Multivariate ρ-Quantiles: a Spatial Approach*

DIMITRI KONEN[†] and DAVY PAINDAVEINE[‡]

[†] *Université libre de Bruxelles, ECARES and Département de Mathématique, Boulevard du Triomphe, Campus Plaine, CP210, B-1050, Brussels, Belgium*
Tel.: +3226505897

[‡] *Université libre de Bruxelles, ECARES and Département de Mathématique, Avenue F.D. Roosevelt, 50, ECARES, CP114/04, B-1050, Brussels, Belgium*
Tel.: +3226503845

By substituting an L_p loss function for the L_1 loss function in the optimization problem defining quantiles, one obtains L_p-quantiles that, as shown recently, dominate their classical L_1-counterparts in financial risk assessment. In this work, we introduce a concept of multivariate L_p-quantiles generalizing the spatial (L_1-)quantiles from Chaudhuri (1996). Rather than restricting to power loss functions, we actually allow for a large class of convex loss functions ρ. We carefully study existence and uniqueness of the resulting ρ-quantiles, both for a general probability measure over $\mathbb{R}^d$ and for a spherically symmetric one. Interestingly, the results crucially depend on ρ and on the nature of the underlying probability measure. Building on an investigation of the differentiability properties of the objective function defining ρ-quantiles, we introduce a companion concept of spatial ρ-depth, that generalizes the spatial depth from Vardi and Zhang (2000). We study extreme ρ-quantiles and show in particular that extreme L_p-quantiles behave in fundamentally different ways for $p \leq 2$ and $p > 2$. Finally, we establish Bahadur representation results for sample ρ-quantiles and derive their asymptotic distributions. Throughout, we impose only very mild assumptions on the underlying probability measure, and in particular we never assume absolute continuity with respect to the Lebesgue measure.

MSC 2010 subject classifications: Primary 62G99, 62H99.
Keywords: Bahadur representation results, Convex objective functions, M-estimation, Multivariate quantiles, Spatial depth, Spatial quantiles.

1. Introduction

The concept of quantile, which is of paramount importance in statistics, has long been limited to probability measures over $\mathbb{R}$. Defining a suitable quantile concept in $\mathbb{R}^d$ is a problem that is intrinsically difficult due to the lack of a canonical ordering in $\mathbb{R}^d$, $d > 1$.

*Research is supported by the Program of Concerted Research Actions (ARC) of the Université libre de Bruxelles and by an Aspirant fellowship from the FNRS (Fonds National pour la Recherche Scientifique), Communauté Française de Belgique.

This has been an active research topic in the last decades; see, among many others, Hallin et al. (2021), Hallin, Paindaveine and Šiman (2010), Serfling (2002), and the references therein. One of the most successful multivariate quantile concepts is the concept of *spatial* (or *geometric*) quantiles from Chaudhuri (1996); for a probability measure P over $\mathbb{R}^d$, the spatial quantile of order α in direction u is defined as the minimizer over $\mathbb{R}^d$ of the map

$$\mu \mapsto M_{\alpha,u}(\mu) = \int_{\mathbb{R}^d} \left\{ \|z-\mu\| - \|z\| - \alpha u'\mu \right\} dP(z), \tag{1}$$

with $\alpha \in [0,1)$ and $u \in \mathcal{S}^{d-1} := \{z \in \mathbb{R}^d : \|z\|^2 := z'z = 1\}$. The success of spatial quantiles is explained by several key distinctive properties, among which: spatial quantiles are easy to compute, even for large d (Mukhopadhyaya and Chatterjee (2011), Vardi and Zhang (2000)). The asymptotic behavior of their sample version is rather standard (Chaudhuri (1996), Zhou and Serfling (2008)). They can easily be extended into regression quantiles (Chakraborty (2003), Cheng and De Gooijer (2007)) or turned into quantiles for functional data (Chakraborty and Chaudhuri (2014), Chowdhury and Chaudhuri (2019), Serfling and Wijesuriya (2017)).

For $d = 1$, spatial quantiles, that minimize the L_1-objective function in (1), reduce to the usual univariate quantiles. In particular, the collection of intervals whose endpoints are the spatial quantiles of order α in direction $u = -1$ and $u = 1$ is a nested family of interquantile intervals, that all contain the univariate median (which is obtained with $\alpha = 0$, irrespective of the direction u). Similarly, *expectiles*, an L_2-analog of quantiles introduced in Newey and Powell (1987), provide a nested family of centrality intervals that all contain the mean of the distribution. Expectiles have met a big success, particularly so in financial risk assessment, where they provide coherent risk measures; see, e.g., Daouia, Girard and Stupfler (2018), Kuan, Yeh and Hsu (2009), Taylor (2008). Quantiles and expectiles belong to the class of L_p-quantiles (associated with L_p loss functions, with $p \geq 1$), or, more generally, of M-quantiles (associated with general convex loss functions); see Breckling and Chambers (1988), Chen (1996). Recently, there has been a growing interest in such generalized quantiles, still with risk assessment as one of the main applications; see, e.g., Daouia, Girard and Stupfler (2019), Gardes, Girard and Stupfler (2020), or Usseglio-Carleve (2018).

The success of spatial quantiles in $\mathbb{R}^d$ and the growing interest in L_p-quantiles in the univariate case $d = 1$ suggest to define a spatial concept of L_p-quantiles. For $p = 2$, this has actually recently been done in Herrmann, Hofert and Mailhot (2018), but the resulting spatial expectiles remain less well understood than their spatial L_1-counterparts from Chaudhuri (1996). To the best of our knowledge, a spatial L_p-quantile concept, or, more generally, a spatial M-quantile concept has not been considered in the literature. In this work, we define such a general concept and we thoroughly investigate its properties. We adopt the following definition.

Definition 1. *Let $\rho : \mathbb{R}^+ \to \mathbb{R}^+$ be a convex function and P be a probability measure over $\mathbb{R}^d$. Fix $\alpha \in [0,1)$ and $u \in \mathcal{S}^{d-1}$. We say that $\mu^\rho_{\alpha,u} = \mu^\rho_{\alpha,u}(P)$ is a* spatial ρ-quantile of order α in direction u for P *if and only if it minimizes the objective function*

$$\mu \mapsto M^\rho_{\alpha,u}(\mu) := \int_{\mathbb{R}^d} \left\{ H^\rho_{\alpha,u}(z-\mu) - H^\rho_{\alpha,u}(z) \right\} dP(z) \tag{2}$$

over $\mathbb{R}^d$, where we let

$$H^\rho_{\alpha,u}(z) := \rho(\|z\|)\left(1 + \alpha \frac{u'z}{\|z\|}\right)\xi_{z,0}, \tag{3}$$

with $\xi_{z_1,z_2} := \mathbb{I}[z_1 \neq z_2]$ (throughout, $\mathbb{I}[A]$ is the indicator function of A).

It might have been natural to write μ^ρ_v, with $v = \alpha u$ rather than $\mu^\rho_{\alpha,u}$, to emphasize the indexation of spatial ρ-quantiles on the unit ball, but we favour the notation $\mu^\rho_{\alpha,u}$ that stresses the heterogenous roles α and u will play in the sequel. The multivariate L_p-quantiles we consider in this work are obtained with $\rho(t) = t^p$, $p \geq 1$. Clearly, these reduce for $p = 1$ to the minimizers of (1), that is, to the spatial quantiles from Chaudhuri (1996). If P has finite second-order moments, then our L_p-quantiles reduce for $p = 2$ to the expectiles introduced in Herrmann, Hofert and Mailhot (2018); as we will show, however, our formulation above only requires that P has finite first-order moments. Note that for $\alpha = 0$, the ρ-quantile $\mu^\rho_{\alpha,u}$, irrespective of u (the direction u does not play any role for $\alpha = 0$), is an M-functional of location, that, for $p = 1$ and $p = 2$, provides the celebrated spatial median (Brown (1983)) and the mean vector of P, respectively. The ρ-quantiles from Definition 1 extend this M-functional of location in the same way the spatial quantiles from Chaudhuri (1996) extend the spatial median: they are thus *M-quantiles*, in the sense of Breckling and Chambers (1988) or Koltchinski (1997) (it should be noted that the only intersection between the M-quantiles in Definition 1 and those from Koltchinski (1997) are the spatial quantiles from Chaudhuri (1996)).

Above, the motivation to consider L_p-quantiles was linked to their relevance for risk assessment, and there is no doubt that spatial L_p-quantiles are natural tools to define suitable risk measures in situations where multidimensional portfolios are considered. The main focus in this work, however, is on a careful study of the probabilistic properties of these L_p-quantiles and, more generally, of the corresponding ρ-quantiles. Quite remarkably, many of these properties crucially depend on the loss function ρ. We provide two examples. (i) Convexity of the objective function in (2) for any order α and direction u is a key property for the study of ρ-quantiles and for their evaluation at empirical distributions, and it may be expected that this convexity is inherited from the convexity of ρ. Our results, however, will show that, in the class of L_p-quantiles, this is the case if and only if $p \leq 2$. For $p > 2$, we will show that convexity holds for $\alpha \leq \alpha_p$ only, where, quite remarkably, α_p is very close to one for any p but does not depend on p monotonically. (ii) The spatial quantiles from Chaudhuri (1996) have recently been criticised because they exit any compact set as $\alpha \to 1$ even for a compactly supported probability measure P; see Girard and Stupfler (2017). As our results will show, this behavior of extreme spatial quantiles is shared by L_p-quantiles with $p \leq 2$, but not by those with $p > 2$. While we discussed these results here for L_p-quantiles only, we will throughout study properties of ρ-quantiles for a virtually arbitrary convex loss function ρ, which will allow us to consider, e.g., exponential loss functions or the celebrated Huber loss functions.

The outline of the paper is as follows. In Section 2, we provide the assumptions under which the objective function $M^{\rho}_{\alpha,u}(\mu)$ in (2) is well-defined for any μ, and we discuss existence of ρ-quantiles. In Section 3, we obtain a necessary and sufficient condition for convexity of $M^{\rho}_{\alpha,u}(\mu)$, we characterize the orders α for which convexity fails when this condition is not satisfied, and we exploit this to derive uniqueness results for ρ-quantiles. In Section 4, we refine these convexity and uniqueness results in the particular case for which the underlying probability measure is spherically symmetric. In Section 5, we study first- and second-order differentiability of the objective function $M^{\rho}_{\alpha,u}(\mu)$, which will play a key role in the subsequent sections. In Section 6, we exploit Robert Serfling's DOQR paradigm to define ρ-depth functions, ρ-outlyingness functions and ρ-rank functions associated with our ρ-quantile functions. We also identify conditions under which ρ-quantile functions are homeomorphisms from the open unit ball (quantiles are indexed by $(\alpha, u) \in [0,1) \times \mathcal{S}^{d-1}$ or, equivalently, by αu in the open unit ball of $\mathbb{R}^d$) to the whole Euclidean space $\mathbb{R}^d$. This will play a major role when studying in Section 7 the behavior of extreme ρ-quantiles. In Section 8, we derive Bahadur representation results for sample ρ-quantiles and deduce their asymptotic distribution. Finally, we briefly discuss some perspectives for future research in Section 9. Proofs are collected in several technical appendices (Appendices S.1 to S.8).

2. Existence

Throughout, we assume that the loss function ρ belongs to the class $\mathcal{C}$ of functions from $\mathbb{R}^+$ to $\mathbb{R}^+$ that are convex, piecewise twice continuously differentiable on $(0,\infty)$, and satisfy $\rho(t) = 0$ only for $t = 0$. Here, ρ *is piecewise twice continuously differentiable on* $(0,\infty)$ means that either (i) there exist a nonnegative integer K and $(0 =: t_0 <) t_1 < t_2 < \ldots < t_K < t_{K+1} := \infty$ such that ρ is twice continuously differentiable on each open interval (t_k, t_{k+1}), $k = 0, \ldots, K$, or (ii) there exists a monotone strictly increasing sequence $(t_0 := 0, t_1, t_2, \ldots)$ in $\mathbb{R}^+$ diverging to infinity such that ρ is twice continuously differentiable on each open interval (t_k, t_{k+1}), $k \in \mathbb{N}$. We let $\mathcal{D}_\rho = (0,\infty) \setminus \{t_1, t_2, \ldots, t_K\}$ in case (i) and $\mathcal{D}_\rho = (0,\infty) \setminus \{t_1, t_2, \ldots\}$ in case (ii). Examples of loss functions in $\mathcal{C}$ are the power loss functions $\rho(t) = t^p$, with $p \geq 1$, the exponential functions $\rho(t) = \exp(ct) - 1$, with $c > 0$, and the Huber loss functions $\rho(t) = (t^2/2)\mathbb{I}[0 < t < c] + c(t - (c/2))\mathbb{I}[t \geq c]$, with $c > 0$. For the Huber loss functions, $\mathcal{D}_\rho = (0, \infty) \setminus \{c\}$, whereas $\mathcal{D}_\rho = (0, \infty)$ for power and exponential loss functions.

For any $\rho \in \mathcal{C}$, we denote as $\mathcal{P}_d^\rho$ the class of probability measures P over $\mathbb{R}^d$ such that for any $\mu \in \mathbb{R}^d$, there exists $\delta > 0$ for which

$$\int_{\mathbb{R}^d} \psi_-(\|z-\mu\| + \delta)\, dP(z) < \infty; \tag{4}$$

throughout, ψ_- and ψ_+ will denote the left- and right-derivative of ρ, respectively (convexity of ρ ensures existence of these one-sided derivatives). For the power loss function $\rho(t) = t^p$, with $p \geq 1$, $P \in \mathcal{P}_d^\rho$ if and only if P has finite moments of order $p-1$ (that is, $\mathrm{E}[\|Z\|^{p-1}] < \infty$, where Z is a random d-vector with distribution P). For $\rho(t) = t$, $\mathcal{P}_d^\rho$ thus collects all probability measures on $\mathbb{R}^d$, which is also the case for Huber loss functions.

We then have the following existence result (see Appendix S.2 for a proof).

Theorem 2.1. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}_d^\rho$. Fix $\alpha \in [0,1]$ and $u \in \mathcal{S}^{d-1}$. Then, (i) $M_{\alpha,u}^\rho(\mu)$ is well-defined for any $\mu \in \mathbb{R}^d$; (ii) if $\alpha < 1$, then P admits at least one ρ-quantile of order α in direction u.*

The existence result in Theorem 2.1(ii) is obtained by establishing that, for any $\alpha \in [0,1)$ and $u \in \mathcal{S}^{d-1}$, the map $\mu \mapsto M_{\alpha,u}^\rho(\mu)$ is both coercive (in the sense that $M_{\alpha,u}^\rho(\mu)$ diverges to infinity as $\|\mu\|$ does) and continuous on $\mathbb{R}^d$. We require that $\rho \in \mathcal{C}$ in Theorem 2.1 to avoid introducing many different collections of loss functions in the sequel, but inspection of the proof reveals that the result actually holds without any differentiability assumption on ρ.

Theorem 2.1(ii) shows that, for $\rho(t) = t^p$, with $p \geq 1$, ρ-quantiles exist for any $\alpha \in [0,1)$ and $u \in \mathcal{S}^{d-1}$ as soon as P has finite moments of order $p-1$ (while subtracting $H_{\alpha,u}^\rho(z)$ in the integrand of (2) in principle has no impact on the corresponding quantile minimizers, not doing so would guarantee existence of a minimizer only under the stronger condition of finite moments of order p). In particular, taking $p = 1$ and $p = 2$, this shows that the quantiles from Chaudhuri (1996) always exist, whereas their expectile analogs from Herrmann, Hofert and Mailhot (2018) only require that P has finite first-order moments (as already mentionned, finite second-order moments are imposed in Herrmann, Hofert and Mailhot (2018)). The quantiles associated with Huber loss functions also always exist for any $\alpha \in [0,1)$ and $u \in \mathcal{S}^{d-1}$.

In Section 7 below, we will study extreme ρ-quantiles, that is, the ρ-quantiles indexed by an order α that is arbitrarily close to one. As we will see, the behavior of such quantiles crucially depends on the existence of the boundary ρ-quantiles indexed by an order $\alpha = 1$ (the term "boundary" results from the fact that Definition 1 imposes that $\alpha \in [0,1)$). Our interest in such boundary quantiles explains why we will be investigating the properties of the map $M^{\rho}_{\alpha,u}(\mu)$ also for $\alpha = 1$, as we already did in Theorem 2.1(ii). At this stage, we stress that Theorem 2.1(ii) remains silent about the existence of such boundary ρ-quantiles, the reason being that coercivity of $\mu \mapsto M^{\rho}_{\alpha,u}(\mu)$ may fail for $\alpha = 1$. For $\rho(t) = t$ and $d \geq 2$, for instance, no such boundary quantiles exist when P is non-atomic and is not supported on a line of $\mathbb{R}^d$; see Girard and Stupfler (2017), Proposition 2.1.

We conclude this section with the following orthogonal- and translation-equivariance result, that will be particularly relevant in Section 4 when considering the particular case for which P is spherically symmetric (the proof readily follows from Definition 1).

Proposition 2.1. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}^d_\rho$. Fix $\alpha \in [0,1)$ and $u \in \mathcal{S}^{d-1}$. Let O be a $d \times d$ orthogonal matrix, b be a d-vector, and denote as $P_{O,b}$ the distribution of $OZ + b$ when Z has distribution P. Then, if μ is a ρ-quantile of P of order α in direction u, then $O\mu + b$ is a ρ-quantile of $P_{O,b}$ of order α in direction Ou.*

Spatial ρ-quantiles are not affine-equivariant, that is, they fail to be equivariant under general affine transformations. However, they can be made affine-equivariant through a transformation-retransformation approach; see, e.g., Serfling (2010) in the case of the Chaudhuri (1996) spatial quantiles.

3. Convexity and uniqueness

The quantiles studied in this work are defined as minimizers of the map $\mu \mapsto M^{\rho}_{\alpha,u}(\mu)$ in (2). Convexity of this map is a most desirable property, that is expected to play a key role when investigating uniqueness of these quantiles and when evaluating them for empirical probability measures. In this section, we therefore study under which conditions on the loss function $\rho \in \mathcal{C}$ the map $\mu \mapsto M^{\rho}_{\alpha,u}(\mu)$ is convex for any $P \in \mathcal{P}^{\rho}_d$.

First note that convexity of $H^{\rho}_{\alpha,u}$ trivially implies convexity of $M^{\rho}_{\alpha,u}$, and that, if $H^{\rho}_{\alpha,u}$ is not convex, then there exists $P \in \mathcal{P}^{\rho}_d$ for which $M^{\rho}_{\alpha,u}$ fails to be convex (simply consider a Dirac probability measure). Therefore, we may focus on studying convexity of $H^{\rho}_{\alpha,u}$. Since any $\rho \in \mathcal{C}$ clearly makes $H^{\rho}_{\alpha,u}$ convex for $d = 1$, we tacitly restrict throughout this section to the case $d \geq 2$. We start with the following preliminary result showing that the larger α, the fewer the functions ρ making $H^{\rho}_{\alpha,u}$ convex for any $u \in \mathcal{S}^{d-1}$.

Lemma 3.1. *For any $\alpha \in [0,1]$, denote as $\mathcal{C}_\alpha$ the collection of functions $\rho \in \mathcal{C}$ such that $H^{\rho}_{\alpha,u}$ is convex for any $u \in \mathcal{S}^{d-1}$. Then, we have the following: (i) $\mathcal{C}_0 = \mathcal{C}$; (ii) if $\alpha_1, \alpha_2 \in [0,1]$ satisfy $\alpha_1 < \alpha_2$, then $\mathcal{C}_{\alpha_2} \subseteq \mathcal{C}_{\alpha_1}$.*

This result suggests considering $\alpha_\rho := \max\{\alpha \in [0,1] : \rho \in \mathcal{C}_\alpha\}$, the largest value of α for which ρ makes $H^{\rho}_{\alpha,u}$ convex for any $u \in \mathcal{S}^{d-1}$ (it is trivial to prove that the maximum exists for any $\rho \in \mathcal{C}$). Ideally, we would like to have that $\alpha_\rho = 1$, as this would ensure that $H^\rho_{\alpha,u}$ is convex for any $\alpha \in [0,1]$ and $u \in \mathcal{S}^{d-1}$. The following result provides a necessary and sufficient condition for $\alpha_\rho = 1$.

Theorem 3.1. *Let $\rho \in \mathcal{C}$. Then, irrespective of $d \geq 2$, $\alpha_\rho = 1$ if and only if the map $t \mapsto t^2/\rho(t)$ is concave on $(0,\infty)$.*

As a corollary, the power loss function $\rho(t) = t^p$ makes $H^\rho_{\alpha,u}$ convex for any $\alpha \in [0,1)$ and any $u \in \mathcal{S}^{d-1}$ if and only if $p \in [1,2]$. For $p > 2$, it is then of interest to determine the corresponding value of $\alpha_\rho (< 1)$. More generally, the following result allows one to determine α_ρ for any loss function ρ that does not satisfy the necessary and sufficient condition in Theorem 3.1.

Theorem 3.2. *Let $\rho \in \mathcal{C}$ be such that the map $t \mapsto t^2/\rho(t)$ is not concave on $(0,\infty)$. Then, irrespective of $d \geq 2$,*

$$\alpha_\rho = \inf_{t \in \mathcal{D}^{\mathrm{cv}}_\rho} \sqrt{\frac{q_t(4p_t^2 - 4p_t - q_t)}{4(p_t - 1)^2(q_t + 1)}} < 1,$$

with

$$p_t := \frac{t\psi_-(t)}{\rho(t)} \quad \textit{and} \quad q_t := \frac{t^2 \psi'_-(t)}{t\psi_-(t) - \rho(t)},$$

where we let $\mathcal{D}^{\mathrm{cv}}_\rho := \{t \in \mathcal{D}_\rho : (t^2/\rho(t))'' > 0\}$ and where ψ'_- is the left-derivative of ψ_- (in this result, $\psi_-(t)$ and $\psi'_-(t)$ are used only for $t \in \mathcal{D}_\rho$, so that we could write $\psi_-(t) = \rho'(t)$ and $\psi'_-(t) = \rho''(t)$ above).

For $\rho(t) = t^p$ with $p > 2$, one readily checks that $\mathcal{D}^{\mathrm{cv}}_\rho = (0,\infty)$ and

$$\alpha_\rho := \sqrt{\frac{p^2(4p-5)}{4(p-1)^2(p+1)}}. \tag{5}$$

Remarkably, α_ρ in (5) exhibits a non-monotonic pattern in p: for $p \in [2,5]$, it decreases monotonically from one to its minimal value $\sqrt{125/128}$ (slightly above .9882), then increases monotonically to one again for $p \in [5,\infty)$; see the left panel of Figure 1. For $p > 2$, it is thus only for most extreme quantile orders α that convexity fails. This is even more the case for the exponential loss function $\rho(t) = \exp(ct) - 1$, for which $\mathcal{D}^{\mathrm{cv}}_\rho = (3.0861/c, \infty)$ and $\alpha_\rho = .9939$. It can be shown that, if the loss function ρ is such that $t \mapsto \rho(t)/t$ is convex, then $\alpha_\rho \geq \sqrt{2/3} \approx .8165$ (for the sake of completeness, we prove this in Appendix S.3; see Corollary S.3.1). Like the power loss functions $\rho(t) = t^p$, $p \in (1,2)$, the Huber loss functions provide a compromise between the L_1 and L_2 loss functions, but since Theorem 3.2 entails that $\alpha_\rho = 0$ for the Huber loss functions, power loss functions clearly should be favoured in terms of convexity.

An important corollary of convexity is the following uniqueness result.

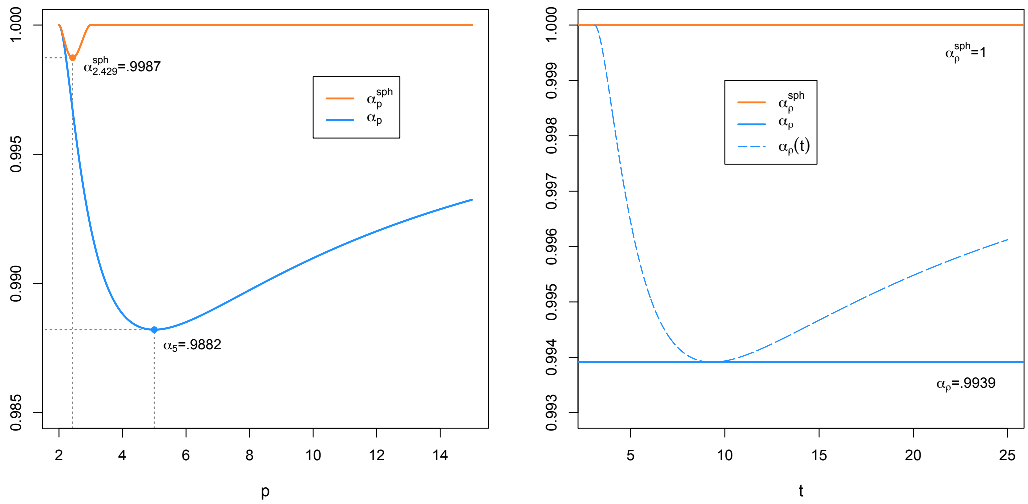

Figure 1. (Left:) For power loss functions $\rho(t) = t^p$, plot of α_ρ (see (5)) and $\alpha_\rho^{\mathrm{sph}}$ (see (6)) (note that (5) shows that $\alpha_\rho \to 1$ as $p \to \infty$). (Right:) For the loss function $\rho(t) = \exp(t) - 1$, plot of the quantity, $\alpha_\rho(t)$ say, of which the infimum is taken in Theorem 3.2; the blue line marks the resulting infimum $\alpha_\rho = .9939$, whereas the orange line stresses that $\alpha_\rho^{\mathrm{sph}} = 1$ (see Section 4).

Theorem 3.3. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}_\rho^d$. Assume that there is no open interval in $(0, \infty)$ on which ψ_- is constant or that P is not concentrated on a line. Then, for any $\alpha \in [0, \alpha_\rho) \cup \{0\}$ and $u \in \mathcal{S}^{d-1}$ (union with $\{0\}$ is needed when $\alpha_\rho = 0$), the map $\mu \mapsto M_{\alpha,u}^\rho(\mu)$ is strictly convex on $\mathbb{R}^d$, so that the ρ-quantile $\mu_{\alpha,u}^\rho$ is unique.*

This covers the well-known result stating that, for $\rho(t) = t$, all ρ-quantiles are unique provided that P is not supported on a line. Remarkably, Theorem 3.3 shows that this structural constraint on P is not needed for the ρ-quantiles associated with other power loss functions: under the corresponding moment assumptions, all ρ-quantiles are unique for $\rho(t) = t^p$, with $p \in (1, 2]$, whereas, for $p > 2$, all ρ-quantiles with an order α that is below (5) are unique. Similarly, for exponential loss functions, ρ-quantiles are unique for any order $\alpha < .9939$.

4. The spherical case

In this section, we consider the special case for which $P (\in \mathcal{P}_d^\rho)$ is spherically symmetric about some location $\mu_0 (\in \mathbb{R}^d)$, in the sense that, for any d-Borel set B and any $d \times d$ orthogonal matrix O, the P-probability of $\mu_0 + OB$ does not depend on O. Since ρ-quantiles are translation-equivariant, we will actually restrict, without any loss of generality, to the case $\mu_0 = 0$ (translation-equivariance here means that if μ is a ρ-quantile of P of order α in direction u, then, for any $h \in \mathbb{R}^d$, $\mu + h$ is a ρ-quantile of P_h of order α in direction u, where P_h is the distribution of $Z + h$ when Z has distribution P).

Note that it follows from Proposition 2.1, if P is spherically symmetric about the origin of $\mathbb{R}^d$ and satisfies $P[\{0\}] < 1$, with $d \geq 2$ say, then any quantile contour $\{\mu_{\alpha,u} : u \in \mathcal{S}^{d-1}\}$, with $\alpha \in [0, \alpha_\rho) \cup \{0\}$, is a hypersphere (uniqueness of these quantiles follows from Theorem 3.3 since P is then not supported on a line). Proposition 2.1 then also implies that, for an arbitrary order $\alpha \in [0, 1)$ and any direction $u \in \mathcal{S}^{d-1}$, the ρ-quantiles of P of order α in direction u form a set that is invariant under all rotations fixing u. In particular, if $\mu^\rho_{\alpha,u}$ is unique, then it belongs to the line spanned by u, which is most natural. For $\alpha \geq \alpha_\rho$, however, uniqueness is not guaranteed, so that it is unclear whether or not quantiles meet this natural property in the spherical case. This motivates the following result.

Theorem 4.1. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}^\rho_d$. Assume that P is spherically symmetric about the origin of $\mathbb{R}^d$. Then, (i) for $\alpha = 0$ and any $u \in \mathcal{S}^{d-1}$, the unique ρ-quantile $\mu^\rho_{\alpha,u}$ is the origin of $\mathbb{R}^d$; (ii) for $\alpha \in (0, 1)$ and $u \in \mathcal{S}^{d-1}$, any ρ-quantile $\mu^\rho_{\alpha,u}$ belongs to the halfline $\{\lambda u : \lambda \geq 0\}$.*

In case (ii), the origin of $\mathbb{R}^d$ may be a ρ-quantile of order $\alpha > 0$ in direction u. Actually, it can be shown that (a) for $\rho(t) = t$, the origin is a ρ-quantile of order $\alpha > 0$ in direction u if and only $\alpha \leq P[\{0\}]$. Moreover, (b) provided that $\psi_+(0)P[\{0\}] + P[\|Z\| \in (0, \infty) \setminus \mathcal{D}_\rho] = 0$ where Z has distribution P (a condition that always holds for $\rho(t) = t^p$ with $p > 1$), the origin cannot be a ρ-quantile of order $\alpha > 0$ in direction u, so that all these quantiles then belong to $\{\lambda u : \lambda > 0\}$ (for the sake of completeness, we prove (a)–(b) in Appendix S.4; see Proposition S.4.1).

If P is spherically symmetric about the origin and satisfies $P[\{0\}] < 1$, Theorem 3.3 shows that ρ-quantiles are unique for any $\alpha < \alpha_\rho$ (as mentioned above) but it remains silent on the case $\alpha \geq \alpha_\rho$. Interestingly, we will be able to say more under sphericity, thanks to the fact that Theorem 4.1 entails that uniqueness will hold if $t \mapsto M^\rho_{\alpha,u}(tu)$ is strictly convex over $[0, \infty)$ for all $u \in \mathcal{S}^{d-1}$, which in turn will hold if $t \mapsto H^\rho_{\alpha,u}(z - tu)$ is convex for any $z \in \mathbb{R}^d$ and any $u \in \mathcal{S}^{d-1}$. Accordingly, for any $\alpha \in [0, 1]$, let $\mathcal{C}^{\mathrm{sph}}_\alpha$ be the collection of functions $\rho \in \mathcal{C}$ such that $t \mapsto H^\rho_{\alpha,u}(z - tu)$ is convex for any $z \in \mathbb{R}^d$ and $u \in \mathcal{S}^{d-1}$. Since $\mathcal{C}^{\mathrm{sph}}_0 = \mathcal{C}$ and $\mathcal{C}^{\mathrm{sph}}_{\alpha_2} \subseteq \mathcal{C}^{\mathrm{sph}}_{\alpha_1}$ for any $\alpha_1 < \alpha_2$ (see the proof of Theorem 4.2 below), we let $\alpha^{\mathrm{sph}}_\rho := \max\{\alpha \in [0, 1] : \rho \in \mathcal{C}^{\mathrm{sph}}_\alpha\}$, parallel to what we did for α_ρ in Section 3. We have the following result.

Theorem 4.2. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}^\rho_d$. Assume that P is spherically symmetric about the origin of $\mathbb{R}^d$. Then, for any $\alpha \in [0, \alpha^{\mathrm{sph}}_\rho) \cup \{0\}$ and $u \in \mathcal{S}^{d-1}$ (again, union with $\{0\}$*

is needed when $\alpha_\rho^{\rm sph} = 0$), the map $t \mapsto M^\rho_{\alpha,u}(tu)$ is strictly convex on $[0,\infty)$, and the ρ-quantile $\mu^\rho_{\alpha,u}$ is unique.

Note that, in the present spherical setup, this uniqueness result may only strengthen the one in Theorem 3.3, since the fact that $\mathcal{C}_\alpha \subseteq \mathcal{C}_\alpha^{\rm sph}$ for any $\alpha \in [0,1]$ implies that $\alpha_\rho^{\rm sph} \geq \alpha_\rho$. Of course, it is natural to wonder under which conditions on ρ all ρ-quantiles are unique ($\alpha_\rho^{\rm sph} = 1$) and, when these conditions are not met, what are the orders α for which uniqueness is guaranteed (that is, what is then the value of $\alpha_\rho^{\rm sph} < 1$). The following result provides a complete answer to these questions.

Theorem 4.3. *Let $\rho \in \mathcal{C}$. Then, (i) $\alpha_\rho^{\rm sph} = 1$ if and only if $4p_t + q_t - p_t q_t \leq 6$ for any $t \in \mathcal{D}_\rho$, where p_t and q_t are as in Theorem 3.2; (ii) if $\alpha_\rho^{\rm sph} < 1$, then, letting $\mathcal{D}_\rho^{\rm sph} := \{t \in \mathcal{D}_\rho : 4p_t + q_t - p_t q_t > 6\} (\subseteq \mathcal{D}_\rho^{\rm cv})$,*

$$\alpha_\rho^{\rm sph} = \inf_{t \in \mathcal{D}_\rho^{\rm sph}} \sqrt{\beta_{p_t,q_t}},$$

where

$$\beta_{p,q} := \frac{2(pq-p-q)^3(\sqrt{c_{p,q}} - q(2p-3)/3)^2}{3(p-1)^2(3-p)(\sqrt{c_{p,q}} - (2p-q))(\sqrt{c_{p,q}} - q(2p-3))^2}$$

involves $c_{p,q} := \frac{q}{3}(3-2p)(2pq-8p+q)$ (if q makes $\beta_{p,q}$ undefined in the expression above, then we let $\beta_{p,q} := \lim_{r\to q} \beta_{p,r}$).

Parallel to α_ρ in Theorems 3.1–3.2, $\alpha_\rho^{\rm sph}$ does not depend on $d (\geq 2)$. For the power loss functions $\rho(t) = t^p$ with $p \geq 1$, it follows from Theorem 4.3 that

$$\alpha_\rho^{\rm sph} = \begin{cases} \sqrt{\frac{p^2(p-2)^3(b_p - (p-\frac{3}{2})^2)}{(p-1)^2(3-p)(b_p - \frac{3}{2})(b_p - 3(p-\frac{3}{2})^2)}}\,(<1) & \text{if } p \in (2,3) \\ 1 & \text{otherwise,} \end{cases} \tag{6}$$

with $b_p := \big(3(p-\frac{3}{2})(\frac{7}{2}-p)\big)^{1/2}$. For $p \in [1,2]$, the result is just a corollary of Theorem 3.1 since we then have $\alpha_\rho^{\rm sph} \geq \alpha_\rho = 1$. For $p > 3$, (6) implies that all ρ-quantiles are uniquely defined under sphericity, while there is no guarantee that this is the case in general (since $\alpha_\rho < 1$ for such values of p). As shown in Figure 1, the values of $\alpha_p^{\rm sph}$ for $p \in (2,3)$ are remarkably close to one (the minimal value, achieved at $p \approx 2.429$, is about .9987), which implies that, also for $p \in (2,3)$, essentially all ρ-quantiles are uniquely defined under sphericity. For the exponential loss functions $\rho(t) = \exp(ct) - 1$, all ρ-quantiles are also unique under sphericity ($\alpha_\rho^{\rm sph} = 1$), while "only" quantiles of order $\alpha < \alpha_\rho = .9939$ are guaranteed to be unique in general.

5. Differentiability of the objective function

For any $\alpha < \alpha_\rho$, the ρ-quantiles $\mu^\rho_{\alpha,u}$ are minimizers of the convex objective function $\mu \mapsto M^\rho_{\alpha,u}(\mu)$. If this objective function is smooth, then ρ-quantiles are characterized by the first-order condition $\nabla M^\rho_{\alpha,u}(\mu^\rho_{\alpha,u}) = 0$. Such a gradient condition will actually play a key role when deriving further properties of ρ-quantiles in the next sections. This provides a strong motivation to study smoothness of the map $\mu \mapsto M^\rho_{\alpha,u}(\mu)$. We start with the following result.

Proposition 5.1. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}^\rho_d$. Fix $\alpha \in [0,1]$ and $u \in \mathcal{S}^{d-1}$. Let Z be a random d-vector with distribution P and write $Z_\mu := Z - \mu$, for any $\mu \in \mathbb{R}^d$. Then, for any $\mu \in \mathbb{R}^d$ and $v \in \mathbb{R}^d \setminus \{0\}$, the directional derivative*

$$\frac{\partial M^\rho_{\alpha,u}}{\partial v}(\mu) = \lim_{t \overset{>}{\to} 0} \frac{M^\rho_{\alpha,u}(\mu + tv) - M^\rho_{\alpha,u}(\mu)}{t}$$

exists and is given by

$$\frac{\partial M^\rho_{\alpha,u}}{\partial v}(\mu) = \psi_+(0)(\|v\| - \alpha u'v)P[\{\mu\}] - \alpha v'\mathrm{E}\bigg[\frac{\rho(\|Z_\mu\|)}{\|Z_\mu\|}\bigg(I_d - \frac{Z_\mu Z'_\mu}{\|Z_\mu\|^2}\bigg)\xi_{Z,\mu}\bigg]u$$
$$-v'\mathrm{E}\bigg[\{\psi_-(\|Z_\mu\|)\mathbb{I}[v'Z_\mu > 0] + \psi_+(\|Z_\mu\|)\mathbb{I}[v'Z_\mu < 0]\}\bigg(1 + \alpha\frac{u'Z_\mu}{\|Z_\mu\|}\bigg)\frac{Z_\mu}{\|Z_\mu\|}\bigg],$$

where I_d is the $d \times d$ identity matrix and ξ_{z_1,z_2} is as in Definition 1.

The objective function thus admits directional derivatives in all directions (hence, is continuous over $\mathbb{R}^d$), but it is not necessarily differentiable. For instance, the classical spatial quantiles obtained with $\rho(t) = t$ provide

$$\frac{\partial M^\rho_{\alpha,u}}{\partial v}(\mu) = (\|v\| - \alpha u'v)P[\{\mu\}] + v'\mathrm{E}\bigg[\bigg(\frac{\mu - Z}{\|\mu - Z\|} - \alpha u\bigg)\xi_{Z,\mu}\bigg],$$

so that $M^\rho_{\alpha,u}$ fails to be differentiable at atoms of P. Clearly, it follows from Theorem 5.1 that a necessary condition for this objective function to be differentiable at μ is $\psi_+(0)P[\{\mu\}] = 0$. The next result provides a necessary and sufficient condition and gives an expression for the corresponding gradient.

Theorem 5.1. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}^\rho_d$. Fix $\alpha \in [0,1]$ and $u \in \mathcal{S}^{d-1}$. Then, (i) $\mu \mapsto M^\rho_{\alpha,u}(\mu)$ is differentiable at $\mu_0 (\in \mathbb{R}^d)$ if and only if $\psi_+(0)P[\{\mu_0\}] + P[\|Z - \mu_0\| \in (0,\infty) \setminus \mathcal{D}_\rho] = 0$, in which case the corresponding gradient is*

$$\nabla M^\rho_{\alpha,u}(\mu_0) = v(\mu_0) - \alpha T(\mu_0)u, \quad \textit{with} \quad v(\mu) := -\mathrm{E}\bigg[\psi_-(\|Z_\mu\|)\frac{Z_\mu}{\|Z_\mu\|}\xi_{Z,\mu}\bigg]$$

and

$$T(\mu) := \mathrm{E}\bigg[\bigg\{\frac{\rho(\|Z_\mu\|)}{\|Z_\mu\|}\bigg(I_d - \frac{Z_\mu Z_\mu'}{\|Z_\mu\|^2}\bigg) + \psi_-(\|Z_\mu\|)\frac{Z_\mu Z_\mu'}{\|Z_\mu\|^2}\bigg\}\xi_{Z,\mu}\bigg],$$

where $Z_\mu := Z - \mu$ *is based on a random d-vector* Z *with distribution* P. *(ii) If* $\psi_+(0)P[\{\mu\}] + P[\|Z-\mu\| \in (0,\infty)\setminus \mathcal{D}_\rho] = 0$ *for any* μ *in an open set* $\mathcal{N}$, *then* $\mu \mapsto M^\rho_{\alpha,u}(\mu)$ *is continuously differentiable on* $\mathcal{N}$.

It follows from this result that, in contrast with $\rho(t) = t$, the power loss functions $\rho(t) = t^p$ with $p > 1$ make the objective function $M^\rho_{\alpha,u}$ (continuously) differentiable even in the atomic case. The corresponding quantiles $\mu^\rho_{\alpha,u}$ are thus the solutions of the first-order equations $\nabla M^\rho_{\alpha,u}(\mu) = 0$, which rewrite

$$-p\mathrm{E}\bigg[\|Z_\mu\|^{p-1}\frac{Z_\mu}{\|Z_\mu\|}\xi_{Z,\mu}\bigg] = \alpha\mathrm{E}\bigg[\|Z_\mu\|^{p-1}\bigg(I_d + (p-1)\frac{Z_\mu Z_\mu'}{\|Z_\mu\|^2}\bigg)\xi_{Z,\mu}\bigg]u.$$

In particular, spatial expectiles ($p = 2$) of order α in direction u are the unique (Theorem 3.3) solutions of

$$2(\mu - \mathrm{E}[Z]) = \alpha\mathrm{E}\bigg[\|Z-\mu\|\bigg(I_d + \frac{(Z-\mu)(Z-\mu)'}{\|Z-\mu\|^2}\xi_{Z,\mu}\bigg)\bigg]u.$$

This is compatible with the fact that the corresponding "median" (that is, the quantile of order $\alpha = 0$, in an arbitrary direction u) is the mean vector $\mathrm{E}[Z]$.

We turn to second-order differentiability, which will be relevant when studying the asymptotic behavior of sample ρ-quantiles in Section 8.

Theorem 5.2. *Let* $\rho \in \mathcal{C}$ *and* $P \in \mathcal{P}^\rho_d$. *Fix* $\alpha \in [0,1]$, $u \in \mathcal{S}^{d-1}$, *and* $\mu_0 \in \mathbb{R}^d$. *Assume that* $P[\|Z-\mu\| \in [0,\infty)\setminus\mathcal{D}_\rho] = 0$ *for any* μ *in an open neighbourhood of* μ_0 *(hence, in particular, that* P *is non-atomic in this neighbourhood). Let further one of the following assumptions hold:*

(A) ψ_- *is concave on* $(0,\infty)$ *and*

$$\int_{\mathbb{R}^d\setminus\{\mu_0\}} \frac{\psi_-(\|z-\mu_0\|)}{\|z-\mu_0\|}\, dP(z) < \infty;$$

(A′) ψ_- *is convex on* $(0,\infty)$, $\psi_+(0) = 0$, *and there exists* $r > 0$ *such that*

$$\int_{\mathbb{R}^d} \psi_-'(\|z-\mu_0\| + r)\, dP(z) < \infty$$

(recall that ψ_-' *is the left-derivative of* ψ_-*).*

Then, for any $v \in \mathbb{R}^d \setminus \{0\}$,

$$\lim_{t \overset{>}{\to} 0} \frac{\nabla M^{\rho}_{\alpha,u}(\mu_0 + tv) - \nabla M^{\rho}_{\alpha,u}(\mu_0)}{t} = \nabla^2 M^{\rho}_{\alpha,u}(\mu_0)v,$$

where the Hessian matrix $\nabla^2 M^{\rho}_{\alpha,u}(\mu)$ is given by

$$\begin{aligned}\nabla^2 M^{\rho}_{\alpha,u}(\mu) &= \big(\partial_i \partial_j M^{\rho}_{\alpha,u}(\mu)\big)_{i,j=1,\ldots,d} \\ &= \mathrm{E}\Bigg[\bigg(\psi'_-(\|Z_\mu\|) - \frac{2\psi_-(\|Z_\mu\|)}{\|Z_\mu\|} + \frac{2\rho(\|Z_\mu\|)}{\|Z_\mu\|^2}\bigg)\bigg(1 + \alpha\frac{u'Z_\mu}{\|Z_\mu\|}\bigg)\frac{Z_\mu Z'_\mu}{\|Z_\mu\|^2}\xi_{Z,\mu} \\ &\quad + \frac{\rho(\|Z_\mu\|)}{\|Z_\mu\|^2}\bigg(I_d - \frac{Z_\mu Z'_\mu}{\|Z_\mu\|^2}\bigg)\xi_{Z,\mu} + \frac{\|Z_\mu\|\psi_-(\|Z_\mu\|) - \rho(\|Z_\mu\|)}{\|Z_\mu\|^2} \\ &\quad \times\bigg\{\bigg(1 + \alpha\frac{u'Z_\mu}{\|Z_\mu\|}\bigg)\bigg(I_d - \frac{Z_\mu Z'_\mu}{\|Z_\mu\|^2}\bigg) + 2\frac{Z_\mu Z'_\mu}{\|Z_\mu\|^2} + \alpha\frac{Z_\mu u' + Z'_\mu u}{\|Z_\mu\|}\bigg\}\xi_{Z,\mu}\Bigg]\end{aligned}$$

(*as in the previous results, $Z_\mu := Z - \mu$, where Z is a random d-vector with distribution P*).

While they may seem complex at first, the assumptions of Theorem 5.2 turn out to be simple (and very weak) when considering specific loss functions ρ. For instance, for $\rho(t) = t^p$ with $p \geq 1$, they only require that $P \in \mathcal{P}^{\rho}_d$ is non-atomic in a neighborhood of μ_0 and is such that $\mathrm{E}[\|Z - \mu_0\|^{p-2}] < \infty$ when Z has distribution P. Note that this last assumption, that cannot be avoided since this expectation is involved in the Hessian matrix $\nabla^2 M^{\rho}_{\alpha,u}(\mu)$, is superfluous for $p \geq 2$. Under the assumptions of Theorems 3.3 and 5.2, this Hessian matrix is positive definite for any $\alpha \in [0, \alpha_\rho) \cup \{0\}$ and any $u \in \mathcal{S}^{d-1}$; since this will be needed in the sequel, we prove it in Appendix S.8 (see Lemma S.8.2).

6. A ρ-version of Robert Serfling's DOQR paradigm

In a series of papers, Robert Serfling introduced the *DOQR paradigm*, that presents *Depth*, *Outlyingness*, *Quantile* and *Rank* functions as interrelated, yet distinct, objects of interest for multivariate nonparametric statistics; see, e.g., Serfling (2010), Serfling (2019), Serfling and Zuo (2010) and the references therein. While this paradigm in principle applies to any multivariate quantile concept, the primary focus when considering this paradigm in the aforementioned papers was on spatial quantiles. This makes it natural to study the paradigm for the generalized spatial quantiles considered in this work, which leads to introducing ρ-depth, ρ-outlyingness, ρ-quantile and ρ-rank functions. As we will see later, some of these functions play a key role to understand the nature of extreme ρ-quantiles.

We start by formally defining ρ-quantile functions. Restricting to the interesting case for which $\alpha_\rho > 0$, Theorems 2.1 and 3.3 imply that ρ-quantiles exist and are unique for any $\alpha \in [0, \alpha_\rho)$ and $u \in \mathcal{S}^{d-1}$, which allows us to adopt the following definition.

Definition 2. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}_\rho^d$. Assume that there is no open interval in $(0, \infty)$ on which ψ_- is constant or that P is not concentrated on a line. Write $\mathcal{B}_r^d = \{z \in \mathbb{R}^d : \|z\| < r\}$. Then, the* ρ-quantile function *of P is the map $Q = Q_P^\rho : \mathcal{B}_{\alpha_\rho}^d \to \mathbb{R}^d$ that is defined through $Q(\alpha u) = \mu_{\alpha,u}^\rho$.*

In dimension $d = 1$ and $\rho(t) = t$, this provides the (centered-outward version of the) usual quantile function. This standard quantile function, that is defined on $\mathcal{B}_1^1 = (-1, 1)$, may of course fail to be continuous (it is discontinuous for empirical probability measures). The multivariate case $d \geq 2$ is different.

Proposition 6.1. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}_\rho^d$, with $d \geq 2$. Assume that there is no open interval in $(0, \infty)$ on which ψ_- is constant or that P is not concentrated on a line. Then, the quantile function $Q = Q_P^\rho : \mathcal{B}_{\alpha_\rho}^d \to \mathbb{R}^d$ is continuous.*

Following Serfling (2010), we associate with the ρ-quantile function Q corresponding concepts of *rank function* R, *depth function* D and *outlyingness function* O. We start with the rank function.

Definition 3. *Let $\rho \in \mathcal{C}$ and assume that $P \in \mathcal{P}_\rho^d$ is not a Dirac probability measure. Then, the* rank function *of P is the map $R = R_P^\rho : \mathbb{R}^d \to \mathbb{R}^d$ defined through $R(\mu) = (T(\mu))^{-1} v(\mu)$, where the $d \times d$ matrix $T(\mu)$ and the d-vector $v(\mu)$ were introduced in Theorem 5.1.*

In the setup of this definition, $T(\mu)$ is positive definite, hence invertible, for any $\mu \in \mathbb{R}^d$ (for the sake of completeness, we prove this in Appendix S.6; see Lemma S.6.1). The natural assumptions under which to study the rank function are those of Theorem 5.1 complemented by conditions ensuring uniqueness of ρ-quantiles (which provides the assumptions in Theorem 6.1 below). Under these assumptions, $\mu \mapsto M_{\alpha,u}^\rho(\mu)$ is continuously differentiable on $\mathbb{R}^d$, with gradient

$$\nabla M_{\alpha,u}^\rho(\mu) = T(\mu)(R(\mu) - \alpha u),$$

so that $\mu = \mu_{\alpha,u} = Q(\alpha u)$ (for $\alpha < \alpha_\rho$) if and only if $R(\mu) = \alpha u$ (recall that, under the assumptions considered, quantiles of order $\alpha \in [0, \alpha_\rho)$ in direction $u \in \mathcal{S}^{d-1}$ are indeed uniquely determined by the gradient condition $\nabla M_{\alpha,u}^\rho(\mu) = 0$). This provides a clear interpretation of the rank function as the inverse map of the quantile function. We have the following result.

Theorem 6.1. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}_d^\rho$. Assume that there is no open interval in $(0,\infty)$ on which ψ_- is constant or that P is not concentrated on a line. Assume further that $\psi_+(0)P[\{\mu\}] + P[\|Z-\mu\| \in (0,\infty) \setminus \mathcal{D}_\rho] = 0$ for any $\mu \in \mathbb{R}^d$. Write $\mathcal{Z}_\rho = Q_P^\rho(\mathcal{B}^d_{\alpha_\rho})$. Then, $Q = Q_P^\rho : \mathcal{B}^d_{\alpha_\rho} \to \mathcal{Z}_\rho$ is a homeomorphism, with inverse $R^\rho_{P|\mathcal{Z}_\rho} : \mathcal{Z}_\rho \to \mathcal{B}^d_{\alpha_\rho}$ (the restriction of R_P^ρ to $\mathcal{Z}_\rho$).*

If $t \mapsto t^2/\rho(t)$ is concave on $(0,\infty)$, then we are in the important particular case $\alpha_\rho = 1$ (Theorem 3.1), for which the quantile function Q is defined on the open unit ball $\mathcal{B}^d = \mathcal{B}^d_1$. We then have the following result.

Theorem 6.2. *Let $\rho \in \mathcal{C}$ be such that $t \mapsto t^2/\rho(t)$ is concave on $(0,\infty)$. Assume that P is not concentrated on a line and that $\psi_+(0)P[\{\mu\}] + P[\|Z-\mu\| \in (0,\infty) \setminus \mathcal{D}_\rho] = 0$ for any $\mu \in \mathbb{R}^d$. Then, $\mathcal{Z}_\rho = Q_P^\rho(\mathcal{B}^d) = \mathbb{R}^d$, so that $Q = Q_P^\rho : \mathcal{B}^d \to \mathbb{R}^d$ is a homeomorphism, with inverse $R = R_P^\rho : \mathbb{R}^d \to \mathcal{B}^d$.*

This result shows in particular that for any power loss function $\rho(t) = t^p$ with $p \in [1,2]$, any non-atomic probability measure that is not concentrated on a line provides ρ-quantiles that span the whole Euclidean space $\mathbb{R}^d$ (the non-atomicity condition is actually needed for $p = 1$ only), whereas the result remains silent for the case $p > 2$. This will have important implications when studying extreme quantiles in Section 7.

Let us turn to depth and outlyingness functions. Clearly, central or "deep" quantiles are indexed by a small order $\alpha \in [0,1)$, whereas exterior or "outlying" ones are rather indexed by a large order α. A natural outlyingness measure for $\mu \in \mathbb{R}^d$ is then the order α of the quantile $\mu_{\alpha,u}$ for which $\mu = \mu_{\alpha,u}$, that is, the oultlyingness of μ is $\|R(\mu)\|$. Any decreasing function of this outlyingness measure is then a natural depth measure. We adopt the following definition.

Definition 4. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}_\rho^d$. Then, (i) the* outlyingness function *of P is the map $O = O_P^\rho$ from $\mathbb{R}^d$ to $[0,1]$ defined through $O(\mu) = \min(\|R(\mu)\|, 1)$, where $R = R_P^\rho$ is the rank function of P. (ii) The* depth function *of P is the map $D = D_P^\rho$ from $\mathbb{R}^d$ to $[0,1]$ defined through $D(\mu) = 1 - O(\mu)$.*

The deepest location, the only one that receives the maximal depth value one, is the ρ-median $\mu^\rho_{0,u}$ of P (the direction u plays no role for $\alpha = 0$). For any direction $u \in \mathcal{S}^{d-1}$, depth decreases along the quantile curves $\{\mu^\rho_{\alpha,u} : \alpha \in [0,\alpha_\rho)\}$ originating from the ρ-median. For $\rho(t) = t$, this depth reduces to the celebrated *spatial depth*; see, e.g., Vardi and Zhang (2000). The depth that are associated with our ρ-quantiles extend this classical depth; in particular, an "expectile spatial depth", whose deepest point is the mean vector of P, is obtained for $\rho(t) = t^2$. For any depth function, the depth regions collecting locations with depth exceeding a given threshold are of interest. The depth regions

$$\mathcal{R}_\alpha^\rho = \mathcal{R}_{P,\alpha}^\rho := \{\mu \in \mathbb{R}^d : D_P^\rho(\mu) \geq \alpha\}$$

are nested "centrality regions"; see, e.g., Mosler (2012) and the references therein. The corresponding depth contours, i.e. the boundaries $\partial\mathcal{R}_\alpha^\rho$ of these depth regions, collect the ρ-quantiles associated with a fixed order α.

For each combination of $\alpha \in \{.25, .50, .75\}$ and $p \in \{1, 1.5, 2, 4\}$, we plot in Figure 2 the depth contours of order α, based on $\rho(t) = t^p$, for the empirical probability measure P_n of six random samples of size $n = 200$ (these were obtained from a uniform grid of 50 directions on the unit circle $\mathcal{S}^{d-1}$, and each quantile was evaluated through the descent method involving the backtracking line search in Section 9.2 of Boyd and Vandenberghe, 2004; R code is available on request). These samples were generated from (i) the bivariate standard normal distribution, (ii)–(iii) the standard t-distributions with $\nu = 4$ and $\nu = 1$ degrees of freedom, (iv) the centered bivariate normal distribution with covariance matrix $\Sigma = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, (v) the bivariate distribution whose marginals are independent exponential distributions with mean one, (vi) the standard skew-t distribution with 4 degrees of freedom and slant vector $\alpha = (10, 10)$; see Azzalini and Capitanio (2014). Figure 2 shows that larger values of p provide contours that are more concentrated about the corresponding median; the only exception is the Cauchy distribution, for which these large-p contours are the most spread ones due to their lack of robustness with respect to extreme observations. As expected, the various medians differ when the underlying distribution is skewed, as it is the case in (v)–(vi).

7. Extreme quantiles

Recently, Girard and Stupfler (2015, 2017) studied the spatial quantiles from Chaudhuri (1996) with a focus on extreme quantiles, that is, those associated with an order α that is close to one. In particular, Girard and Stupfler (2017) derived striking results on extreme quantiles showing that (i) spatial quantiles exit any compact set as $\alpha \to 1$ and that (ii) they do so in a direction that eventually coincides with the direction u in which quantiles are computed. Surprisingly, this typically also happens when the underlying distribution P is compactly supported. As shown in Paindaveine and Virta (2021), the result even holds under atomic probability measures P, so that this unexpected behavior also shows in the sample case (provided that not all observations lie on a line of $\mathbb{R}^d$).

Of course, it is natural to ask whether or not this behavior of extreme quantiles shows for other ρ-quantiles. We tackle this question in the present section. Our first result is the following.

Theorem 7.1. *Let $\rho \in \mathcal{C}$ be such that $t \mapsto t^2/\rho(t)$ is concave on $(0, \infty)$. Assume that P is not concentrated on a line and that $\psi_+(0)P[\{\mu\}] + P[\|Z - \mu\| \in (0, \infty) \setminus \mathcal{D}_\rho] = 0$ for*

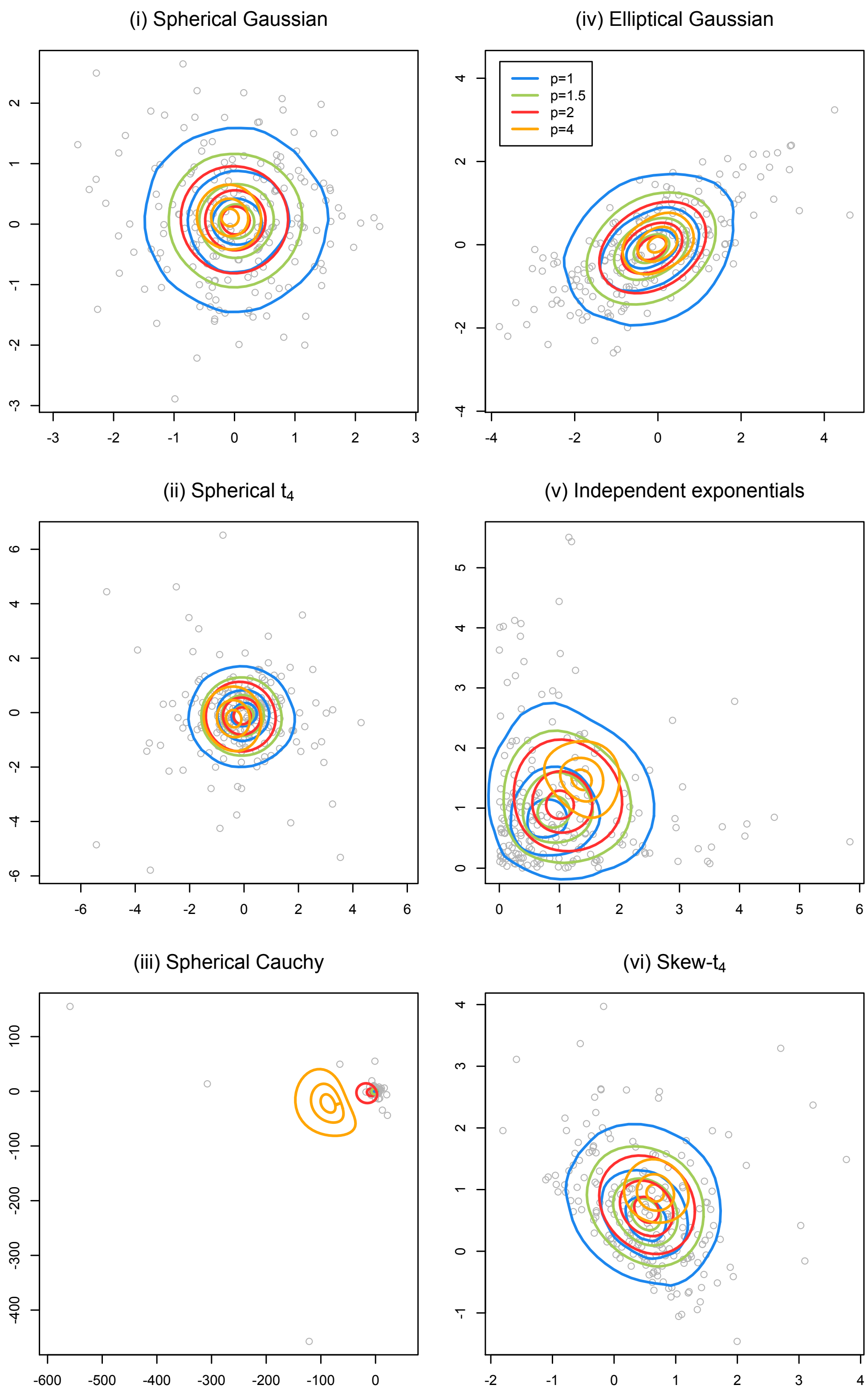

Figure 2. For $\rho(t) = t^p$ with $p = 1, 1.5, 2, 4$, ρ-depth contours of order $\alpha = .25, .50, .75$ for random samples of size $n = 200$ drawn from six bivariate distributions; see Section 6 for details.

any $\mu \in \mathbb{R}^d$. Let (α_n) be a sequence in $[0,1)$ that converges to one and (u_n) be a sequence in $\mathcal{S}^{d-1}$. Then, (i) $\|\mu^\rho_{\alpha_n,u_n}\| \to \infty$; (ii) if $u_n \to u$, then $\mu^\rho_{\alpha_n,u_n}/\|\mu^\rho_{\alpha_n,u_n}\| \to u$.

This result shows that all ρ-quantiles for which $\alpha_\rho = 1$, hence in particular those associated with $\rho(t) = t^p$ for $p \in (1,2]$, will show the behavior of the extreme quantiles from Chaudhuri (1996) described above. Note that for $p \in (1,2]$, we have $\psi_+(0) = 0$ and $\mathcal{D}_\rho = (0,\infty)$, so that Theorem 7.1 does not require that P is non-atomic, hence also allows for empirical distributions. We illustrate this in Figure 3 for $P = P_n$, the empirical distribution of a random sample of size $n = 10$ drawn from the bivariate standard normal distribution. For $\rho(t) = t^p$, with $p \in \{1, 1.5, 2, 2.25, 3, 4\}$, the figure shows the ρ-quantiles $\mu^\rho_{\alpha,u}$, for $\alpha \in [0,1)$ and $u = (\cos(\pi\ell/6), \sin(\pi\ell/6))$, with $\ell = 0,1,2,3$. Clearly, for the values of p that are covered by Theorem 7.1, namely $p = 1, 1.5, 2$, quantiles exit any compact set and do so eventually in the corresponding direction u. In contrast, the figure suggests that, for $p > 2$, the Euclidean norm of extreme ρ-quantiles remains bounded. This is indeed the case, as the following result shows.

Theorem 7.2. *Let $\rho \in \mathcal{C}$ such that $\rho(t)/t^2 \to \infty$ as $t \to \infty$. Assume that $P \in \mathcal{P}^\rho_d$ (a) is not concentrated on a line of $\mathbb{R}^d$ and (b) satisfies $\int_{\mathbb{R}^d} \rho(\|z\|)\, dP(z) < \infty$ (if $\rho(t)/t^3$ is bounded away from 0 as $t \to \infty$, then Condition (b) is superfluous). Then, there exists a bounded set $S \subset \mathbb{R}^d$ such that, for any $\alpha \in [0,1)$ and $u \in \mathcal{S}^{d-1}$, all ρ-quantiles of order α in direction u belong to S (moreover, $D(\mu) = 0$ for any $\mu \in \mathbb{R}^d \setminus S$).*

Under the condition of this result, all ρ-quantiles of order α in direction u may fail to be unique for $\alpha \in (\alpha_\rho, 1)$, which is the reason why Theorem 7.2 states that all ρ-quantiles of order α in direction u belong to S. The result implies that for $\rho(t) = t^p$ with $p \geq 3$, extreme ρ-quantiles are bounded as soon as $P \in \mathcal{P}^\rho_d$ is not concentrated on a line and that, for $\rho(t) = t^p$ with $p \in (2,3)$, the same holds provided that P further has finite moments of order p rather than finite moments of order $p-1$ only (we conjecture that this stronger moment assumption for $p \in (2,3)$ is actually superfluous, but we were not able to avoid this assumption when proving Theorem 7.2). Note that Theorem 7.2 confirms in particular that, in Figure 3, the ρ-quantiles associated with $p > 2$ form a bounded set.

As mentioned in Section 2, ρ-quantiles in principle are not defined for $\alpha = 1$, but of course Definition 1 may still be adopted to define possible quantiles for order $\alpha = 1$. We have the following existence result.

Proposition 7.1. *Let the assumptions of Theorem 7.2 hold. Then, for any $u \in \mathcal{S}^{d-1}$, there exists a quantile $\mu^\rho_{1,u}$.*

In contrast, it directly follows from Theorem 6.2 that, under the assumptions of Theorem 7.1, there is no $u \in \mathcal{S}^{d-1}$ for which a quantile $\mu^\rho_{1,u}$ exist (this result is already known

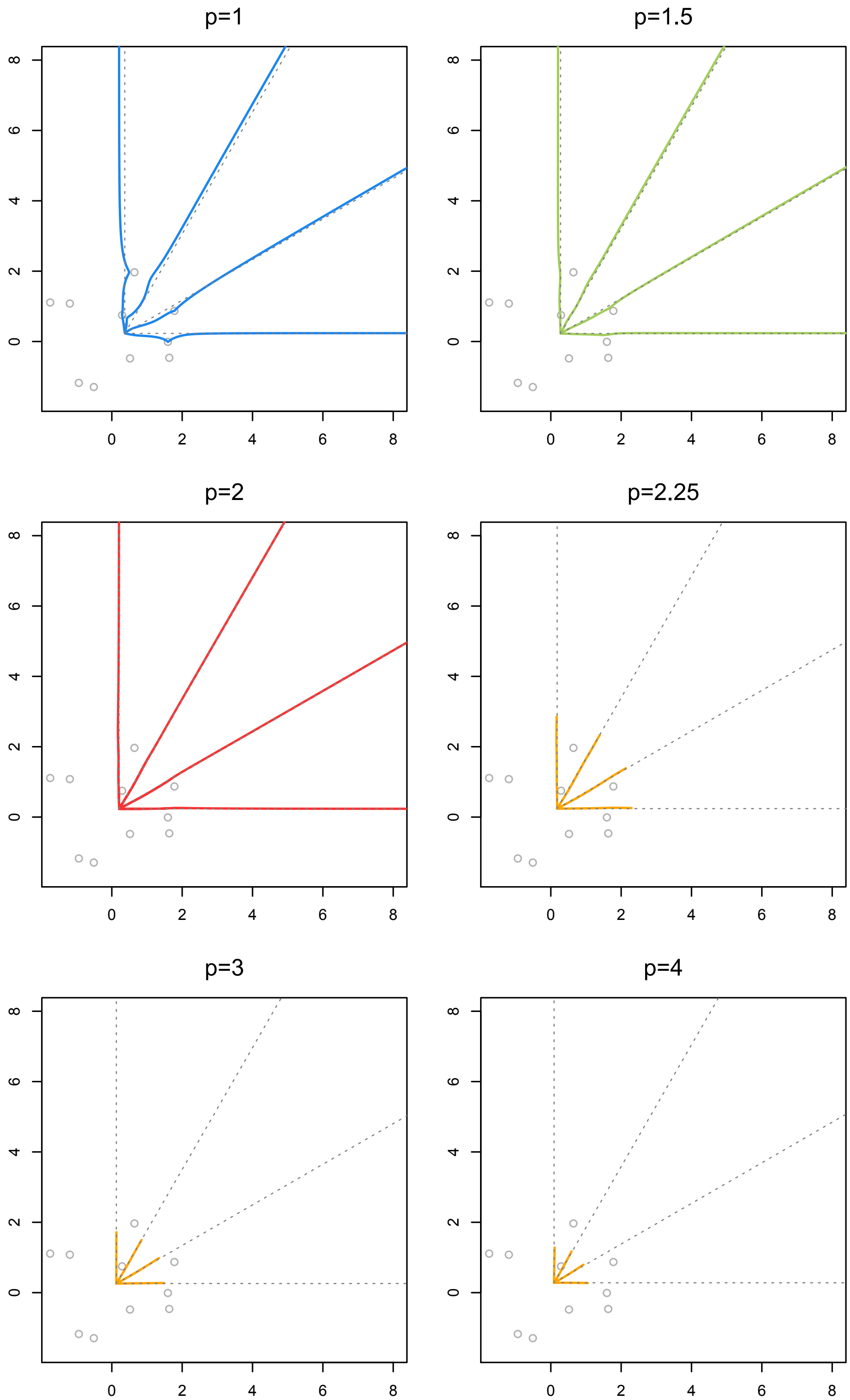

Figure 3. For the loss functions $\rho(t) = t^p$ with $p = 1, 1.5, 2, 2.25, 3, 4$, the plots show the ρ-quantiles $\mu^{\rho}_{\alpha,u}$ for $\alpha \in [0, 1)$ and $u = (\cos(\pi\ell/6), \sin(\pi\ell/6))$, with $\ell = 0, 1, 2, 3$; the underlying probability measure P is the empirical distribution P_n associated with a random sample of size $n = 10$ from the bivariate standard normal distribution. Dashed lines are showing the halflines with the corresponding directions u originating from the median $\mu^{\rho}_{0,u}$.

for $\rho(t) = t$; see Proposition 2.1 in Girard and Stupfler, 2017).

The following corollary of Theorem 7.2 extends in some sense the continuity of the quantile function (Proposition 6.1) to the framework where quantiles of order $\alpha = 1$ exist.

Corollary 7.1. *Let the assumptions of Theorem 7.2 hold. Let (α_n) be a sequence in $[0, 1)$ that converges to $\alpha \in [0, 1]$ and (u_n) be a sequence in $\mathcal{S}^{d-1}$ that converges to $u (\in \mathcal{S}^{d-1})$. Fix an arbitrary sequence $(\mu^\rho_{\alpha_n,u_n})$ of ρ-quantiles. Then, (i) any converging subsequence of $(\mu^\rho_{\alpha_n,u_n})$ converges to a ρ-quantile $\mu^\rho_{\alpha,u}$; (ii) if $\mu^\rho_{\alpha,u}$ is unique, then $\mu^\rho_{\alpha_n,u_n} \to \mu^\rho_{\alpha,u}$.*

This result further confirms that the quantile functions—hence also the rank, depth and outlyingness functions—associated with the loss functions ρ covered by Theorem 7.1 and Theorem 7.2 are very different in nature. In particular, in the framework of Theorem 7.2, the depth of μ will be exactly zero if $\|\mu\|$ is large enough. Some recent research efforts in the statistical depth literature aimed at defining depth functions—or at modifying existing depth functions—that do not show this "vanishing property"; see, e.g., Francisci, Nieto-Reyes and Agostinelli (2019) and the many references therein. This vanishing property is indeed undesirable in some inferential applications, such as, e.g., supervised classification based on the max-depth approach; see Francisci, Nieto-Reyes and Agostinelli (2019), Ghosh and Chaudhuri (2005), and Li, Cuesta-Albertos and Liu (2012). Quite nicely, the ρ-depths associated with loss functions ρ compatible with Theorem 7.1 will not exhibit this vanishing property. Yet, as in Girard and Stupfler (2017), some might find it shocking that the corresponding ρ-quantiles span the whole Euclidean space even when P is compactly supported. This can be avoided by adopting a loss function ρ meeting the conditions of Theorem 7.2. As a conclusion, while Theorems 7.1–7.2 discriminate between two fundamentally different classes of DOQR functions, none of these two worlds is "the good one" and the choice of ρ, hence the choice among both worlds, should be performed based on the inferential problem at hand.

8. Asymptotics for point estimation

We now consider estimation of the ρ-quantiles $\mu^\rho_{\alpha,u} = \mu^\rho_{\alpha,u}(P)$ based on a random sample $Z_1, \ldots, Z_n$ from P. As usual, the natural estimator is obtained by replacing P with the corresponding empirical probability measure. In this section, we study the asymptotic properties of the resulting sample ρ-quantiles. We start with the following consistency result.

Theorem 8.1. *Fix $\rho \in \mathcal{C}$ and assume that there is no open interval in $(0, \infty)$ on which ψ_- is constant or that $P \in \mathcal{P}^d_\rho$ is not concentrated on a line. Denote as P_n the* *empirical probability measure associated with a random sample of size n from P. Fix $\alpha \in [0, \alpha_\rho) \cup \{0\}$, $u \in \mathcal{S}^{d-1}$, and write $\hat{\mu}^\rho_{\alpha,u} = \mu^\rho_{\alpha,u}(P_n)$. Then,*

$$\hat{\mu}^\rho_{\alpha,u} \to \mu^\rho_{\alpha,u}$$

almost surely as $n \to \infty$.

The sample spatial median, that is, the median obtained with the loss function $\rho(t) = t$, satisfies a classical asymptotic normality result (see, e.g., Möttönen et al. (2010)), which, as usual, allows one to perform hypothesis testing or to build confidence zones for the population spatial median. This is an important advantage over competing multivariate medians, that exhibit so complicated asymptotic distributions that it is not possible to base inference on them (this is in particular the case for the celebrated Tukey median; see Massé (2002)). Quite nicely, all sample ρ-quantiles enjoy a standard asymptotic normality result, relying on a neat Bahadur representation result (that typically may itself have further applications, such as the derivation of LIL results). We have the following result.

Theorem 8.2. *Let $\rho \in \mathcal{C}$ and $P \in \mathcal{P}^\rho_d$. Assume that there is no open interval in $(0, \infty)$ on which ψ_- is constant or that P is not concentrated on a line. Fix $\alpha \in [0, \alpha_\rho) \cup \{0\}$ and $u \in \mathcal{S}^{d-1}$. Assume that*

$$\int_{\mathbb{R}^d} \psi^2_-(\|z - \mu^\rho_{\alpha,u}\|)\, dP(z) < \infty$$

and that $P[\|Z - \mu\| \in [0, \infty) \setminus \mathcal{D}_\rho] = 0$ for any μ in an open neighborhood of $\mu^\rho_{\alpha,u}$ (hence, in particular, that P is non-atomic in this neighborhood). Let further one of the following assumptions hold:

(A) ψ_- is concave on $(0, \infty)$ and

$$\int_{\mathbb{R}^d \setminus \{\mu^\rho_{\alpha,u}\}} \frac{\psi_-(\|z - \mu^\rho_{\alpha,u}\|)}{\|z - \mu^\rho_{\alpha,u}\|}\, dP(z) < \infty;$$

(A′) ψ_- is convex on $(0, \infty)$, $\psi_+(0) = 0$, and there exists $r > 0$ such that

$$\int_{\mathbb{R}^d} \psi'_-(\|z - \mu^\rho_{\alpha,u}\| + r)\, dP(z) < \infty$$

(recall that ψ'_- is the left-derivative of ψ_-).

Let $\hat{\mu}^\rho_{\alpha,u} = \mu^\rho_{\alpha,u}(P_n)$, where P_n is the empirical probability measure associated with a *random sample $Z_1, \ldots, Z_n$ of size n from P. Then,*

$$\begin{aligned}\sqrt{n}(\hat{\mu}^{\rho}_{\alpha,u} - \mu^{\rho}_{\alpha,u}) &\\ &= \frac{1}{\sqrt{n}} A^{-1} \sum_{i=1}^{n} \nabla H^{\rho}_{\alpha,u}(Z_i - \mu^{\rho}_{\alpha,u}) \mathbb{I}[\|Z_i - \mu^{\rho}_{\alpha,u}\| \in \mathcal{D}_\rho] + o_{\mathrm{P}}(1) \\ &\stackrel{\mathcal{D}}{\to} \mathcal{N}_d(0, V)\end{aligned}$$

as $n \to \infty$, where $V := A^{-1}BA^{-1}$ involves $A := \nabla^2 M^{\rho}_{\alpha,u}(\mu^{\rho}_{\alpha,u})$ and $B := \mathrm{E}[(\nabla H^{\rho}_{\alpha,u}(Z_1 - \mu^{\rho}_{\alpha,u}))(\nabla H^{\rho}_{\alpha,u}(Z_1 - \mu^{\rho}_{\alpha,u}))' \mathbb{I}[\|Z_1 - \mu^{\rho}_{\alpha,u}\| \in \mathcal{D}_\rho]]$.

We stress that this result requires very mild assumptions only. In particular, for the power loss functions $\rho(t) = t^p$ with $p \geq 2$, it only requires that P is non-atomic in a neighborhood of $\mu^{\rho}_{\alpha,u}$ and admits finite moments of order $2(p-1)$ (for the median obtained with $p = 2$, namely the mean, this is the usual finite second-order moment assumption, and the result only restate the usual multivariate central limit theorem, but for the mild local non-atomicity assumption). For $p \in [1, 2)$, Theorem 8.2 further requires that $\mathrm{E}[\|Z - \mu^{\rho}_{0,u}\|^{p-2}]$ exists and is finite (note that, for the spatial median ($p = 1$), Möttönen et al. (2010) derives the result under assumptions that are more stringent, since it is imposed there that $\mathrm{E}[\|Z - \mu^{\rho}_{0,u}\|^{-r}]$ exists and is finite for any $r \in [0, 2)$. Invertibility of A is always guaranteed; see Lemma S.8.2 in Appendix S.8.

To illustrate the result, we focus on ρ-medians ($\alpha = 0$) under sphericity. If P is spherically symmetric about the origin of $\mathbb{R}^d$, then all ρ-medians $\mu^{\rho}_{0,u}$ are equal to each other (they coincide with the origin of $\mathbb{R}^d$; see Theorem 4.1), which makes it valid to compare the asymptotic variances of sample ρ-medians. We consider the power loss functions $\rho(t) = t^p$ with $p \geq 1$, for which

$$\nabla H^{\rho}_{\alpha,u}(x)(\nabla H^{\rho}_{\alpha,u}(x))' \mathbb{I}[x \in \mathcal{D}_\rho] = p^2 \|x\|^{2(p-1)} \frac{xx'}{\|x\|^2} \xi_{x,0}$$

and

$$\nabla^2 H^{\rho}_{\alpha,u}(x) \mathbb{I}[x \in \mathcal{D}_\rho] = \|x\|^{p-2} \left\{ pI_d + p(p-2) \frac{xx'}{\|x\|^2} \right\} \xi_{x,0};$$

see Lemma S.5.1 in Appendix S.5. If P is spherically symmetric about the origin of $\mathbb{R}^d$, then $\|Z\|$ and $Z/\|Z\|$ are mutually independent, with $Z/\|Z\|$ uniformly distributed over $\mathcal{S}^{d-1}$, which yields

$$B = \frac{p^2}{d} \mathrm{E}[\|Z\|^{2(p-1)}] I_d \quad \text{and} \quad A = \frac{p(d+p-2)}{d} \mathrm{E}[\|Z\|^{p-2} \xi_{Z,0}] I_d.$$

Thus, the asymptotic covariance matrix V is given by

$$V = A^{-1}BA^{-1} = \frac{d\mathrm{E}[\|Z\|^{2(p-1)}]}{(d+p-2)^2 (\mathrm{E}[\|Z\|^{p-2}])^2} I_d =: v_p(P) I_d. \tag{7}$$

For $p = 1$, this reduces to the asymptotic covariance matrix of the spatial median (see Möttönen et al. (2010)), whereas, for $p = 2$, this provides the asymptotic covariance matrix $V = \mathrm{E}[ZZ']$ of the sample mean. Let us consider various spherical distributions. If $P = P^t_\nu$ is the d-variate t-distribution with ν degrees of freedom, then $\|Z\|^2/d$ is Fisher–Snedecor with d and ν degrees of freedom, which yields

$$v_p(P^t_\nu) = \frac{\Gamma(\frac{d+2}{2})\Gamma(\frac{d+2p-2}{2})\Gamma(\frac{\nu+2}{2})\Gamma(\frac{\nu-2p+2}{2})}{\Gamma^2(\frac{d+p}{2})\Gamma^2(\frac{\nu-p+2}{2})} \tag{8}$$

for $\nu > 2(p-1)$, whereas if $P = P^e_\eta$ is the d-variate power-exponential distribution with tail parameter $\eta (> 0)$, then

$$v_p(P^e_\eta) = \frac{2^{(1-\eta)/\eta}\Gamma(\frac{d+2\eta}{2\eta})\Gamma(\frac{d+2p-2}{2\eta})}{\eta\Gamma^2(\frac{d+p+2\eta-2}{2\eta})}; \tag{9}$$

the power-exponential distribution with tail parameter η refers to the distribution admitting the density $z \mapsto f^e_\eta(z) := c_{d,\eta} \exp(-\|z\|^{2\eta}/2)$ with respect to the Lebesgue measure over $\mathbb{R}^d$ ($c_{d,\eta}$ is a normalizing constant). The asymptotic variance at the standard d-variate normal distribution is obtained by taking $\nu \to \infty$ in (8) or, alternatively, by taking $\eta = 1$ in (9).

The factors $v_p(P^t_\nu)$ and $v_p(P^e_\eta)$, that completely characterize the asymptotic covariance matrix of the sample ρ-median associated with $\rho(t) = t^p$ under the corresponding distributions, are plotted in Figure 4. For heavy tails, the medians associated with a small value of p dominate their competitors, whereas the opposite happens for light tails (lighter-than-normal tails are obtained for $\eta > 1$ in the power-exponential case). Note that the sample ρ-median associated with $\rho(t) = t^p$ is the maximum likelihood estimator of the symmetry center in the location family generated by power-exponential distributions with parameter $\eta = p/2$, which explains that large values of p ($p > 2$) will behave well under lighter-than-normal tails. All in all, the median associated with $p = 1.5$ seems to provide a nice balance between the spatial median and sample mean associated with $p = 1$ and $p = 2$, respectively. While these considerations are specific to the spherical case, the efficiency of ρ-medians in the elliptical case could be studied following the analysis in Magyar and Tyler (2011), where the focus was exclusively on the spatial median ($p = 1$).

To check correctness of Theorem 8.2, we performed a Monte-Carlo study involving the bivariate ($d = 2$) t-distributions with ν degrees of freedom with $\nu \in \{3, 5, 7, \ldots, 21\}$, and the bivariate power-exponential distributions with parameter $\eta \in \{.8, 1.2, 1.6, \ldots, 4\}$. For each of these distributions, we generated $M = 10{,}000$ random samples of size $n = 200$ and evaluated the ρ-medians $\hat{\mu}^\rho_{0,u} = \hat{\mu}^\rho_{0,u}(m)$ associated with $\rho(t) = t^p$ for $p \in \{1, 1.5, 2, 4\}$ in

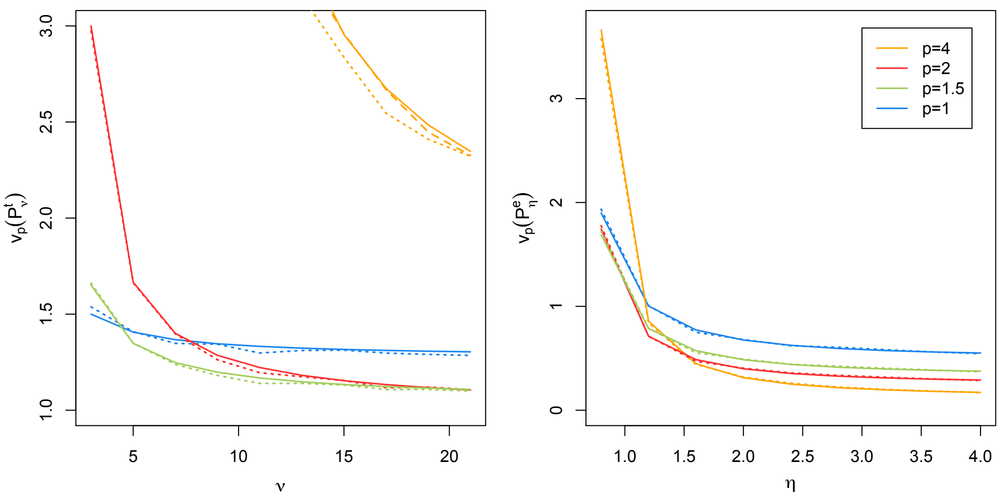

Figure 4. (Left:) For $p \in \{1, 1.5, 2, 4\}$, plots of $\nu \mapsto v_p(P^t_\nu)$ for $\nu = 3, 5, 7, \ldots, 21$, where $v_p(P^t_\nu)$ (see (8)) is the factor characterizing the asymptotic covariance matrix, at the bivariate t-distribution with ν degrees of freedom, of the sample ρ-median based on $\rho(t) = t^p$. Dotted lines are estimates of $v_p(P^t_\nu)$ computed from $M = 10{,}000$ random samples of size $n = 200$. For $p = 4$, the dashed line provides the same result for random samples of size $n = 1{,}000$. (Right:) Still for $p \in \{1, 1.5, 2, 4\}$, plots of $\eta \mapsto v_p(P^e_\eta)$ for $\eta = .8, 1.2, 1.6, \ldots, 4$, where $v_p(P^e_\eta)$ (see (9)) is the factor characterizing the asymptotic covariance matrix, at the bivariate power-exponential distribution with tail parameter η, of the sample ρ-median based on $\rho(t) = t^p$. Dotted lines are estimates of $v_p(P^e_\eta)$ computed from $M = 10{,}000$ random samples of size $n = 200$; see Section 8 for details.

each sample $m = 1, \ldots, M$. In Figure 4, we report the quantities

$$\hat{v}_p := n\bigg(\frac{1}{M}\sum_{m=1}^{M}\big(\hat{\mu}^\rho_{0,u}(m) - \bar{\mu}^\rho_{0,u}\big)\big(\hat{\mu}^\rho_{0,u}(m) - \bar{\mu}^\rho_{0,u}\big)'\bigg)_{11},$$

with

$$\bar{\mu}^\rho_{0,u} := \frac{1}{M}\sum_{m=1}^{M}\hat{\mu}^\rho_{0,u}(m).$$

These quantities estimate the upper-left entry in the corresponding asymptotic covariance matrix V, namely the corresponding factor $v_p(P)$ in (7). Clearly, the results are in perfect agreement with Theorem 8.2. It is only for $p = 4$ and t-distributions that some deviation from the asymptotic theory is seen, but this deviation vanishes for larger sample sizes (for $p = 4$ and t-distributions, Figure 4 also provides the results for sample size $n = 1{,}000$).

Obviously, using Theorem 8.2 to conduct inference based on ρ-quantiles (i.e., performing hypothesis testing or building confidence zones) requires estimating consistently the corresponding asymptotic covariance matrix V. A natural estimator is of course $\hat{V}_n = \hat{A}_n^{-1} \hat{B}_n \hat{A}_n^{-1}$, with

$$\hat{A}_n := \frac{1}{n} \sum_{i=1}^{n} \nabla^2 H^{\rho}_{\alpha,u}(Z_i - \hat{\mu}^{\rho}_{\alpha,u}) \mathbb{I}[\|Z_i - \hat{\mu}^{\rho}_{\alpha,u}\| \in \mathcal{D}_\rho]$$

and

$$\hat{B}_n := \frac{1}{n} \sum_{i=1}^{n} (\nabla H^{\rho}_{\alpha,u}(Z_i - \hat{\mu}^{\rho}_{\alpha,u}))(\nabla H^{\rho}_{\alpha,u}(Z_i - \hat{\mu}^{\rho}_{\alpha,u}))' \mathbb{I}[\|Z_i - \hat{\mu}^{\rho}_{\alpha,u}\| \in \mathcal{D}_\rho].$$

One may proceed as in Haberman (1989) to establish that $\hat{V}_n$ converges in probability to V as the sample size n diverges to infinity.

9. Perspectives for future research

In this paper, we investigated the properties of the spatial ρ-quantiles in Definition 1. While this arguably settles the probabilistic study of these quantiles in the setup considered, our work naturally calls for an extension to more general setups and for applications of these quantiles. As mentioned in the introduction, the spatial quantiles from Chaudhuri (1996) are flexible objects that can cope with more exotic types of data, such as functional data. This is associated with the fact that these quantiles are defined as minimizers of an objective function (see (1)) that involves norms and inner products only, hence that also makes sense in Hilbert spaces. This, however, is also the case for the objective function defining ρ-quantiles in (2), so that it would be natural to investigate the properties of ρ-quantiles for random variables taking values in Hilbert spaces and to compare their properties with those of the classical spatial quantiles; we refer to Cardot, Cénac and Godichon-Baggioni (2017), Cardot, Cénac and Zitt (2013) and Chakraborty and Chaudhuri (2014) for results on the spatial median and spatial quantiles in infinite-dimensional spaces.

Another direction for future research is related to inferential applications. As already mentioned in the introduction, the spatial quantiles from Chaudhuri (1996) and the companion spatial depth have been much used in a quantile regression framework (Chakraborty (2003), Cheng and De Gooijer (2007), Chowdhury and Chaudhuri (2019)), and it would be of interest to consider ρ-quantiles in this setup. In particular, this would provide a spatial concept of multiple-output expectile regression, which would be quite natural since expectiles were originally introduced, in Newey and Powell (1987), as an L_2-alternative to the traditional L_1-concept of quantile regression (Koenker and Bassett (1978)). Another natural venue for application of ρ-quantiles and ρ-depth is supervised classification. In the last decade, supervised classification based on depth, where a new observation is classified into the population with respect to which it is deepest, has met much success in the literature; see, e.g., Li, Cuesta-Albertos and Liu (2012), Pokotylo, Mozharovskyi and Dyckerhoff (2019), and the references therein. In this framework, L_p-depths provide natural tools to implement this max-depth approach where p might be chosen through cross-validation. Such applications, or the application of ρ-quantiles in risk assessment, deserve a full-fledged paper, hence are left for future work.

Acknowledgement

The authors would like to thank the Editor-In-Chief, Mark Podolskij, the Associate Editor, and two anonymous referees for their insightful comments and suggestions. This research work was supported by a research fellowship from the Francqui Foundation and by the Program of Concerted Research Actions (ARC) of the Université libre de Bruxelles.

References

Azzalini, A. and Capitanio, A. (2014). *The Skew-Normal and Related Families. IMS Monograph series*. Cambridge University Press.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, Cambridge.

Breckling, J. and Chambers, R. (1988). M-Quantiles. *Biometrika* **75** 761–771.

Brown, B. M. (1983). Statistical uses of the spatial median. *J. R. Statist. Soc. Ser. B* **45** 25–30.

Cardot, H., Cénac, P. and Zitt, P.-A. (2013). Efficient and fast estimation of the geometric median in Hilbert spaces with an averaged stochastic gradient algorithm. *Bernoulli* **19** 18–43.

Cardot, H., Cénac, P. and Godichon-Baggioni, A. (2017). Online estimation of the geometric median in Hilbert spaces: Nonasymptotic confidence balls. *Ann. Statist.* **45** 591–614.

Chakraborty, B. (2003). On multivariate quantile regression. *J. Statist. Plann. Inference* **110** 109–132.

Chakraborty, A. and Chaudhuri, P. (2014). The spatial distribution in infinite dimensional spaces and related quantiles and depths. *Ann. Statist.* **42** 1203–1231.

Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *J. Amer. Statist. Assoc.* **91** 862–872.

Chen, Z. (1996). Conditional L_p-quantiles and their application to the testing of symmetry in non-parametric regression. *Statist. Probab. Lett.* **29** 107–115.

Cheng, Y. and De Gooijer, J. G. (2007). On the uth geometric conditional quantile. *J. Statist. Plann. Inference* **137** 1914–1930.

Chowdhury, J. and Chaudhuri, P. (2019). Nonparametric depth and quantile regression for functional data. *Bernoulli* **25** 395–423.

Daouia, A., Girard, S. and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *J. Roy. Statist. Soc. Ser. B* **80** 263–292.

Daouia, A., Girard, S. and Stupfler, G. (2019). Extreme M-quantiles as risk measures: from L_1 to L_p optimization. *Bernoulli* **25** 264–309.

Francisci, G., Nieto-Reyes, A. and Agostinelli, C. (2019). Generalization of the simplicial depth: no vanishment outside the convex hull of the distribution support. *arXiv preprint arXiv:1909.02739.*

Gardes, L., Girard, S. and Stupfler, G. (2020). Beyond tail median and conditional tail expectation: extreme risk estimation using tail L_p-optimization. *Scand. J. Statist.* **47** 922–949.

Ghosh, A. K. and Chaudhuri, P. (2005). On maximum depth and related classifiers. *Scand. J. Stat.* **32** 327–350.

Girard, S. and Stupfler, G. (2015). Extreme geometric quantiles in a multivariate regular variation framework. *Extremes* **18** 629–663.

Girard, S. and Stupfler, G. (2017). Intriguing properties of extreme geometric quantiles. *REVSTAT* **15** 107–139.

Haberman, S. J. (1989). Concavity and estimation. *Ann. Statist.* **17** 1631–1661.

Hallin, M., Paindaveine, D. and Šiman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: From L_1 optimization to halfspace depth (with discussion). *Ann. Statist.* **38** 635–703.

Hallin, M., del Barrio, E., Albertos, J. C. and Matran, C. (2021). Center-outward distribution/quantile functions, ranks, and signs in $\mathbb{R}^d$: a measure transportation approach. *Ann. Statist.* **49(2)** 1139–1165.

Herrmann, K., Hofert, M. and Mailhot, M. (2018). Multivariate geometric expectiles. *Scand. Actuar. J.* **2018** 629–659.

Koenker, R. and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46** 33–50.

Koltchinski, V. I. (1997). M-estimation, convexity and quantiles. *Ann. Statist.* **25** 435–477.

Kuan, C. M., Yeh, J. H. and Hsu, Y. C. (2009). Assessing value at risk with CARE, the Conditional Autoregressive Expectile models. *J. Econometrics* **150** 261–270.

Li, J., Cuesta-Albertos, J. A. and Liu, R. Y. (2012). DD-classifier: Nonparametric classification procedure based on DD-plot. *J. Amer. Statist. Assoc* **107** 737–753.

Magyar, A. and Tyler, D. E. (2011). The asymptotic efficiency of the spatial median for elliptically symmetric distributions. *Sankhyā* **73** 165–192.

Massé, J.-C. (2002). Asymptotics for the Tukey median. *J. Multivariate Anal.* **81** 286–

300.

MOSLER, K. (2012). *Multivariate Dispersion, Central Regions, and Depth: The Lift Zonoid Approach* **165**. Springer Science & Business Media.

MÖTTÖNEN, J., NORDHAUSEN, K., OJA, H. et al. (2010). Asymptotic theory of the spatial median. In *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in honor of Professor Jana Jurečková* 182–193. Institute of Mathematical Statistics.

MUKHOPADHYAYA, N. D. and CHATTERJEE, S. (2011). High dimensional data analysis using multivariate generalized spatial quantiles. *J. Multivariate Anal.* **102** 768–780.

NEWEY, W. K. and POWELL, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica* **55** 819–847.

PAINDAVEINE, D. and VIRTA, J. (2021). On the behavior of extreme spatial quantiles under minimal assumptions. In *Advances in Contemporary Statistics and Econometrics* (A. DAOUIA and A. RUIZ-GAZEN, eds.) 243–259. Springer, Cham.

POKOTYLO, O., MOZHAROVSKYI, P. and DYCKERHOFF, R. (2019). Depth and depth-based classification with R-package ddalpha. *J. Statist. Softw.* **91** 1–46.

SERFLING, R. (2002). Quantile functions for multivariate analysis: approaches and applications. *Stat. Neerl.* **56** 214–232.

SERFLING, R. (2010). Equivariance and invariance properties of multivariate quantile and related functions, and the role of standardisation. *J. Nonparametr. Stat.* **22** 915–936.

SERFLING, R. (2019). Depth functions on general data spaces, I. Perspectives, with consideration of "density" and "local" depths. Submitted.

SERFLING, R. and WIJESURIYA, U. (2017). Depth-based nonparametric description of functional data, with emphasis on use of spatial depth. *Comput. Statist. Data Anal.* **105** 24–45.

SERFLING, R. and ZUO, Y. (2010). Discussion of "Multivariate quantiles and multiple-output regression quantiles: From L_1 optimization to halfspace depth", by M. Hallin, D. Paindaveine, and M. Šiman. *Ann. Statist.* **38** 676–684.

TAYLOR, J. (2008). Estimating value at risk and expected shortfall using expectiles. *J. Financ. Econometrics* **6** 231–252.

USSEGLIO-CARLEVE, A. (2018). Estimation of conditional extreme risk measures from heavy-tailed elliptical random vectors. *Electron. J. Stat.* **12** 4057–4093.

VARDI, Y. and ZHANG, C.-H. (2000). The multivariate L_1-median and associated data depth. *Proc. Natl. Acad. Sci. USA* **97** 1423–1426.

ZHOU, W. and SERFLING, R. (2008). Multivariate spatial U-quantiles: a Bahadur-Kiefer representation, a Theil–Sen estimator for multiple regression, and a robust dispersion estimator. *J. Statist. Plann. Inference* **138** 1660–1678.