mathbb{R}$, let

$$\Psi(z) = \int_{-\infty}^z G(t) dt,$$

and $\varphi(x) = \Psi(\|x\|_p)$. Then, φ is convex by the composition between the non-decreasing function Ψ and $\|\cdot\|_p$, both convex, see [54][Theorem 5.1]. In addition, $\nabla\varphi(x) = F_q(x)$, which means that F_q is the gradient of a convex function. It remains to show that $F_q(X)$ follows the distribution $U_{d,q}^+$. From Lemma 3.1, $F_q(X)$ is the product of two independent random variables, $G(\|X\|_p)$ and $(\|X\|_p^{-1} X)^{\otimes(p-1)}$. On the one hand, $G(\|X\|_p)$ is uniformly distributed on $[0, 1]$, by definition of G. On the other hand, the distribution of $\|X\|_p^{-1} X$ is uniform on $\mathbb{S}_+^{d,p}$, which implies (3.1). In order to compute $Q_q = F_q^{-1}$, it remains to invert (3.1). If $u = (\|x\|_p^{-1} x)^{\otimes(p-1)} G(\|x\|_p)$,

$$\|u\|_q^q = G(\|x\|_p)^q \sum_{k=1}^d \left(\frac{x_k}{\|x\|_p} \right)^p.$$

Thus, $\|u\|_q = G(\|x\|_p)$ which yields $\|u\|_q^{-1} u = (\|x\|_p^{-1} x)^{\otimes(p-1)}$. This rewrites $\|u\|_q^{-1} u \|x\|_p^{p-1} = x^{\otimes(p-1)}$, where $\|x\|_p^{p-1} = G^{-(p-1)}(\|u\|_q)$, so

$$\|u\|_q^{-1} u G^{-(p-1)}(\|u\|_q) = x^{\otimes(p-1)}.$$

Finally, (3.2) follows by applying the exponent $q - 1 = 1/(p-1)$, which implies (3.3) and (3.4), completing the proof of Proposition 3.1. $\qquad\square$

Remark 3.1. *It is well-known that in the special case $p = 2$, the distribution of $(\|X\|_p^{-1} X)^{\otimes(p-1)}$ is uniform on $\mathbb{S}_+^{d,q}$ where $q = p = 2$, see [13]. One may wonder if this property holds true for other choices of $p > 1$. This is actually not the case as we shall see now. As q is the Hölder conjugate of p, we clearly have $\|X\|_p^{p-1} = \|X^{\otimes(p-1)}\|_q$ which implies that*

$$\left(\frac{X}{\|X\|_p} \right)^{\otimes(p-1)} = \frac{X^{\otimes(p-1)}}{\|X^{\otimes(p-1)}\|_q}.$$

All the components of X share the same $\mathcal{L}_p$ distribution. Consequently, each X_i^p has a Gamma distribution with shape and scale parameters $1/p$ and 1. It implies that $X_i^{p-1} = X_i^{p/q}$ is the power $1/q$ of a Gamma distribution. We immediately deduce from Lemma 3.1 that, as soon as $p \neq 2$, $(\|X\|_p^{-1} X)^{\otimes(p-1)}$ is not uniformly distributed on $\mathbb{S}_+^{d,q}$.

Figure 3 illustrates Proposition 3.1. Each column contains the reference measure $U_{d,q}^+$ for $q = p/(p-1)$ in the first line, and the associated distribution with *i.i.d.* components $X_i \sim \mathcal{L}_p$ below. Reference contours of orders $0.25, 0.5, 0.75$ are represented as well as the explicit MK quantile contours. Whereas the center-outward ordering intrinsic to U_d is natural for elliptical models, [13][Section 2.4], the left-to-right ordering may be more relevant under generalized Gamma models in $\mathbb{R}_+^d$.

4. Multivariate values and vectors at risk

Considering several univariate risks separately leads to neglecting the correlation structure between the components of a random vector. Hence, studying a notion of risk in a multivariate way is of major importance. We refer to [26][Section 2.1] for motivations for intrinsically multivariate risk analysis problems. The risk framework considers a vector of dependent losses $X \in \mathbb{R}_+^d$, where each component is a positive measure whose unit may be different from one to another. Naturally, the larger the value of a component, the greater the associated risk. On the one hand, we argue that real-valued risk measures are not sufficient in such setting. With real-valued random variables, real-valued risk measures are able to catch all the information needed. But, with the increase of the dimension, one may need vectors in $\mathbb{R}^d$ to get the directional information of the multivariate tails. With this in mind, we make here a proposal for vector-valued risk measures, namely *Vectors-at-Risk* and *Conditional-Vectors-at-Risk*, which aim to sum up the relevant information contained in the center-outward quantiles and superquantiles. Related to these vector-valued measures, we define *multivariate values-at-risk* and *multivariate conditional-values-at-risk* which are measures of the form of $\rho : \mathbb{R}_+^d \to \mathbb{R}$ able to compare X and Y by $\rho(X) > \rho(Y)$. Importantly, the computation of these real-valued measures gives the vector-valued ones, hitting two targets with one shot without added complexity. On the other hand, a multivariate extension of the concepts of VaR and CVaR shall have the same interpretation as in $\mathbb{R}$. In dimension $d = 1$, the VaR of some risk at a level α is

(a) $U_{d,q}^+$, $q = 3$ (b) $U_{d,q}^+$, $q = 2$ (c) $U_{d,q}^+$, $q = 5/4$

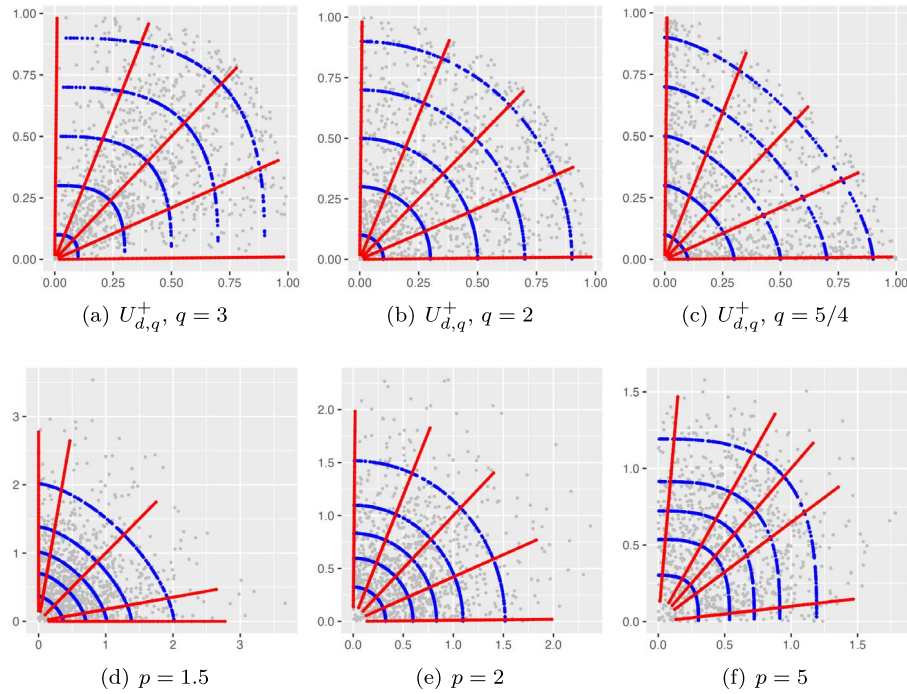(d) $p = 1.5$ (e) $p = 2$ (f) $p = 5$

FIG 3. *(First line) Reference measures $U_{d,q}^+$ for indices q conjugate with p. (Second line) Generalized Gamma distributions $\mathcal{L}_p$ and explicit quantile contours and sign curves.*

simply the quantile of order α. Accordingly, the CVaR is the superquantile of order α. These are to be understood respectively as the worst risk encountered with ν-probability α, and as the averaged risk beyond this quantile. Being an observation from the underlying distribution, it shall be vector-valued in our multivariate framework.

By construction, the MK quantile contour of order $\alpha \in [0, 1]$ contains the points having the most outward position with ν-probability α. Hence, we argue that a Vector-at-Risk of order α should belong to this contour. Even if it is always adapted to the geometry of ν, a MK quantile contour can either describe a *central* or a *bottom-left* area, by changing the reference measure. We explore both these possibilities, leaving it to the reader to decide which tool to adopt. In what follows, we rely on U_d and U_d^+, its restriction to $\mathbb{R}_+^d$. Of course, other p-spherical uniform distributions could be used, but we do not know if this would induce any benefit. Starting from the choice between *central* or *bottom-left* areas for MK quantile contours, the worst vectors of losses are the furthest from the origin in our context. Then, we suggest to select points from the center-outward quantile contour of order α with maximal norm, using Definition 2.2. *Conditional-Vectors-at-Risk* (CVaRs) are defined in the same way, but considering the center-outward superquantiles, with Definition 2.6.

Definition 4.1. *The Vector-at-Risk at level α is defined by*

$$VaR_\alpha(X) \in \operatorname{argsup} \{ \|X\|_1 \; ; \; X \in \mathcal{C}_\alpha \} .$$

Similarly, the Conditional-Vector-at-Risk is defined by

$$CVaR_\alpha(X) \in \operatorname{argsup} \{ \|X\|_1 \; ; \; X \in \mathcal{C}_\alpha^s \} .$$

The choice of $\|\cdot\|_1$ in the above definition depends on how to compare different risks. We emphasize that this choice is very distinct to the one of the reference measure: this consideration takes place once the contours are fixed. Without added information, we believe that two observations of same 1-norm shall be considered of same importance. Intuitively, when comparing multivariate risks, the risks per component add up. For example, if

$$x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad x_2 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

belong to the same quantile contour of level α, choosing $\|\cdot\|_2$ instead of $\|\cdot\|_1$ in Definition 4.1 would induce to consider that x_1 is worst than x_2. This choice is coherent with the common practice of computing univariate VaR and CVaR on the random variable $\|X\|_1$ in financial applications, but a major difference is that it encodes the multivariate joint probability of the vector X before applying the sum. Thus, it takes into account the correlations, while providing more information, because typical values for component-wise risks can be retrieved through our Vectors-at-Risk. With this in mind, the Vector-at-Risk of order α is to be understood as the "worst" risk encountered with probability α, whereas the Conditional-Vector-at-Risk is the average risk beyond this "worst" observation, where the notion of "worst" is characterized here by $\|\cdot\|_1$ and the focus is either on central or on bottom-left areas. Thus, this generalizes the understanding of univariate VaR and CVaR. We displayed in Figure 4 the VaR_α and $CVaR_\alpha$ with quantile and superquantile contours estimated via the numerical scheme presented in Section 5.1 for $\varepsilon = 10^{-3}$. Each red point is the worst observation inside the corresponding quantile or superquantile regions.

Intuitively, Definition 4.1 already gives a real-valued risk measure. Note that the set $\{ \|X\|_1 \; ; \; X \in \mathcal{C}_\alpha \}$ is the same as $\{ \|Q_X(u)\|_1 \; ; \; \|u\|_2 = \alpha \}$ for both the reference distribution U_d and its restriction U_d^+ to $\mathbb{R}_+^d$.

Definition 4.2. *The multivariate value-at-risk at level α is defined as*

$$\rho_\alpha^Q(X) = \sup_u \{ \|Q_X(u)\|_1 \; ; \; \|u\|_2 = \alpha \} .$$

Similarly, the multivariate conditional value-at-risk is defined as

$$\rho_\alpha^S(X) = \sup_u \{ \|S_X(u)\|_1 \; ; \; \|u\|_2 = \alpha \} .$$

These *multivariate (conditional) values-at-risk* indicate on the spreadness of the point cloud drawn from a vector of losses, and contain information about its centrality and the correlations between its components. Empirical experiments have been conducted in Section 5.2 where they are compared with the *maximal correlation risk measure* from [5].
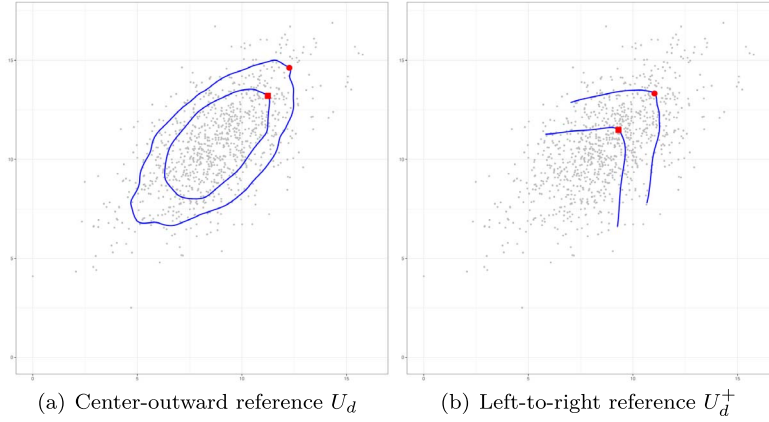
(a) Center-outward reference U_d (b) Left-to-right reference U_d^+

FIG 4. *Gaussian distribution with VaR_α and $CVaR_\alpha$ in red and associated MK quantile and superquantile contours in blue, for $\alpha = 0.5$.*

5. Numerical experiments

5.1. On the choice of the empirical quantile map

Hereafter, we chose to use the entropic map as a regularized empirical map, as advocated in [8, 12, 43]. Regularizing the quantile function enables to reach a smooth function $T : u \mapsto T(u)$ in practice, and has been the concern of several works, including [5, 37]. Our regularization choice is thus motivated by the known computational advantages of entropic optimal transport, [17, 30], with a very recent line of work focusing on the entropic map [33, 48, 49, 51]. A relaxed version of Monge problem (2.1) is the Kantorovich problem, see the introduction of [61]. This allows for easier computations, even more so when one uses entropically regularized optimal transport. We only state here the semi-dual formulation of Kantorovich regularized problem, that one can find for instance in [30]. The notation differ from traditional ones in optimal transport theory, to be consistent with the context of center-outward quantiles. Let $\mathcal{C}(\mathcal{X})$ be the space of continuous functions from $\mathcal{X}$ to $\mathbb{R}$. Let $\varepsilon > 0$ be the regularization parameter, which must be low in order to approximate better the true OT. For $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ with finite second-order moments and with respective supports $\mathcal{U}, \mathcal{X}$, and for a given $x_0 \in \mathcal{X}$, the semi-dual problem has a unique solution and writes

$$\max_{v \in \mathcal{C}(\mathcal{X}):v(x_0)=0} \int_{\mathcal{U}} v^{c,\epsilon}(u)d\mu(u) + \int_{\mathcal{X}} v(x)d\nu(x) - \epsilon, \tag{5.1}$$

where $v^{c,\epsilon} \in \mathcal{C}(\mathcal{U})$ is the smooth c-transform of v

$$\forall u \in \mathcal{U}, v^{c,\epsilon}(u) = -\epsilon \log \Big(\int_{\mathcal{X}} \exp \Big(\frac{v(x) - \frac{1}{2}\|u - x\|^2}{\epsilon} \Big) d\nu(x) \Big). \tag{5.2}$$

In practice, ν, the law whose we consider the quantiles, is not known and is approached by its empirical counterpart, so it always has a finite moment of order 2. On the contrary, μ, the reference law, is known and absolutely continuous. In this semi-discrete setting, one can use stochastic algorithms [7, 8, 30] to solve (5.1) and obtain the optimal Kantorovich potential v. From this, one can deduce the entropic map, [49], an approximation of the Monge map partially legitimated by an entropic analog of Brenier's theorem, [10, 16]

$$Q_\varepsilon(u) = \nabla\Big(\frac{1}{2}\|u\|^2 - v^{c,\epsilon}(u)\Big).$$

Additionally, neither Assumption 2.1 nor Assumption 2.2 are necessary to its definition or to ensure its continuity. Using (5.2), one has the analytic formula

$$Q_\epsilon(u) = \int_{\mathcal{X}} x g_\epsilon(u,x) d\nu(x), \tag{5.3}$$

where

$$g_\epsilon(u,x) = \frac{\exp\Big(\frac{v_\epsilon(x) - \frac{1}{2}\|u-x\|^2}{\epsilon}\Big)}{\int_{\mathcal{X}} \exp\Big(\frac{v_\epsilon(z) - \frac{1}{2}\|u-z\|^2}{\epsilon}\Big) d\nu(z)} = \exp\Big(\frac{v_\epsilon^{c,\epsilon}(u) + v_\epsilon(x) - \frac{1}{2}\|u-x\|^2}{\epsilon}\Big).$$

This formula can be read as a conditional expectation of ν w.r.t. the conditional law associated with the transport plan $d\pi_\varepsilon(u,x) = g_\epsilon(u,x)d\nu(x)d\mu(u)$, which appears to be solution of the Kantorovich regularized problem, [30]. Plugging Q_ε into Definitions 2.4 and 2.5 induces the entropic analogs

$$S_\epsilon(u) = \frac{1}{1 - \|u\|} \int_{\|u\|}^1 Q_\varepsilon\Big(t\frac{u}{\|u\|}\Big) dt \tag{5.4}$$

and

$$E_\epsilon(u) = \frac{1}{\|u\|} \int_0^{\|u\|} Q_\varepsilon\Big(t\frac{u}{\|u\|}\Big) dt. \tag{5.5}$$

The empirical counterparts $\widehat{Q}_{\varepsilon,n}$, $\widehat{S}_{\varepsilon,n}$ and $\widehat{E}_{\varepsilon,n}$ of definitions (5.3) and (5.4) are obtained by plug-in estimators of the problem (5.1) for the empirical measure $\widehat{\nu}_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ and for the reference distribution μ. In the experiments of the next section, we use the stochastic Robbins-Monro algorithm taken from [7]. The entropic map $\widehat{Q}_{\varepsilon,n}$ defined in (5.3) is considered as a regularized empirical quantile function with $\varepsilon = 10^{-3}$, and the integrals in (5.4) are estimated by Riemann sums. Because the entropic map is continuous and it is the gradient of a convex function, even empirically, it shall provide nested and smooth contours. Moreover, estimating Monge maps is known to suffer from the curse of dimensionality, see [41, 49]. But, when the objective is the entropic map, for any $\varepsilon > 0$, convergence rates independent from d were obtained in [51]. Furthermore, the empirical version of the entropic map approaches the true Monge map ($\varepsilon = 0$) with a rate independent from d if ν is assumed to be discrete, see [48]. Note that, even for low ε, the entropic map does not exactly push-forward the reference distribution onto ν, and is rather an approximation of it.

(a) Expected shortfalls, $\varepsilon =$ 0.001

(b) Expected shortfalls, $\varepsilon =$ 0.005

(c) Expected shortfalls, $\varepsilon =$ 0.01

(d) Superquantiles, $\varepsilon = 0.001$ (e) Superquantiles, $\varepsilon = 0.005$ (f) Superquantiles, $\varepsilon = 0.01$
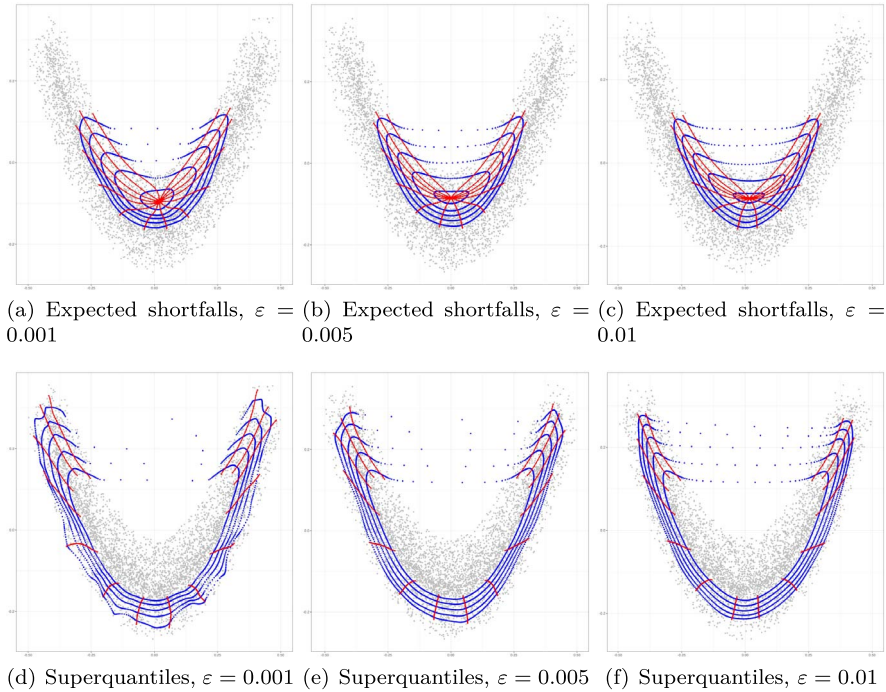
Fig 5.   *Center-outward expected shortfall and superquantile contours of levels α in $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ in blue and averaged sign curves C_u in red.*

5.2. Empirical study on simulated data

5.2.1. Descriptive plots for quantiles and superquantiles

Descriptive plots associated with Definition 2.6 are given in Figure 5, to describe all the information contained in our new concepts. The continuous reference distribution U_d is transported to the discrete banana-shaped measure ν with support of size $n = 5000$, via our empirical center-outward expected shortfall (resp. superquantile) function. Note how the data splits into a central area and a periphery area by the range of points covered by both maps. These are satisfying estimators for the well-suited $\varepsilon = 10^{-3}$, that is to say the first column of Figure 5. As the regularization parameter ε for E_ε and S_ε grows, the contours concentrate around the mean vector of ν, which is a known feature of entropic optimal transport. One can observe that our regularized approach yields smooth interpolation between image points. For visualization purposes, the red points of averaged sign curves are linked by straight paths. The blue points are not, to illustrate how the contours capture the empty space in the middle of the non convex point cloud.

5.2.2. Risk measurements on toy examples

In dimension $d = 2$, an empirical distribution can be represented with a scatter plot and the riskiest observations are visible with the naked eye: they are located furthest from the origin. From this principle, we selected easy-to-handle situations to evaluate our risk measures in Figure 6. Each row refers to a situation where a blue scatter plot (first column) is to be compared with an orange one (second column). For each situation, our *(Conditional) Vectors-at-Risk* of order $\alpha = 0.75$ is illustrated on each scatter plot. Moreover, the associated ρ_α^Q and ρ_α^S are computed and compared to the maximal correlation risk mesure ρ_C from [5] in the third column. The higher the bar, the riskier the corresponding vector of losses. In view of their comparison, the measurements $\rho \in \{\rho_\alpha^Q, \rho_\alpha^S, \rho_C\}$ are rescaled. For Y_1 the distribution of the blue scatter plot and Y_2 the orange one, one considers, for the height of the bars,

$$\rho(Y_1)/\max\left(\rho(Y_1), \rho(Y_2)\right) \quad \text{and} \quad \rho(Y_2)/\max\left(\rho(Y_1), \rho(Y_2)\right).$$

For each situation, our VaR and CVaR provide typical observations in the multivariate tails. The *maximal-correlation risk measure* ρ_C performs as well as expected, while our CVaR succeeds in more situations. Indeed, ρ_C benefits from several theoretical properties but only measures the riskiness of $X - \mathbb{E}(X)$, hence it neglects the shift effects.

Figure 6(a) contains two Gaussian distributions with identical mean vectors and covariance matrices related through the multiplication by a positive real. This is well tackled by each real-valued risk measurement. The existence of more outliers must be taken into account similarly, as in figure 6(b), where the two scatter plots originate from the same underlying distribution, but some outliers are added to the orange one. The CVaR is much more sensitive to these outliers than the VaR, as in dimension $d = 1$. Note that ρ_C ignores the outliers and leads to the wrong decision in the sense than the blue distribution is considered as the riskiest one. In the situation of Figure 6(c), the orange scatterplot is identical to the blue one, but shifted to the right side, so that it must be the riskiest. As ρ_C ignores this shift, it induces the wrong decision. Finally, in the last example of Figure 6(d), making a decision on the relative risk between the underlying distributions requires a preference for one of the components. Here, the two situations reveal risks of same intensity but directed towards different directions. To inform on the underlying directional information, vector-valued risk measures such as our VaRs and CVaRs are needed in complement.

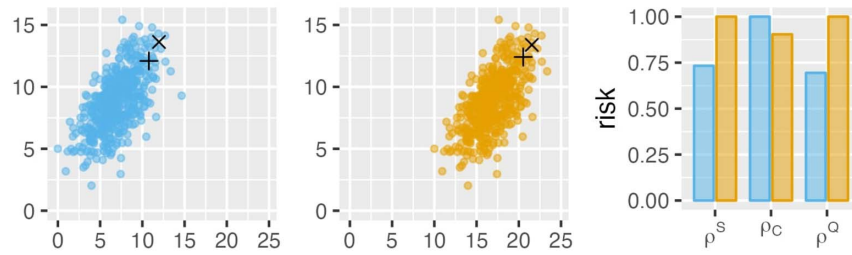### 5.3. Risk measurements on wind gusts data

In this section, we illustrate our new multivariate risk measures on the analysis of a real dataset provided by the ExtremalDep R package, [6], and dedicated to the study of strong wind gusts. This dataset has previously been studied in [25, 32, 42] in a context of risk measurement. The three variables are hourly wind
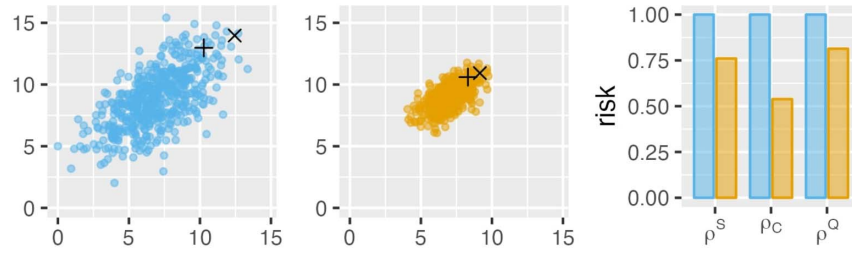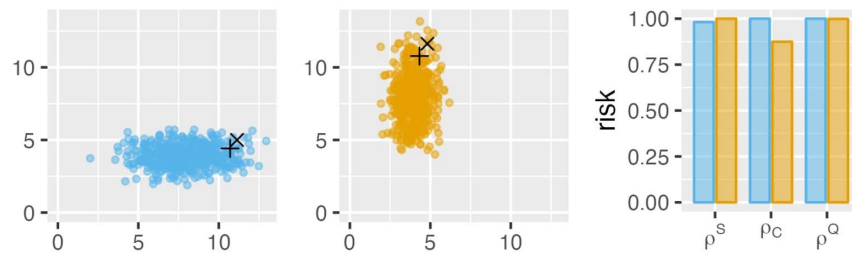
(a) Same Gaussians, reduced covariance matrix.
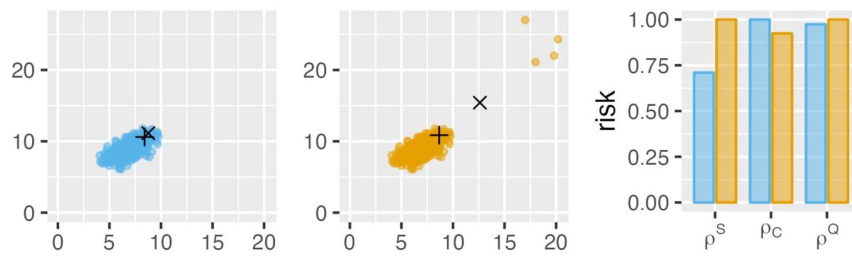
(b) Same Gaussians, added outliers.

(c) Same Gaussians with shift.

(d) Gaussians pointed in the vertical or horizontal direction

FIG 6. *VaRs (+) and CVaRs (×) on toy examples. Third column : real-valued risk measurements for the point clouds of the same line.*
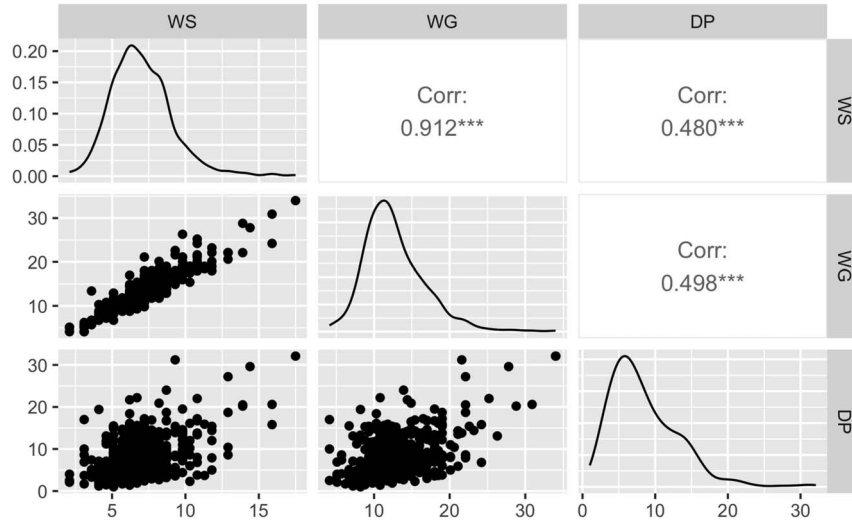
Fig 7. *Three dimensional wind gust data set.*

gust (WG) in meters per second, wind speed (WS) in meters per second, and air pressure at sea level (DP) in millibars, recorded at Parcay-Meslay (France) between July 2004 and July 2013. We consider the 1450 weekly maximum of each measurement. Because the variables are of different nature, it is the precise framework where multivariate risk analysis is useful, rather than the aggregation of several variables.

Figure 7 represents our three-dimensional dataset with pair scatterplots under the diagonal and Pearson correlation values above. The diagonal represents empirical density functions of each variable. Upper-right dependence can be observed and has physical explanations. Strong wind gusts occur with stormy weather, during which strong wind speed and high air pressure are frequently recorded.

Figure 8 represents the three dimensional empirical distribution in red together with our vectorial risk measures. With the increase of the dimension, such representative plots are no longer convenient. Rather, these measurements can be retrieved as in Table 1.

TABLE 1
Components of (Conditional) Vectors-at-Risk for several risk levels.

	$VaR_{0.25}$	$VaR_{0.5}$	$VaR_{0.75}$			$CVaR_{0.25}$	$CVaR_{0.5}$	$CVaR_{0.75}$
WS	8.21	9.98	12.36		WS	11.40	12.37	14.26
WG	15.06	18.45	21.58		WG	21.28	22.43	26.24
DP	11.92	13.65	17.84		DP	16.68	19.86	23.15

This summarizes the targeted information contained in the dataset. For in-

(a) Risk level 0.25   (b) Risk level 0.5   (c) Risk level 0.75
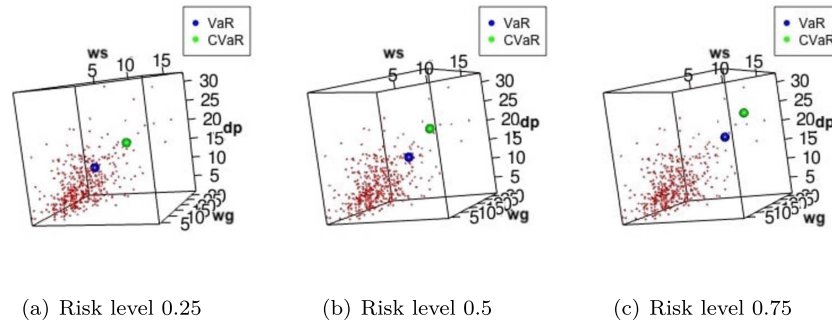
Fig 8. *Vectors-at-Risk and Conditional-Vectors-at-Risk.*

stance, with a given probability 0.25, 0.5 or 0.75, one shall expect, at worst, wind gusts of respective speed 15.06, 18.45, 21.58 m/s. With respectively same probability, averaged observations beyond these worst cases shall lie around 21.28, 22.43, 26.24 m/s. As in [42], we interpret these results thanks to the Beaufort scale. Note that this scale does not capture the speed of wind gusts, as it usually averages over 10 minutes, by convention. Values between 13.9 and 17.1 m/s can be considered as high winds. Strong winds begin with 17.2 m/s, with severely strong winds above 20.7 m/s up to 24.4 m/s. Severely strong winds can cause slight structural damage, but less than storms for values around 24.5-28.4 m/s. Above and up to 32.6 m/s, violent storms are very rarely experienced and cause widespread damage. Wind speeds greater than 32.7 m/s correspond to hurricanes. Thanks to Table 1, with probability 0.75, the worst scenarios in Parcay-Meslay for wind gusts are severely strong winds. Moreover, along the tail events corresponding to 25% of occurrences, one shall expect wind gusts of same speed as storms. Also, to illustrate the fact that considering univariate measures leads to underestimating the risk, we display in Figure 2 traditional univariate Values-at-Risk of WG, the wind gusts. For example, its median is 11.80, (strong breeze), and must be compared with 18.45, (strong winds), the second coordinate of $VaR_{0.5}$.

TABLE 2
Univariate quantiles of our variable WG.

Min	1st quartile	Median	3rd quartile	Max
4.10	9.80	11.80	14.90	34

Such differences have statistical foundings. The median depends on the *univariate* empirical distribution of WG to describe a probability of 1/2. Conversely, our multivariate $VaR_{0.5}$ encodes the *multivariate* joint probability of the whole point cloud (WG,WS,DP). This can be summarized by the fact that univariate risk measures *neglect the correlations*, which legitimates the use of multivariate risk analysis.

6. Conclusion and perspectives

In this paper, we provided new concepts of superquantiles and expected short-falls, in a meaningful multivariate way, based on the Monge-Kantorovich quantiles. We expect that our definitions allow the extension of the univariate applications to the multivariate case. Also, we introduced new transport-based risk measures, extending the concepts of Value-at-Risk and Conditional-Value-at-Risk to the dimension $d > 1$. The definitions of this paper are mainly motivated by practical concerns and the coherence of the proposed concepts with respect to their interpretation in dimension $d = 1$. Interestingly, from our center-outward expected shortfall function, one can retrieve Lorenz functions defined in [38]. This is discussed in Section 6.1, and other perspectives are also presented in Section 6.2.

6.1. On the class of integrated quantile functions

The role of integrated quantile functions in dimension $d = 1$ has been highlighted in [35]. Obviously, we belong to the line of works trying to extend such functions to the setting $d > 1$. In Section 4 of [38], two different approaches are discussed, in view of generalizing

$$\alpha \mapsto \int_0^\alpha Q(t)dt. \tag{6.1}$$

Hereafter, we bridge their concepts with ours, namely Definition 2.5 that extends (6.1). Recall that $\mathbb{S}^{d-1} = \{\varphi \in \mathbb{R}^d : \|\varphi\|_2 = 1\}$ and $\mathbb{P}_S$ is the uniform probability measure on $\mathbb{S}^{d-1}$. There, the (absolute) center-outward Lorenz function from [38][Definition 4] writes as

$$L_{\mathbf{X}\pm} : \alpha \mapsto \mathbb{E}[X\mathbb{1}_{X \in \mathbb{C}_\alpha}] = \int_{\mathbb{S}^{d-1}} \alpha E(\alpha\varphi)d\mathbb{P}_S(\varphi). \tag{6.2}$$

This being said, we believe that our proposed concepts of integration along sign curves contain more information, namely directional. When characterizing the contributions of central regions to the expectation, $L_{\mathbf{X}\pm}$ provides meaningful concepts of Lorenz curves, but this approach is insufficient for superquantiles and multivariate tails, as illustrated in Example 2.1.

Another important generalization of (6.1) is the one from [27] about multivariate Lorenz curves. A main difference between the concepts of [27] and [38] is the reference distribution, either the uniform on the unit hypercube or the spherical uniform. Here, Lorenz curves aim to visualize inequalities within a given population. In this context, one could focus either on the contribution of middle classes as in [38], or on the one of the bottom of the population as in [27], in a way that these works are in fact complementary. Furthermore, another approach is proposed in [38][Definition 6], that is real-valued, namely the (absolute) center-outward Lorenz potential function, for $\varphi \in \mathbb{S}^{d-1}$,

$$\alpha \mapsto \mathbb{E}_\varphi[\psi(\alpha\varphi)]. \tag{6.3}$$

This is quite natural because, at the core of the univariate considerations of [35], one can retrieve the property that the quantile function is the gradient of a convex potential, that is the definition of the center-outward quantile function Q. Also, it has a physical interpretation, with a measurement of the *work* of the quantile function. There, one might note that our center-outward expected shortfall function draws a connection between [38][Definition 4] and [38][Definition 6], through (6.2) and (2.11). Somehow, this strengthens the relation between $L_{\mathbf{X}\pm}$ and the potential function ψ. Studying deeper such connections between existing notions might lead to interesting results, but is left for further work.

6.2. Other perspectives

Perspectives are both theoretical and practical. Statistical tests associated with our real-valued measures would be of great interest, to automatically detect when to use the vector-valued ones. Also, these risk measures appeal for being used in practical settings. Extending our definitions to extreme quantile levels and building related procedures is another line of study. For instance, [18][Theorem 4.3] ensures the existence of a MK quantile map which is stable as moving further to the tail contours, which is not the case with the spherical uniform U_d. Our proposed center-outward superquantiles can be defined with other spherical reference measures, thus one can also calibrate these superquantiles as the work of [18] suggests. In addition, the many applications of the univariate superquantile function in other fields than risk measurement, such as superquantile regression or optimization with a superquantile loss, appeal for further work, even more so with Theorems 2.3 and 2.4. Finally, the convergence of empirical entropic optimal transport towards its population counterpart is an active field, [21, 34, 46, 49], and it remains to study to what extent existing results adapt to the convergence of $\widehat{S}_{\varepsilon,n}$ and $\widehat{E}_{\varepsilon,n}$.

Funding

References

[1] ACERBI, C., NORDIO, C. and SIRTORI, C. (2001). Expected Shortfall as a Tool for Financial Risk Management. arXiv:cond-mat/0102304.

[2] ACERBI, C. and TASCHE, D. (2002). On the coherence of Expected Shortfall. *Journal of Banking & Finance* **26** 1487–1503.

[3] ARMAUT, S., DIEL, R. and LALOË, T. (2023). On some depth-based risk measurement for high losses. hal-03933578.

[4] BARTHE, F., GUÉDON, O., MENDELSON, S. and NAOR, A. (2005). A probabilistic approach to the geometry of the ℓ_p^n-ball. *The Annals of Prob-*

ability **33** 480–513. https://doi.org/10.1214/009117904000000874 MR2123199

[5] BEIRLANT, J., BUITENDAG, S., DEL BARRIO, E., HALLIN, M. and KAMPER, F. (2020). Center-outward quantiles and the measurement of multivariate risk. *Insurance: Mathematics and Economics* **95** 79–100. https://doi.org/10.1016/j.insmatheco.2020 MR4154446

[6] BERANGER, B., PADOAN, S. and MARCON, G. (2021). ExtremalDep: extremal dependence models. R Package.

[7] BERCU, B. and BIGOT, J. (2021). Asymptotic distribution and convergence rates of stochastic algorithms for entropic optimal transportation between probability measures. *The Annals of Statistics* **49** 968–987. https://doi.org/10.1214/20-AOS1987 MR4255115

[8] BERCU, B., BIGOT, J. and THURIN, G. (2023). Stochastic optimal transport in Banach Spaces for regularized estimation of multivariate quantiles. arXiv:2302.00982.

[9] BOGACHEV, V. I. and RUAS, M. A. S. (2007). *Measure theory* **1**. Springer.

[10] BRENIER, Y. (1991). Polar Factorization and Monotone Rearrangement of Vector-Valued Functions. *Communications on Pure and Applied Mathematics* **44** 375–417. MR1100809

[11] CAI, J., JIA, H. and MAO, T. (2022). A multivariate CVaR risk measure from the perspective of portfolio risk management. *Scandinavian Actuarial Journal* **2022** 189–215. https://doi.org/10.1080/03461238.2021.1944905 MR4402541

[12] CARLIER, G., CHERNOZHUKOV, V., DE BIE, G. and GALICHON, A. (2022). Vector quantile regression and optimal transport, from theory to numerics. *Empirical Economics* **62** 35–62. https://doi.org/10.1007/s00181-020-01919-y

[13] CHERNOZHUKOV, V., GALICHON, A., HALLIN, M. and HENRY, M. (2017). Monge–Kantorovich depth, quantiles, ranks and signs. *The Annals of Statistics* **45** 223–256. https://doi.org/10.1214/16-AOS1450 MR3611491

[14] COUSIN, A. and DI BERNARDINO, E. (2013). On multivariate extensions of Value-at-Risk. *Journal of Multivariate Analysis* **119** 32–46. https://doi.org/10.1016/j.jmva.2013.03.016 MR3061413

[15] COUSIN, A. and DI BERNARDINO, E. (2014). On Multivariate Extensions of Conditional-Tail-Expectation. *Insurance: Mathematics and Economics* **55** 272–282. MR3179817

[16] CUESTA, J. A. and MATRÁN, C. (1989). Notes on the Wasserstein Metric in Hilbert Spaces. *Annals of Probability* **17** 1264–1276. MR1009457

[17] CUTURI, M. (2013). Sinkhorn Distances: Lightspeed Computation of Optimal Transport. *Advances in Neural Information Processing Systems* **26**.

[18] DE VALK, C. F. and SEGERS, J. (2018). Stability and tail limits of transport-based quantile contours LIDAM Discussion Papers ISBA No. 2018031, Université catholique de Louvain, Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA).

[19] DEB, N. and SEN, B. (2023). Multivariate rank-based distribution-free

nonparametric testing using measure transportation. *Journal of the American Statistical Association* **118** 192–207. MR4571116

[20] Del Barrio, E., González-Sanz, A. and Hallin, M. (2020). A note on the regularity of optimal-transport-based center-outward distribution and quantile functions. *J. Multivar. Anal.* **180** 104671. MR4147635

[21] del Barrio, E., Gonzalez-Sanz, A., Loubes, J.-M. and Niles-Weed, J. (2023). An improved central limit theorem and fast convergence rates for entropic transportation costs. *SIAM Journal on Mathematics of Data Science* **3** 639–669. MR4616887

[22] del Barrio, E. and Loubes, J.-M. (2019). Central limit theorems for empirical transportation cost in general dimension. *Ann. Probab.* **47** 926–951. https://doi.org/10.1214/18-AOP1275 MR3916938

[23] del Barrio, E., Sanz, A. G. and Hallin, M. (2022). Nonparametric Multiple-Output Center-Outward Quantile Regression. arXiv:2204.11756.

[24] Di Bernardino, E., Laloë, T., Maume-Deschamps, V. and Prieur, C. (2013). Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. *ESAIM: Probability and Statistics* **17** 236–256. MR3021318

[25] Di Bernardino, E. and Prieur, C. (2018). Estimation of the multivariate conditional tail expectation for extreme risk levels: Illustration on environmental data sets. *Environmetrics* **29** e2510. MR3868561

[26] Ekeland, I., Galichon, A. and Henry, M. (2012). Comonotonic measures of multivariate risks. *Mathematical Finance* **22** 109–132. MR2881882

[27] Fan, Y., Henry, M., Pass, B. and Rivero, J. A. (2022). Lorenz map, inequality ordering and curves based on multidimensional rearrangements. arXiv:2203.09000.

[28] Figalli, A. (2018). On the continuity of center-outward distribution and quantile functions. *Nonlinear Analysis* **117** 413–421. MR3886582

[29] Galichon, A. and Henry, M. (2012). Dual theory of choice with multivariate risks. *Journal of Economic Theory* **147** 1501–1516. MR2942620

[30] Genevay, A., Cuturi, M., Peyré, G. and Bach, F. (2016). Stochastic Optimization for Large-scale Optimal Transport. *Advances in neural information processing systems* **29**.

[31] Ghosal, P. and Sen, B. (2022). Multivariate Ranks and Quantiles using Optimal Transport: Consistency, Rates, and Nonparametric Testing. *The Annals of Statistics* **50** 1012–1037. MR4404927

[32] Goegebeur, Y., Guillou, A. and Qin, J. (2023). Dependent conditional tail moments for extreme levels. hal-03948352v3. MR4712386

[33] Goldfeld, Z., Kato, K., Rioux, G. and Sadhu, R. (2022). Limit Theorems for Entropic Optimal Transport Maps and the Sinkhorn Divergence. arXiv:2207.08683. MR4718466

[34] Gonzalez-Sanz, A., Loubes, J.-M. and Niles-Weed, J. (2022). Weak limits of entropy regularized Optimal Transport: potentials, plans and divergences. arXiv:2207.07427.

[35] Gushchin, A. A. and Borzykh, D. A. (2017). Integrated quantile functions: properties and applications. *Modern Stochastics: Theory and Appli-

cations **4** 285–314. MR3739012

[36] HALLIN, M. (2022). Measure Transportation and Statistical Decision Theory. *Annual Review of Statistics and Its Application* **9** 401–424. MR4394914

[37] HALLIN, M., DEL BARRIO, E., CUESTA-ALBERTOS, J. and MATRÁN, C. (2021). Distribution and quantile functions, ranks and signs in dimension d: A measure transportation approach. *The Annals of Statistics* **49** 1139–1165. https://doi.org/10.1214/20-AOS1996 MR4255122

[38] HALLIN, M. and MORDANT, G. (2022). Center-Outward Multiple-Output Lorenz Curves and Gini Indices a measure transportation approach Working Papers ECARES, ULB – Universite Libre de Bruxelles.

[39] HALLIN, M., VECCHIA, D. and LIU, H. (2023). Rank-based testing for semiparametric VAR models: A measure transportation approach. *Bernoulli* **29** 229–273. https://doi.org/10.3150/21-BEJ1456 MR4497246

[40] 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** 497–546. MR2088289

[41] HÜTTER, J.-C. and RIGOLLET, P. (2021). Minimax estimation of smooth optimal transport maps. *The Annals of Statistics* **49** 1166–1194. https://doi.org/10.1214/20-AOS1997 MR4255123

[42] MARCON, G., NAVEAU, P. and PADOAN, S. (2017). A semi-parametric stochastic generator for bivariate extreme events. *Stat* **6** 184–201. MR3653052

[43] MASUD, S. B., WERENSKI, M., MURPHY, J. M. and AERON, S. (2023). Multivariate soft rank via entropic optimal transport: sample efficiency and generative modeling. *Journal of Machine Learning Research* **24** 1–65. MR4596107

[44] MCCANN, R. J. (1995). Existence and uniqueness of monotone measure-preserving maps. *Duke Mathematical Journal* **80** 309–323. https://doi.org/10.1215/S0012-7094-95-08013-2 MR1369395

[45] MCCANN, R. J. (2001). Polar factorization of maps on Riemannian manifolds. *Geometric & Functional Analysis GAFA* **11** 589–608. MR1844080

[46] MENA, G. and NILES-WEED, J. (2019). Statistical bounds for entropic optimal transport: sample complexity and the central limit theorem. In *Advances in Neural Information Processing Systems* **32**.

[47] NIU, Z. and BHATTACHARYA, B. B. (2022). Distribution-free joint independence testing and robust independent component analysis using optimal transport. arXiv:2211.15639.

[48] POOLADIAN, A.-A., DIVOL, V. and NILES-WEED, J. (2023). Minimax estimation of discontinuous optimal transport maps: The semi-discrete case. arXiv:2301.11302.

[49] POOLADIAN, A.-A. and NILES-WEED, J. (2021). Entropic estimation of optimal transport maps. arXiv:2109.12004.

[50] PRÉKOPA, A. (2012). Multivariate value at risk and related topics. *Annals of Operations Research* **193(1)** 49–69. MR2874756

[51] RIGOLLET, P. and STROMME, A. J. (2022). On the sample complexity of

entropic optimal transport. arXiv:2206.13472.

[52] ROCKAFELLAR, R. T. (1970). *Convex analysis.* Princeton University Press. MR0274683

[53] ROCKAFELLAR, R. T. (1981). Favorable classes of Lipschitz continuous functions in subgradient optimization. MR0704977

[54] ROCKAFELLAR, R. T. and ROYSET, J. O. (2013). Superquantiles and their applications to risk, random variables and regression. In *Theory Driven by Influential Applications* 151–167. MR3357644

[55] ROCKAFELLAR, R. T. and ROYSET, J. O. (2014). Random variables, monotone relations, and convex analysis. *Mathematical Programming* **148** 297–331. MR3274854

[56] SCHECHTMAN, G. and ZINN, J. (1990). On the volume of the intersection of two L_p^n balls. *Proceedings of the American Mathematical Society* **110** 217–224. MR1015684

[57] SHI, H., DRTON, M., HALLIN, M. and HAN, F. (2021). Distribution-free tests of multivariate independence based on center-outward quadrant, Spearman, Kendall, and van der Waerden statistics. arXiv:2111.15567v4.

[58] SHI, H., DRTON, M. and HAN, F. (2022). Distribution-free consistent independence tests via center-outward ranks and signs. *Journal of the American Statistical Association* **117** 395–410. MR4399094

[59] TORRES, R., LILLO, R. E. and LANIADO, H. (2015). A directional multivariate value at risk. *Insurance: Mathematics and Economics* **65** 111–123. MR3430401

[60] UĞURLU, K. (2023). A new coherent multivariate average-value-at-risk. *Optimization* **72** 493–519. MR4548360

[61] VILLANI, C. (2003). *Topics in optimal transportation. Graduate Studies in Mathematics* **58**. American Mathematical Society. MR1964483