REVSTAT – Statistical Journal
Volume 15, Number 1, January 2017, 107–139

INTRIGUING PROPERTIES OF EXTREME GEOMETRIC QUANTILES

Authors: Stéphane Girard
– Team Mistis, Inria Grenoble Rhône-Alpes & LJK, France
Stephane.Girard@inria.fr

Gilles Stupfler
– GREQAM UMR 7316, CNRS, EHESS, Centrale Marseille, Aix Marseille Université, France
gilles.stupfler@univ-amu.fr

Received: January 2015 Revised: May 2015 Accepted: May 2015

Abstract:

- Central properties of geometric quantiles have been well-established in the recent statistical literature. In this study, we try to get a grasp of how extreme geometric quantiles behave. Their asymptotics are provided, both in direction and magnitude, under suitable moment conditions, when the norm of the associated index vector tends to one. Some intriguing properties are highlighted: in particular, it appears that if a random vector has a finite covariance matrix, then the magnitude of its extreme geometric quantiles grows at a fixed rate. We take profit of these results by defining a parametric estimator of extreme geometric quantiles of such a random vector. The consistency and asymptotic normality of the estimator are established, and contrasted with what can be obtained for univariate quantiles. Our results are illustrated on both simulated and real data sets. As a conclusion, we deduce from our observations some warnings which we believe should be known by practitioners who would like to use such a notion of multivariate quantile to detect outliers or analyze extremes of a random vector.

Key-Words:

- *extreme quantile; geometric quantile; consistency; asymptotic normality.*

AMS Subject Classification:

- 62H05, 62G20, 62G32.

1. INTRODUCTION

Let X be a random vector in $\mathbb{R}^d$. Up to now, several definitions of multivariate quantiles of X have been proposed in the statistical literature. We refer to [25] for a review of various possibilities for this notion. Here, we focus on the notion of "spatial" or "geometric" quantiles, introduced by [14], which generalises the characterisation of a univariate quantile shown in [22]. For a given vector u belonging to the unit open ball B^d of $\mathbb{R}^d$, where $d \geq 2$, a geometric quantile with index vector u is any solution of the optimisation problem defined by

$$\arg\min_{q\in\mathbb{R}^d} \mathbb{E}\big(\|X-q\| - \|X\|\big) - \langle u, q\rangle \ , \tag{1.1}$$

where $\langle\cdot,\cdot\rangle$ is the usual scalar product on $\mathbb{R}^d$ and $\|\cdot\|$ is the associated Euclidean norm. Note that $q(u) \in \mathbb{R}^d$ possesses both a direction and magnitude. It can be seen that geometric quantiles are in fact special cases of M-quantiles introduced by [3] which were further analysed by [23]. Besides, such quantiles have various strong properties. First, the quantile with index vector $u \in B^d$ is unique whenever the distribution of X is not concentrated on a single straight line in $\mathbb{R}^d$ (see [14] or Theorem 2.17 in [21]). Second, although they are not fully affine equivariant, they are equivariant under any orthogonal transformation [14]. Third, geometric quantiles characterise the associated distribution. Namely, if two random variables X and Y yield the same quantile function q, then X and Y have the same distribution [23]. Finally, for $u = 0$, the well-known L^2-geometric median is obtained, which is the simplest example of a "central" quantile [28]. We point out that one may compute an estimation of the geometric median in an efficient way, see [8].

These properties make geometric quantiles reasonable candidates when trying to define multivariate quantiles, which is why their estimation was studied in several papers. We refer for instance to [14], who established a Bahadur expansion for the estimator of geometric quantiles obtained by solving the sample counterpart of problem (1.1). [10] then introduced a transformation–retransformation procedure to obtain affine equivariant estimates of multivariate quantiles. This notion was extended to a multiresponse linear model by [11]. Recently, [16] defined a multivariate quantile–quantile plot using geometric quantiles. Conditional geometric quantiles can also be defined by substituting a conditional expectation to the expectation in (1.1). We refer to [6] for the estimation of the conditional geometric median and to [15] for the estimation of an arbitrary conditional geometric quantile. The estimation of a conditional median when there is an infinite-dimensional covariate is considered in [13].

Let us note though that the previous papers focus on central properties of geometric quantiles and of their sample versions. While some of them label geometric quantiles as "extreme" when $\|u\|$ is close to 1 ([14, 15]) and use it in real applications (see *e.g.* [12] for an application to outlier detection), the specific properties of these extreme geometric quantiles have not been investigated yet. In this study, we provide the asymptotics of the direction and magnitude of the extreme geometric quantile $q(u)$ when $\|u\| \to 1$, under suitable moment conditions. There are well-known analogue results for univariate extreme quantiles in the right tail of a distribution, see *e.g.* [18]. A particular corollary of our results is that the magnitude of the extreme geometric quantiles of a random vector X having a finite covariance matrix grows at a fixed rate. Moreover, in this case, the magnitude of the extreme geometric quantiles is asymptotically characterised by the covariance matrix of X. This is an intriguing property, which opens the door to a parametric estimation of extreme quantiles whose asymptotic properties are studied in this work.

The outline of the paper is as follows. Asymptotic properties of geometric quantiles are stated in Section 2. An illustrative application to the estimation of extreme geometric quantiles is given in Section 3. Some examples and numerical illustrations of our results, including a study of a real data set, are presented in Section 4. Section 5 offers a couple of concluding remarks, in which some warnings are given to practitioners who would like to use such geometric quantiles to detect outliers or analyze extremes of a random vector. Proofs are deferred to Section 6.

2. ASYMPTOTIC BEHAVIOUR OF EXTREME GEOMETRIC QUANTILES

From now on, we assume that the distribution of X is not concentrated on a single straight line in $\mathbb{R}^d$ and non-atomic. [14] proved that, in this context, the solution $q(u)$ of (1.1), namely the geometric quantile with index vector u, exists and is unique for every $u \in B^d$. Let $\psi\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be defined as $\psi(u,q) = \mathbb{E}(\|X-q\| - \|X\|) - \langle u, q\rangle$ and assume further that $t/\|t\| = 0$ if $t = 0$. If $u \in \mathbb{R}^d$ is such that there is a solution $q(u) \in \mathbb{R}^d$ to problem (1.1), then the gradient of $q \mapsto \psi(u,q)$ must be zero at $q(u)$, that is

$$u + \mathbb{E}\left(\frac{X - q(u)}{\|X - q(u)\|}\right) = 0 \,. \tag{2.1}$$

This condition immediately entails that if $u \in \mathbb{R}^d$ is such that problem (1.1) has a solution $q(u)$, then $\|u\| \le 1$. In fact, we can prove a stronger result:

Proposition 2.1. *The optimisation problem (1.1) has a solution if and only if $u \in B^d$.*

Moreover, remarking that the function $\psi(u,\cdot)$ is strictly convex, [14] proved the following characterisation of a geometric quantile: for every $u \in B^d$, $q(u)$ is the solution of problem (1.1) if and only if it satisfies equation (2.1). In particular, this entails that the function $G\colon \mathbb{R}^d \to B^d$ defined by

$$\forall q \in \mathbb{R}^d\,, \qquad G(q) = -\mathbb{E}\left(\frac{X-q}{\|X-q\|}\right)$$

is a continuous bijection. Proposition 2.6(iii) in [23] shows that the inverse of the function G, *i.e.* the geometric quantile function $u \mapsto q(u)$, is also continuous on B^d.

In most cases however, computing explicitly the function G is a hopeless task, which makes it impossible to obtain a closed-form expression for the geometric quantile function. It is thus of interest to prove general results about the geometric quantile $q(u)$, especially regarding its direction and magnitude. Our first main result focuses on the special case of spherically symmetric distributions.

Proposition 2.2. *If X has a spherically symmetric distribution then:*

- **(i)** *The map $u \mapsto q(u)$ commutes with every linear isometry of $\mathbb{R}^d$. Especially, the norm of a geometric quantile $q(u)$ only depends on the norm of u.*
- **(ii)** *For all $u \in B^d$, the geometric quantile $q(u)$ has direction u if $u \neq 0$ and $q(0) = 0$ otherwise.*
- **(iii)** *The function $\|u\| \mapsto \|q(u)\|$ is a continuous strictly increasing function on $[0,1)$.*
- **(iv)** *It holds that $\|q(u)\| \to \infty$ as $\|u\| \to 1$.*

Although Proposition 2.2(i,iii) cannot be expected to hold true for a random variable which is not spherically symmetric, one may wonder if (ii,iv), namely that a geometric quantile shares the direction of its index vector and that the norm of the geometric quantile function tends to infinity on the unit sphere, can be extended to the general case. The next result, which examines the behaviour of the geometric quantile function near the boundary of the open ball B^d, provides an answer to this question.

Theorem 2.1. *Let S^{d-1} be the unit sphere of $\mathbb{R}^d$.*

- **(i)** *It holds that $\|q(v)\| \to \infty$ as $\|v\| \to 1$.*
- **(ii)** *Moreover, if $v \to u$ with $u \in S^{d-1}$ and $v \in B^d$ then $q(v)/\|q(v)\| \to u$.*

Theorem 2.1 shows two properties of geometric quantiles: first, the norm of the geometric quantile $q(v)$ with index vector v diverges to infinity as $\|v\| \uparrow 1$. In other words, Proposition 2.2(iv) still holds for any distribution. This is a rather intriguing property of geometric quantiles, since it holds even if the distribution of X has a compact support (for instance, when X is uniformly distributed on a square). A related point is the fact that sample geometric quantiles do not necessarily lie within the convex hull of the sample, see [4] for a counter-example. Second, if $v \to u \in S^{d-1}$, then the geometric quantile $q(v)$ has asymptotic direction u. Proposition 2.2(ii) thus remains true asymptotically for any distribution.

It is possible to specify the convergences obtained in Theorem 2.1 under moment assumptions. Theorem 2.2 provides a first-order expansion of both the direction and the magnitude of an extreme geometric quantile $q(\alpha u)$ in the direction u, where u is a unit vector and α tends to 1.

Theorem 2.2. *Let $u \in S^{d-1}$.*

(i) *If $\mathbb{E}\|X\| < \infty$ then $q(\alpha u) - \big\{\|q(\alpha u)\| u + \mathbb{E}(X - \langle X, u\rangle u)\big\} \to 0$ as $\alpha \uparrow 1$.*

(ii) *If $\mathbb{E}\|X\|^2 < \infty$ and Σ denotes the covariance matrix of X then*

$$\|q(\alpha u)\|^2 (1-\alpha) \;\to\; \frac{1}{2}\big(\operatorname{tr}\Sigma - u'\Sigma u\big) > 0 \qquad \text{as} \quad \alpha \uparrow 1\,.$$

Let us note that the integrability conditions of Theorem 2.2 exclude any random vector $\|X\|$ whose distribution possesses a right tail which is too heavy. For instance, condition $\mathbb{E}\|X\| < \infty$ in (i) excludes the multivariate Student distribution with less than one degree of freedom, while condition $\mathbb{E}\|X\|^2 < \infty$ in (ii) excludes the multivariate Student distribution with less than two degrees of freedom.

Consequence 1. It appears that, if X has a finite covariance matrix Σ, then the magnitude of an extreme geometric quantile is determined (in the asymptotic sense) by Σ. In other words, since the asymptotic direction of an extreme geometric quantile in the direction u is exactly u by Theorem 2.1, it follows that the extreme geometric quantiles of two probability distributions which admit the same finite covariance matrix are asymptotically equivalent. This phenomenon is illustrated on simulated data in Section 4 below. This is surprising from the extreme value perspective: one could expect the behaviour of extreme geometric quantiles not to be driven by a central parameter such as the covariance matrix, as happens in the univariate context where the value of an extreme quantile depends on the tail heaviness of the probability density function of X.

Consequence 2. The map $\lambda \mapsto \|q((1-\lambda^{-1})u)\|$ is a regularly varying function with index 1/2 (see Bingham *et al.*, 1987) and therefore:

$$\frac{\|q(\beta u)\|}{\|q(\alpha u)\|} = \left(\frac{1-\alpha}{1-\beta}\right)^{1/2} \big(1 + \mathrm{o}(1)\big)$$

when $\alpha \to 1$ and $\beta \to 1$. In other words, given an arbitrary extreme geometric quantile, one can deduce the asymptotic behaviour of every other extreme geometric quantile sharing its direction, independently of the distribution.

Again, this is fundamentally different from the univariate case when deducing the value of an extreme quantile from another one then requires the knowledge (or an estimate) of the extreme-value index of the distribution, see [18], Chapter 4. A further, perhaps unexpected, consequence is that our results can actually be used to define a consistent and asymptotically Gaussian estimator of extreme geometric quantiles by using the standard empirical estimator of the covariance matrix of X, see Section 3 below.

Consequence 3. Finally, Theorem 2.2 provides some information on the shape of an extreme quantile contour. It is readily seen that the global maximum of the function $h_1(u) := \operatorname{tr}\Sigma - u'\Sigma u$ on S^{d-1} is reached at a unit eigenvector $u_{\min}$ of Σ associated with its smallest eigenvalue $\lambda_{\min} > 0$. Thus, the norm of an extreme geometric quantile is asymptotically the largest in the direction where the variance is the smallest. Similarly, the global minimum of h_1 is reached at a unit eigenvector $u_{\max}$ of Σ associated with its largest eigenvalue $\lambda_{\max} > 0$. In particular, if f is the probability density function associated with an elliptically contoured distribution [7], the level sets of f coincide with the level sets of the function $h_2(u) := u'\Sigma u$. The global maximum of h_2 is reached at the eigenvector $u_{\max}$ while the global minimum is reached at $u_{\min}$. The extreme geometric quantile is therefore furthest from the origin in the direction where the density level set is closest to the origin, see Section 4 for an illustration on real data. In such a case, the extreme geometric quantile contour plot and the density level plots are in some sense orthogonal (even though they agree when the distribution of X is spherically symmetric). Of course, one should not expect a direct geometric match between quantile contours and density contours, but this phenomenon should be kept in mind when designing outlier detection procedures. In our view, this can be seen as a consequence of the lack of affine-equivariance of geometric quantiles. To tackle this issue, one may apply a transformation–retransformation procedure, see [27]. Such procedures admit sample analogues, see for instance [9, 10], at the possible loss of geometric interpretation, see [26].

3. AN ESTIMATOR OF EXTREME GEOMETRIC QUANTILES

In this paragraph, our focus is to illustrate Consequence 2 of Theorem 2.2 at the sample level. Let $X_1, ..., X_n$ be independent random copies of a random vector X having a finite covariance matrix Σ. It follows from Theorem 2.2 that any extreme geometric quantile $q(\alpha u)$ of X, with $\alpha \uparrow 1$ and $u \in S^{d-1}$ can be approximated by:

$$q_{\mathrm{eq}}(\alpha u) := (1-\alpha)^{-1/2} \left[\frac{1}{2}\left(\operatorname{tr}\Sigma - u'\Sigma u\right)\right]^{1/2} u \,. \tag{3.1}$$

This can be used to define an estimator of the extreme geometric quantiles of X: let $\overline{X}_n = n^{-1}\sum_{k=1}^n X_k$ be the sample mean and

$$\widehat{\Sigma}_n = \frac{1}{n}\sum_{k=1}^n (X_k - \overline{X}_n)\,(X_k - \overline{X}_n)'$$

be the empirical estimator of the covariance matrix Σ of X. Let further (α_n) be an increasing sequence of positive real numbers tending to 1. Our estimator $\widehat{q}_n(\alpha_n u)$ of $q(\alpha_n u)$ is then

$$\widehat{q}_n(\alpha_n u) = (1-\alpha_n)^{-1/2}\left[\frac{1}{2}\left(\operatorname{tr}\widehat{\Sigma}_n - u'\widehat{\Sigma}_n u\right)\right]^{1/2} u\,.$$

The consistency of $\widehat{q}_n(\alpha_n u)$ is examined in the next result.

Theorem 3.1. *Let $u \in S^{d-1}$ and assume that $\alpha_n \uparrow 1$. If $\mathbb{E}\|X\|^2 < \infty$ then*

$$\sqrt{1-\alpha_n}\,\big(\widehat{q}_n(\alpha_n u) - q(\alpha_n u)\big) \to 0 \qquad \text{almost surely as} \quad n \to \infty\,.$$

This result actually means that the extreme geometric quantile estimator is relatively consistent in the sense that

$$\frac{\widehat{q}_n(\alpha_n u) - q(\alpha_n u)}{\|q(\alpha_n u)\|} \to 0 \qquad \text{almost surely as} \quad n \to \infty\,,$$

since $\|q(\alpha_n u)\|^{-1} = \mathrm{O}(\sqrt{1-\alpha_n})$, see Theorem 2.2(ii). This normalisation could be expected since the quantity to be estimated diverges in magnitude. Under the additional assumption that X has a finite fourth moment, an asymptotic normality result can be established for this estimator:

Theorem 3.2. *Let $u \in S^{d-1}$ and assume that $\alpha_n \uparrow 1$ is such that $n(1-\alpha_n) \to 0$. If $\mathbb{E}\|X\|^4 < \infty$ then*

$$\sqrt{n(1-\alpha_n)}\,\big(\widehat{q}_n(\alpha_n u) - q(\alpha_n u)\big) \xrightarrow{d} Z \qquad \text{as} \quad n \to \infty$$

where Z is a Gaussian centred random vector.

Let us highlight that the covariance matrix of the Gaussian limit in Theorem 3.2 essentially depends on the covariance matrix M of the Gaussian limit of $\sqrt{n}(\widehat{\Sigma}_n - \Sigma)$, see the proof in Section 6. Although M has a complicated expression (see *e.g.* [24]), it can be estimated when $\mathbb{E}\|X\|^4 < \infty$, which makes it possible to construct asymptotic confidence regions for extreme geometric quantiles.

Extreme geometric quantiles can thus be consistently estimated by $\widehat{q}_n(\alpha_n u)$, whatever the "order" α_n, and an asymptotic normality result is obtained when

$\alpha_n \uparrow 1$ *quickly enough.* The proposed estimator is therefore able to extrapolate arbitrarily far from the original sample. This is very different from the univariate case, where the empirical quantile $\widehat{q}_n(\alpha_n) = \inf\{t \in \mathbb{R} \mid \widehat{F}_n(t) \geq \alpha_n\}$, deduced from the empirical cumulative distribution function $\widehat{F}_n$, estimates the true quantile $q(\alpha_n)$ consistently only if α_n converges to 1 *slowly enough.* The extrapolation with faster rates α_n is then handled assuming that the underlying distribution function is heavy-tailed and by using adapted estimators, see *e.g.* [29] and the monograph [18].

4. NUMERICAL ILLUSTRATIONS

4.1. Simulation study

In this section, our main results are illustrated, particularly Theorems 2.2, 3.1 and 3.2 in the bivariate case $d = 2$ to make the display easier. In this framework, $u \in S^1$ can be represented by an angle: $u = u_\theta = (\cos\theta, \sin\theta)$, $\theta \in [0, 2\pi)$. The iso-quantile curves $\mathcal{C}q(\alpha) = \{q(\alpha u_\theta),\ \theta \in [0, 2\pi)\}$ and their estimates $\mathcal{C}\widehat{q}_n(\alpha) = \{\widehat{q}_n(\alpha u_\theta),\ \theta \in [0, 2\pi)\}$ can then be considered in order to get a grasp of the behaviour of extreme quantiles in every direction. The following two distributions are considered for the random vector X:

- The centred Gaussian multivariate distribution $\mathcal{N}(0, v_X, v_Y, v_{XY})$, with probability density function: $\forall x, y \in \mathbb{R}$,

$$f(x,y) = \frac{1}{2\pi\sqrt{\det\Sigma}} \exp\left(-\frac{1}{2}\begin{pmatrix} x \\ y \end{pmatrix}' \Sigma^{-1} \begin{pmatrix} x \\ y \end{pmatrix}\right) \quad \text{with } \Sigma = \begin{pmatrix} v_X & v_{XY} \\ v_{XY} & v_Y \end{pmatrix}.$$

- A double exponential distribution $\mathcal{E}(\lambda_-, \mu_-, \lambda_+, \mu_+)$, with $\lambda_-, \mu_-, \lambda_+, \mu_+ > 0$, whose probability density function is: $\forall x, y \in \mathbb{R}$,

$$f(x,y) = \frac{1}{4}\begin{cases} \lambda_+ \mu_+ e^{-\lambda_+|x| - \mu_+|y|} & \text{if } xy > 0\,, \\ \lambda_- \mu_- e^{-\lambda_-|x| - \mu_-|y|} & \text{if } xy \leq 0\,. \end{cases}$$

  In this case, X is centred and has covariance matrix

$$\Sigma = \begin{pmatrix} \dfrac{1}{\lambda_-^2} + \dfrac{1}{\lambda_+^2} & \dfrac{1}{2}\left[\dfrac{1}{\lambda_+\mu_+} - \dfrac{1}{\lambda_-\mu_-}\right] \\ \dfrac{1}{2}\left[\dfrac{1}{\lambda_+\mu_+} - \dfrac{1}{\lambda_-\mu_-}\right] & \dfrac{1}{\mu_-^2} + \dfrac{1}{\mu_+^2} \end{pmatrix}.$$

Three different sets of parameters were used for each distribution, in order that the related covariance matrices coincide:

- $\mathcal{N}(0,1/2,1/2,0)$ and $\mathcal{E}(2,2,2,2)$ with spherical covariance matrices;
- $\mathcal{N}(0,1/8,3/4,0)$ and $\mathcal{E}(4,2\sqrt{2/3},4,2\sqrt{2/3})$ with diagonal covariance matrices;
- $\mathcal{N}(0,1/2,1/2,1/6)$ and $\mathcal{E}(2\sqrt{3},2\sqrt{3},2\sqrt{3/5},2\sqrt{3/5})$ with full covariance matrices.

In each case, we carry out the following computations:

- For each $\alpha \in \{0.99, 0.995, 0.999\}$, the true quantile curves $\mathcal{C}q(\alpha)$ obtained by solving problem (1.1) numerically, as well as their analogues $\mathcal{C}q_{\text{eq}}(\alpha)$ using approximation (3.1) are computed. The normalised squared approximation error

$$e(\alpha) \;=\; (1-\alpha)\int_0^{2\pi} \|q_{\text{eq}}(\alpha u_\theta) - q(\alpha u_\theta)\|^2 \, d\theta$$

  is then recorded.
- For each value of α, we draw $N = 1000$ replications of an n-sample $(X_1, ..., X_n)$ of independent copies of X, with $n \in \{100, 200, 500\}$. The estimated quantile curves $\mathcal{C}\widehat{q}_n^{(j)}(\alpha)$ corresponding to the j-th replication and the associated normalised squared error

$$E_n^{(j)}(\alpha) \;=\; (1-\alpha)\int_0^{2\pi} \left\|\widehat{q}_n^{(j)}(\alpha u_\theta) - q(\alpha u_\theta)\right\|^2 d\theta$$

  are computed as well as the mean squared error $E_n(\alpha) = N^{-1}\sum_{j=1}^N E_n^{(j)}(\alpha)$.

The true quantile curves, as well as the approximated and the estimated ones are displayed on Figures 1–6 in the case $n = 200$ and $\alpha = 0.995$. The true quantile curves look very similar in Figures 1 and 4, in Figures 2 and 5 and Figures 3 and 6 (in which the words "best", "median" and "worst" are to be understood with respect to the L^2 error). This is in accordance with Theorem 2.2: eventually, extreme geometric quantiles only depend on the covariance matrix of the underlying distribution. Moreover, the approximated quantile curves are close to the true ones in all cases, and the estimated quantile curves are satisfying in all situations with a moderate variability. Similar results were observed for $n = 100, 500$ and $\alpha = 0.99, 0.999$. We do not report the graphs here for the sake of brevity; we do however display the approximation and estimation errors in Table 1. Unsurprisingly, the estimation error $E_n(\alpha)$ decreases as the sample size n increases. Both approximation and estimation errors $e(\alpha)$ and $E_n(\alpha)$ have a stable behaviour with respect to α.

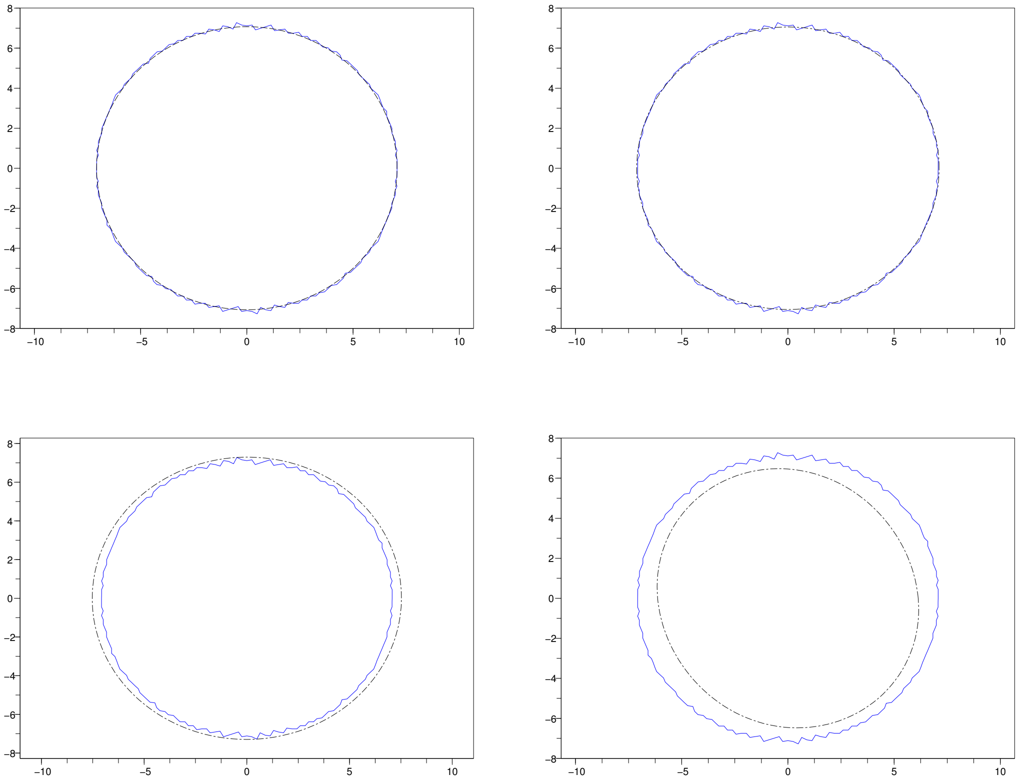

Figure 1: Spherical Gaussian distribution $\mathcal{N}(0, 1/2, 1/2, 0)$ for $\alpha = 0.995$. Top left: comparison between a numerical method and the use of the equivalent (3.1) for the computation of the iso-quantile curve, full line: numerical method, dashed line: asymptotic equivalent. Top right, bottom left and bottom right: best, median and worst estimates of the iso-quantile curve for $n = 200$, full line: numerical method, dashed-dotted line: estimator $\widehat{q}_n$.

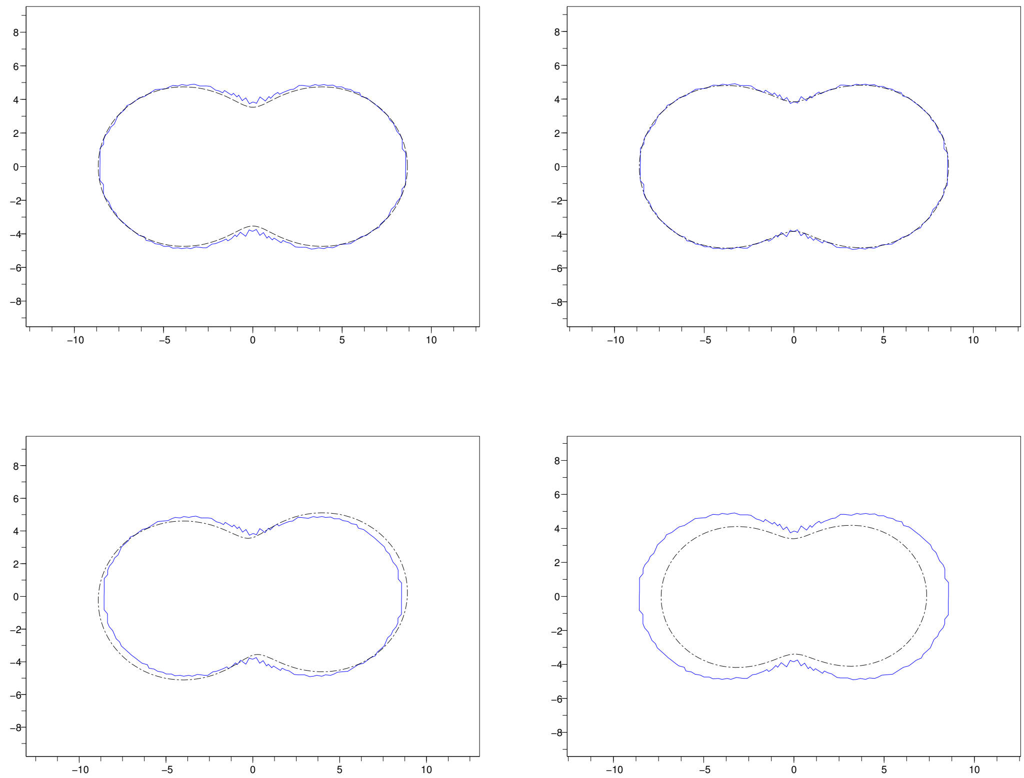

Figure 2: Diagonal Gaussian distribution $\mathcal{N}(0, 1/8, 3/4, 0)$ for $\alpha = 0.995$. Top left: comparison between a numerical method and the use of the equivalent (3.1) for the computation of the iso-quantile curve, full line: numerical method, dashed line: asymptotic equivalent. Top right, bottom left and bottom right: best, median and worst estimates of the iso-quantile curve for $n = 200$, full line: numerical method, dashed-dotted line: estimator $\widehat{q}_n$.

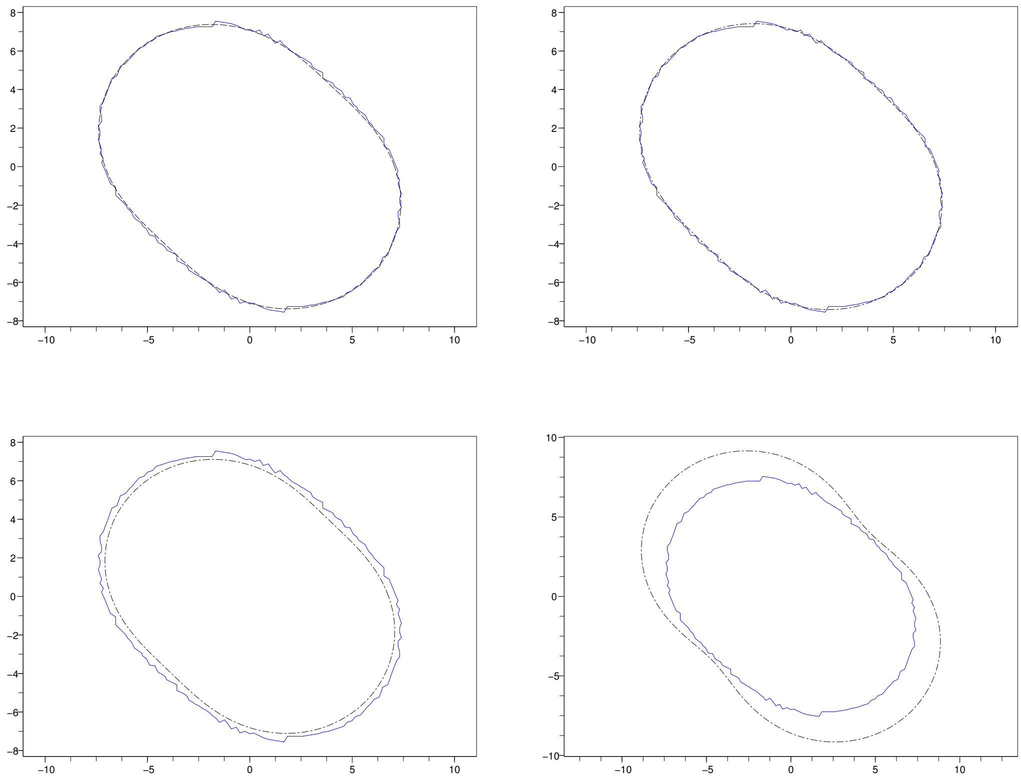

Figure 3: Full Gaussian distribution $\mathcal{N}(0, 1/2, 1/2, 1/6)$ for $\alpha = 0.995$. Top left: comparison between a numerical method and the use of the equivalent (3.1) for the computation of the iso-quantile curve, full line: numerical method, dashed line: asymptotic equivalent. Top right, bottom left and bottom right: best, median and worst estimates of the iso-quantile curve for $n = 200$, full line: numerical method, dashed-dotted line: estimator $\widehat{q}_n$.

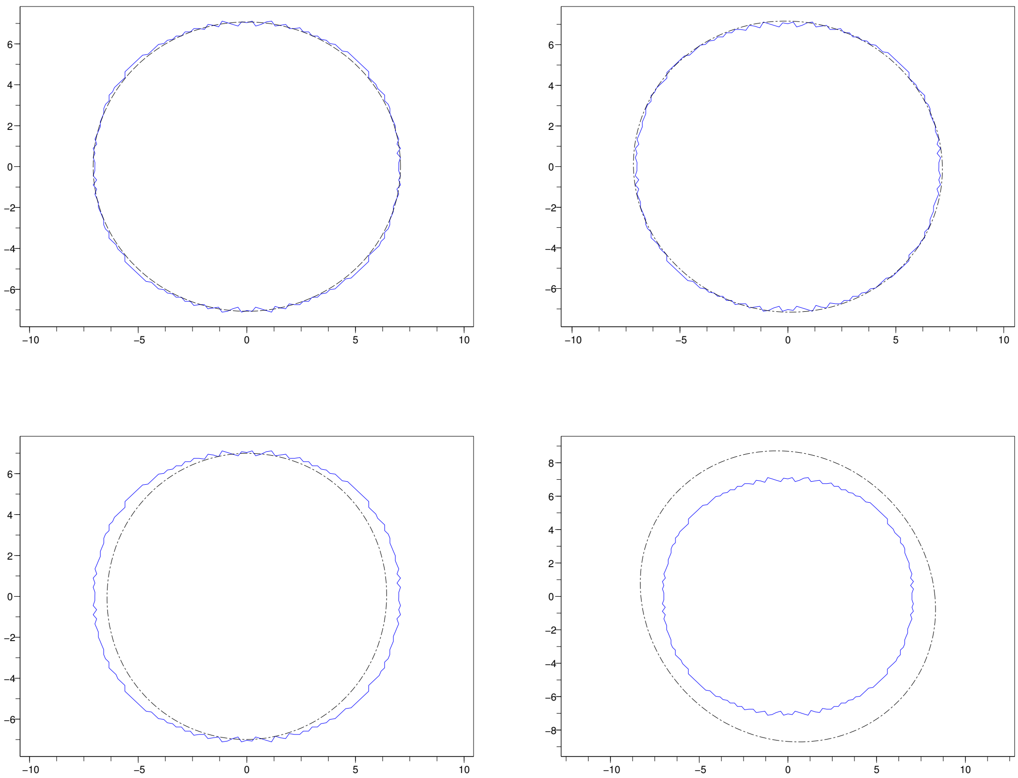

Figure 4: Spherical double exponential distribution $\mathcal{E}(2,2,2,2)$ for $\alpha = 0.995$. Top left: comparison between a numerical method and the use of the equivalent (3.1) for the computation of the iso-quantile curve, full line: numerical method, dashed line: asymptotic equivalent. Top right, bottom left and bottom right: best, median and worst estimates of the iso-quantile curve for $n = 200$, full line: numerical method, dashed-dotted line: estimator $\widehat{q}_n$.

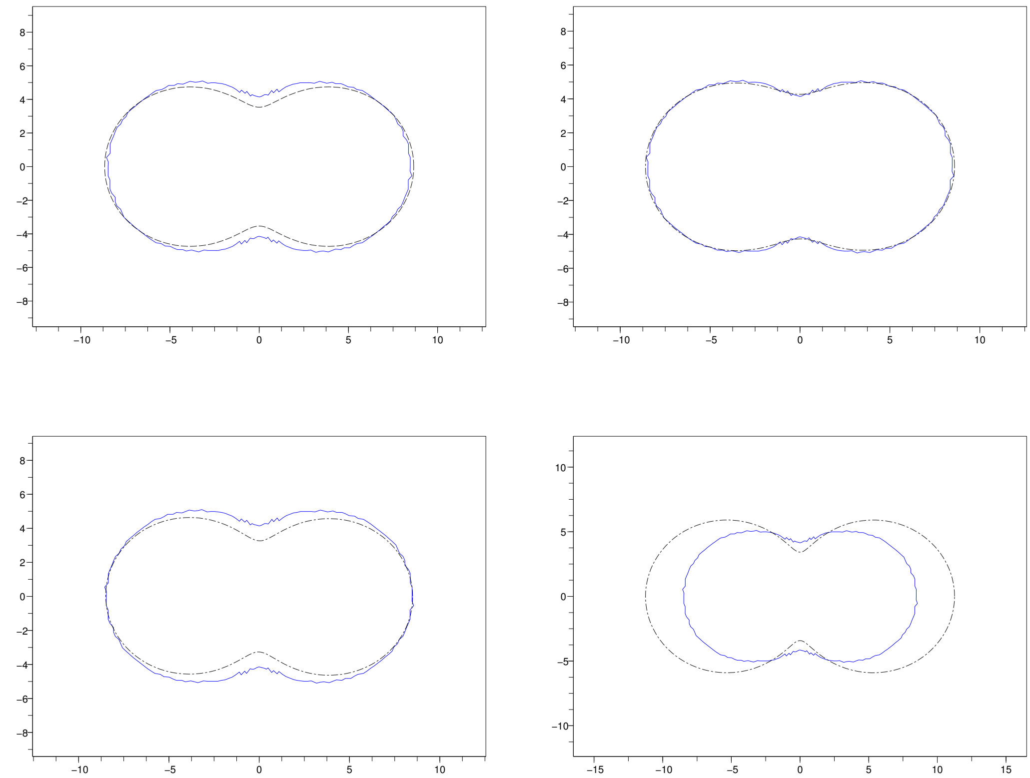

Figure 5: Diagonal double exponential distribution $\mathcal{E}(4, 2\sqrt{2/3}, 4, 2\sqrt{2/3})$ for $\alpha = 0.995$. Top left: comparison between a numerical method and the use of the equivalent (3.1) for the computation of the iso-quantile curve, full line: numerical method, dashed line: asymptotic equivalent. Top right, bottom left and bottom right: best, median and worst estimates of the iso-quantile curve for $n = 200$, full line: numerical method, dashed-dotted line: estimator $\widehat{q}_n$.

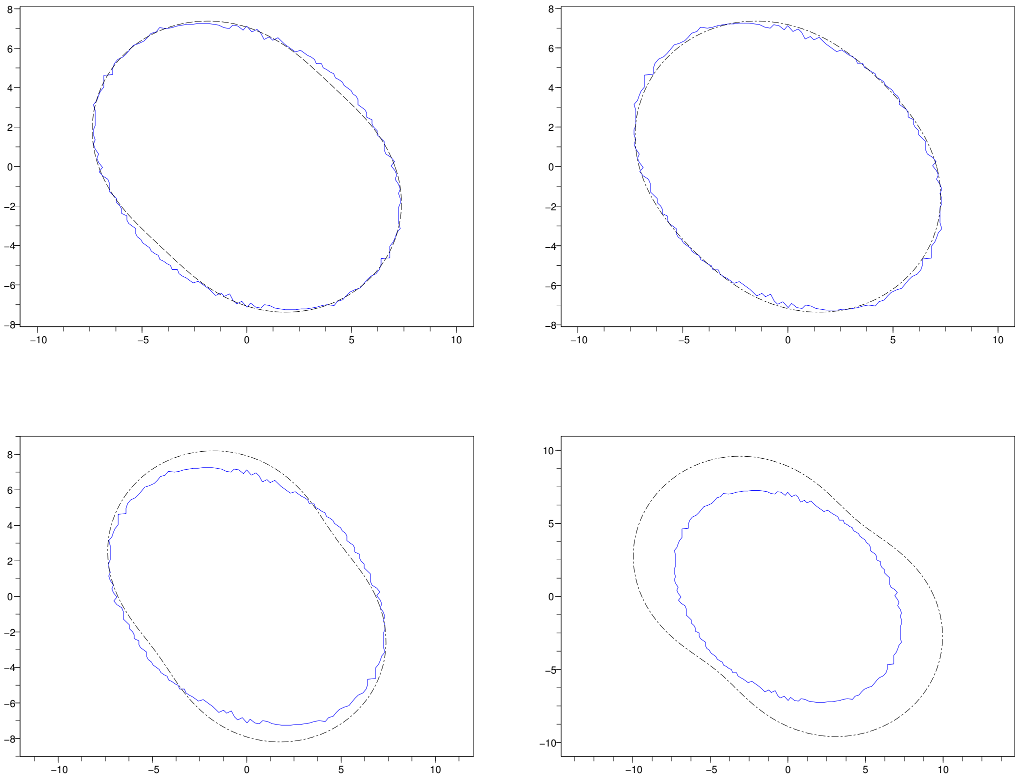

Figure 6: Full double exponential distribution $\mathcal{E}(2\sqrt{3}, 2\sqrt{3}, 2\sqrt{3/5}, 2\sqrt{3/5})$ for $\alpha = 0.995$. Top left: comparison between a numerical method and the use of the equivalent (3.1) for the computation of the iso-quantile curve, full line: numerical method, dashed line: asymptotic equivalent. Top right, bottom left and bottom right: best, median and worst estimates of the iso-quantile curve for $n = 200$, full line: numerical method, dashed-dotted line: estimator $\widehat{q}_n$.

Table 1: Errors $e(\alpha)$ and $E_n(\alpha)$ in all cases.

Distribution	Value of α	Error $e(\alpha)$	Error $E_n(\alpha)$		
			$n=100$	$n=200$	$n=500$
Centred Gaussian $\mathcal{N}(0,1/2,1/2,0)$	0.990	$2.55\cdot10^{-5}$	$1.29\cdot10^{-3}$	$6.50\cdot10^{-4}$	$2.93\cdot10^{-4}$
	0.995	$2.43\cdot10^{-5}$	$1.28\cdot10^{-3}$	$6.44\cdot10^{-4}$	$2.88\cdot10^{-4}$
	0.999	$5.75\cdot10^{-5}$	$1.30\cdot10^{-3}$	$6.70\cdot10^{-4}$	$3.16\cdot10^{-4}$
Centred Gaussian $\mathcal{N}(0,1/2,1/2,1/6)$	0.990	$1.05\cdot10^{-4}$	$1.45\cdot10^{-3}$	$7.32\cdot10^{-4}$	$3.57\cdot10^{-4}$
	0.995	$4.34\cdot10^{-5}$	$1.37\cdot10^{-3}$	$6.65\cdot10^{-4}$	$2.89\cdot10^{-4}$
	0.999	$6.34\cdot10^{-5}$	$1.38\cdot10^{-3}$	$6.83\cdot10^{-4}$	$3.05\cdot10^{-4}$
Centred Gaussian $\mathcal{N}(0,1/8,3/4,0)$	0.990	$6.05\cdot10^{-4}$	$1.79\cdot10^{-3}$	$1.17\cdot10^{-3}$	$8.23\cdot10^{-4}$
	0.995	$1.77\cdot10^{-4}$	$1.34\cdot10^{-3}$	$7.31\cdot10^{-4}$	$3.91\cdot10^{-4}$
	0.999	$5.96\cdot10^{-5}$	$1.20\cdot10^{-3}$	$6.02\cdot10^{-4}$	$2.70\cdot10^{-4}$
Double exponential $\mathcal{E}(2,2,2,2)$	0.990	$9.30\cdot10^{-5}$	$2.69\cdot10^{-3}$	$1.47\cdot10^{-3}$	$6.37\cdot10^{-4}$
	0.995	$5.46\cdot10^{-5}$	$2.63\cdot10^{-3}$	$1.41\cdot10^{-3}$	$5.93\cdot10^{-4}$
	0.999	$6.32\cdot10^{-5}$	$2.63\cdot10^{-3}$	$1.39\cdot10^{-3}$	$5.97\cdot10^{-4}$
Double exponential $\mathcal{E}(2\sqrt{3},2\sqrt{3},2\sqrt{3/5},2\sqrt{3/5})$	0.990	$6.17\cdot10^{-4}$	$4.37\cdot10^{-3}$	$2.71\cdot10^{-3}$	$1.42\cdot10^{-3}$
	0.995	$2.24\cdot10^{-4}$	$3.89\cdot10^{-3}$	$2.26\cdot10^{-3}$	$9.96\cdot10^{-4}$
	0.999	$2.27\cdot10^{-4}$	$3.77\cdot10^{-3}$	$2.16\cdot10^{-3}$	$9.62\cdot10^{-4}$
Double exponential $\mathcal{E}(4,2\sqrt{2/3},4,2\sqrt{2/3})$	0.990	$1.64\cdot10^{-3}$	$4.13\cdot10^{-3}$	$2.81\cdot10^{-3}$	$2.16\cdot10^{-3}$
	0.995	$8.13\cdot10^{-4}$	$3.27\cdot10^{-3}$	$1.98\cdot10^{-3}$	$1.33\cdot10^{-3}$
	0.999	$6.62\cdot10^{-5}$	$2.40\cdot10^{-3}$	$1.23\cdot10^{-3}$	$5.62\cdot10^{-4}$

4.2. Real data illustration

The finite sample behaviour of extreme geometric quantiles is illustrated on a two-dimensional dataset extracted from the Pima Indians Diabetes Database. This data set[1] was already considered by [15] and [12], among others. In the latter study, geometric iso-quantile curves with a high α are used to detect outliers in the data set. Using extreme quantiles for outlier detection was advocated in *e.g.* [5, 20] in the univariate case and [19] using depth-based quantile regions in the multivariate case; see also the monograph [1].

After working on the data set so as to eliminate missing values, the data set consists of $n = 392$ pairs (X_i, Y_i), where X_i is the body mass index (BMI) of the i-th individual and Y_i is its diastolic blood pressure. The centered data cloud is represented in Figure 7 with blue crosses, along with the geometric iso-quantile curve with $\alpha = 0.95$. While geometric quantiles with a moderate α tend to give a fair idea of the shape of the data cloud (see *e.g.* [12]), the same cannot be said for extreme geometric quantiles on this example. This is an illustration of the phenomenon described in Consequence 3 in Section 2: the norm of an extreme geometric quantile is the largest in the direction where the variance is the smallest.

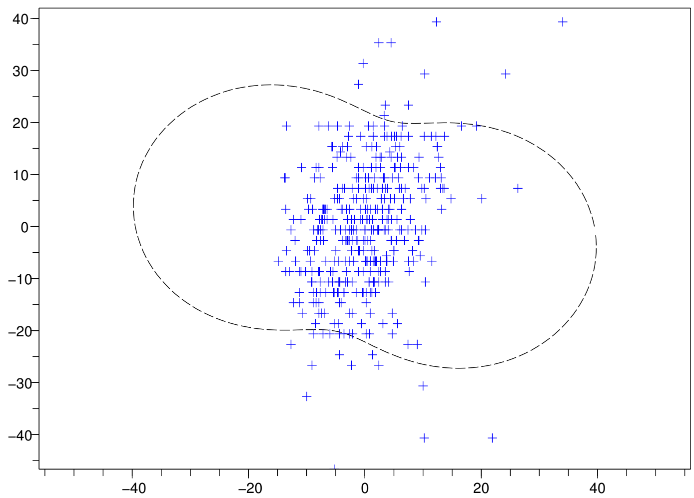

Figure 7: Pima Indians Diabetes data set. Black dashed line: estimate of the iso-quantile curve for $\alpha = 0.95$, with the estimator $\widehat{q}_n$.

We are thus led to think that here, outlier detection would be dangerous without a preliminary transformation–retransformation procedure [10].

[1] Available at `ftp.ics.uci.edu/pub/machine-learning-databases/pima-indians-diabetes`

5. CONCLUDING REMARKS

In this paper, we established the asymptotics of extreme geometric quantiles. A particular consequence of our results is that, if the underlying distribution possesses a finite covariance matrix Σ, then an extreme geometric quantile may be estimated accurately, no matter how extreme it is, with the help of the standard empirical estimator of Σ. This result is supported by our numerical study. The situation is very different from the univariate case, in which the asymptotic decay of a survival function can be linked to the asymptotic behaviour of an extreme quantile.

An additional issue, illustrated on a real data set, is that although central geometric quantile contours may roughly match the shape of the data cloud, this does not necessarily stay true for extreme iso-quantile curves. This is why we would advise practitioners to be cautious when using such a notion of multivariate quantile to detect outliers or analyze the extremes of a random vector. We believe that one can tackle this problem by applying a transformation–retransformation procedure, see [27] at the population level, and [9, 10] at the sample level. Future work on extreme geometric quantiles thus includes building and studying their analogues for transformed–retransformed data.

Finally, let us underline again that this work was carried out under moment conditions such as the existence of finite first and second-order moments for $\|X\|$. The case when these assumptions are violated is investigated in [17].

6. PROOFS

Some preliminary results are collected in Paragraph 6.1, their proofs are postponed to Paragraph 6.3. The proofs of the main results are provided in Paragraph 6.2.

6.1. Preliminary results

The first lemma provides some technical tools necessary to show Theorem 2.2(ii).

Lemma 6.1. *Let* $\varphi\colon \mathbb{R}^d \times \mathbb{R}_+ \times S^{d-1} \to \mathbb{R}$ *be the function defined by*

$$\varphi(x, r, v) \;=\; r^2 \left[1 + \frac{\langle x - rv,\, v\rangle}{\|x - rv\|} \right].$$

Then, for all $v \in S^{d-1}$, $\varphi(\cdot,\cdot,v)$ *is nonnegative and*

$$\forall x \in \mathbb{R}^d,\ \ \forall r \le \|x\|,\ \ \varphi(x,r,v) \le 2\,r^2 \quad \textit{and} \quad \forall r > \|x\|,\ \ \varphi(x,r,v) \le \|x\|^2\,.$$

In particular, $\varphi(x,r,v) \le 2\|x\|^2$ *for every* $(x,r,v) \in \mathbb{R}^d \times \mathbb{R}_+ \times S^{d-1}$.

The next lemma is the first step to prove Theorem 2.2(i).

Lemma 6.2. *Let* $u \in S^{d-1}$. *If* $\mathbb{E}\|X\| < \infty$ *then, for all* $v \in \mathbb{R}^d$,

$$\|q(\alpha u)\| \left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, v \right\rangle \to -\mathbb{E}\big\langle X - \langle X,u\rangle\, u,\, v\big\rangle \qquad \text{as} \quad \alpha \uparrow 1\,.$$

Lemma 6.3 below is a result which is similar to Lemma 6.2.

Lemma 6.3. *Let* $u \in S^{d-1}$. *If* $\mathbb{E}\|X\|^2 < \infty$ *then*

$$\|q(\alpha u)\|^2 \left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\rangle \to -\frac{1}{2}\, \mathbb{E}\big\|X - \langle X,u\rangle\, u\big\|^2 \qquad \text{as} \quad \alpha \uparrow 1\,.$$

Lemma 6.4 is the first step to prove Theorem 3.2. It is essentially a refinement of Lemma 6.2.

Lemma 6.4. *Let* $u \in S^{d-1}$. *If* $\mathbb{E}\|X\|^2 < \infty$ *then, for all* $v \in \mathbb{R}^d$,

$$\begin{aligned} &\|q(\alpha u)\| \left[\|q(\alpha u)\| \left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, v \right\rangle + \mathbb{E}\big\langle X - \langle X,u\rangle\, u,\, v\big\rangle \right] \to \\ &\quad \to\ \langle u,v\rangle\ \mathrm{Var}\langle X,u\rangle - \frac{1}{2}\,\langle u,v\rangle\ \mathbb{E}\big\|X - \langle X,u\rangle\, u\big\|^2 + \langle u,v\rangle\ \big\|\mathbb{E}\big(X - \langle X,u\rangle\, u\big)\big\|^2 \\ &\qquad -\ \mathrm{Cov}\big(\langle X,u\rangle, \langle X,v\rangle\big) \end{aligned}$$

as $\alpha \uparrow 1$.

Lemma 6.5 below is a refinement of Lemma 6.3. It is the second step to prove Theorem 3.2.

Lemma 6.5. *Let* $u \in S^{d-1}$. *If* $\mathbb{E}\|X\|^3 < \infty$ *then*

$$\begin{aligned} &\|q(\alpha u)\| \left(\|q(\alpha u)\|^2 \left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\rangle + \frac{1}{2}\, \mathbb{E}\big\|X - \langle X,u\rangle\, u\big\|^2 \right) \to \\ &\quad \to\ \mathbb{E}\Big(\langle X,u\rangle \Big[\big\langle X,\, \mathbb{E}\big(X - \langle X,u\rangle\, u\big) \big\rangle - \big\|X - \langle X,u\rangle\, u\big\|^2 \Big] \Big) \qquad \text{as} \quad \alpha \uparrow 1\,. \end{aligned}$$

6.2. Proofs of the main results

Proof of Proposition 2.1: From [14], it is known that if $u \in B^d$ then problem (1.1) has a unique solution $q(u) \in \mathbb{R}^d$. To prove the converse part of this result, use equation (2.1) to get

$$\left\| \mathbb{E}\left(\frac{X - q(u)}{\|X - q(u)\|} \right) \right\| = \|u\| \, .$$

Let us introduce the coordinate representations $X = (X_1, ..., X_d)$ and $q(u) = (q_1(u), ..., q_d(u))$. The Cauchy–Schwarz inequality yields

$$\begin{aligned} \|u\|^2 &= \left\| \mathbb{E}\left(\frac{X - q(u)}{\|X - q(u)\|} \right) \right\|^2 = \sum_{i=1}^{d} \left[\mathbb{E}\left(\frac{X_i - q_i(u)}{\|X - q(u)\|} \right) \right]^2 \\ &\le \sum_{i=1}^{d} \mathbb{E}\left(\frac{(X_i - q_i(u))^2}{\|X - q(u)\|^2} \right) = 1 \, . \end{aligned}$$

Furthermore, equality holds if and only if for all $i \in \{1, ..., d\}$, there exists $\mu_i \in \mathbb{R}$ such that

$$\frac{X_i - q_i(u)}{\|X - q(u)\|} = \mu_i$$

almost surely. In particular, if $w = (\mu_1, ..., \mu_d)$, this entails $X \in D = q(u) + \mathbb{R}w$ almost surely, which cannot hold since the distribution of X is not concentrated in a single straight line in $\mathbb{R}^d$. It follows that necessarily $\|u\|^2 < 1$, which is the result. □

Proof of Proposition 2.2:

(**i**) Note that (2.1) implies that, for any linear isometry h of $\mathbb{R}^d$ and every $u \in B^d$,

$$h(u) + \mathbb{E}\left(\frac{h(X) - h \circ q(u)}{\|X - q(u)\|} \right) = 0 \, .$$

Since h is a linear isometry, the random vectors X and $h(X)$ have the same distribution and the equality $\|X - q(u)\| = \|h(X) - h \circ q(u)\|$ holds almost surely. It follows that

$$h(u) + \mathbb{E}\left(\frac{X - h \circ q(u)}{\|X - h \circ q(u)\|} \right) = 0 \, .$$

Since $h(u) \in B^d$, it follows that $h \circ q(u) = q \circ h(u)$, which completes the proof of the first statement.

(**ii**) To prove the second part of Proposition 2.2, start by noting that since X and $-X$ have the same distribution, it holds that $\mathbb{E}\left(X/\|X\|\right) = 0$. The case $u = 0$ is then obtained via (2.1). If $u \neq 0$, up to using the first part of the result with a suitable linear isometry, we shall assume without loss of generality that

$u = (u_1, 0, ..., 0)$ for some constant $u_1 \in (0,1)$. It is then enough to prove that there exists some constant $q_1(u) > 0$ such that $q(u) = (q_1(u), 0, ..., 0)$. To this end, let us remark that, on the one hand, if $v_1 \in \mathbb{R}$ and $w = (1, 0, ..., 0)$ then

$$\forall j \in \{2, ..., d\}, \quad \mathbb{E}\left(\frac{X_j}{\|X - v_1 w\|}\right) = 0 , \tag{6.1}$$

since, for all $j \in \{2, ..., d\}$, the random vectors $(X_1, ..., X_{j-1}, -X_j, X_{j+1}, ..., X_d)$ and X have the same distribution. On the other hand, the dominated convergence theorem entails that the function

$$v_1 \mapsto \mathbb{E}\left(\frac{X_1 - v_1}{\|X - v_1 w\|}\right)$$

is continuous, converges to 1 at $-\infty$, is equal to 0 at 0 and converges to -1 at $+\infty$. Thus, the intermediate value theorem yields that there exists some constant $q_1(u) > 0$ such that

$$u_1 + \mathbb{E}\left(\frac{X_1 - q_1(u)}{\|X - q_1(u) w\|}\right) = 0 . \tag{6.2}$$

Consequently, collecting (6.1) and (6.2) yields

$$u + \mathbb{E}\left(\frac{X - q_1(u) w}{\|X - q_1(u) w\|}\right) = 0$$

and it only remains to apply (2.1) to finish the proof of the second statement.

(iii) To show the third statement, use the first result to obtain that the function $g\colon \|u\| \mapsto \|q(u)\|$ is indeed well-defined; since the geometric quantile function is continuous, so is g. Assume that g is not strictly increasing: namely, there exist $u_1,\ u_2 \in B^d$ such that $\|u_1\| < \|u_2\|$ and $\|q(u_1)\| \geq \|q(u_2)\|$. Since $q(0) = 0$, it is a consequence of the intermediate value theorem that one may find $u,\ v \in B^d$ such that $\|u\| < \|v\|$ and $\|q(u)\| = \|q(v)\|$. Let h be an isometry such that $h(u/\|u\|) = h(v/\|v\|)$; then

$$\|q(h(u))\| = \|q(u)\| = \|q(v)\| = \|q(h(v))\|$$

and

$$\frac{q(h(u))}{\|q(h(u))\|} = \frac{h(u)}{\|h(u)\|} = \frac{h(v)}{\|h(v)\|} = \frac{q(h(v))}{\|q(h(v))\|} .$$

In other words, $q(h(u))$ and $q(h(v))$ have the same direction and magnitude, so that they are necessarily equal, which entails that $h(u) = h(v)$ because the geometric quantile function is one-to-one. This is a contradiction because $\|h(u)\| = \|u\| < \|v\| = \|h(v)\|$, and the third statement is proven.

(iv) Assume that $\|q(u)\|$ does not tend to infinity as $\|u\| \to 1$; since g is increasing, it tends to a finite positive limit r. In other words, $\|q(u)\| \leq r$ for every $u \in B^d$, which is a contradiction since the geometric quantile function maps B^d onto $\mathbb{R}^d$, and the proof is complete. □

Proof of Theorem 2.1:

(i) If the first statement were false, then one could find a sequence (v_n) contained in B^d such that $\|v_n\| \to 1$ and such that $(\|q(v_n)\|)$ does not tend to infinity. Up to extracting a subsequence, one can assume that $(\|q(v_n)\|)$ is bounded. Again, up to extraction, one can assume that (v_n) converges to some $v_\infty \in S^{d-1}$ and that $(q(v_n))$ converges to some $q_\infty \in \mathbb{R}^d$. Moreover, it is straightforward to show that for every $u_1, u_2, q_1, q_2 \in \mathbb{R}^d$

$$\big|\psi(u_1, q_1) - \psi(u_2, q_2)\big| \;\le\; \big\{1 + \|u_2\|\big\}\, \|q_2 - q_1\| + \|q_1\|\, \|u_2 - u_1\|$$

so that the function ψ is continuous on $\mathbb{R}^d \times \mathbb{R}^d$. Recall then that the definition of $q(v_n)$ implies that for every $q \in \mathbb{R}^d$, $\psi(v_n, q(v_n)) \le \psi(v_n, q)$ and let n tend to infinity to obtain

$$q_\infty \;=\; \arg\min_{q \in \mathbb{R}^d} \psi(v_\infty, q)\;.$$

Because $v \in S^{d-1}$, this contradicts Proposition 2.1, and the proof of the first statement is complete: $\|q(v)\| \to \infty$ as $\|v\| \to 1$.

(ii) Pick a sequence (v_n) of elements of B^d converging to u and remark that from (2.1),

$$v_n + \mathbb{E}\left(\frac{X - q(v_n)}{\|X - q(v_n)\|}\right) \;=\; 0$$

for every integer n. Hence, for n large enough, the following equality holds:

$$v_n + \mathbb{E}\left(\left\|\frac{X}{\|q(v_n)\|} - \frac{q(v_n)}{\|q(v_n)\|}\right\|^{-1} \left[\frac{X}{\|q(v_n)\|} - \frac{q(v_n)}{\|q(v_n)\|}\right]\right) \;=\; 0\;. \tag{6.3}$$

Since the sequence $(q(v_n)/\|q(v_n)\|)$ is bounded it is enough to show that its only accumulation point is u. Let then u^* be an accumulation point of this sequence. Since $\|q(v_n)\| \to \infty$, we may let $n \to \infty$ in (6.3) and use the dominated convergence theorem to obtain $u - u^* = 0$, which completes the proof. □

Proof of Theorem 2.2:

(i) Let $(u, w_1, ..., w_{d-1})$ be an orthonormal basis of $\mathbb{R}^d$ and consider the following expansion:

$$\frac{q(\alpha u)}{\|q(\alpha u)\|} \;=\; b(\alpha)\, u + \sum_{k=1}^{d-1} \beta_k(\alpha)\, w_k \tag{6.4}$$

where $b(\alpha), \beta_1(\alpha), ..., \beta_{d-1}(\alpha)$ are real numbers. It immediately follows that

$$\frac{q(\alpha u)}{\|q(\alpha u)\|} - u - \frac{1}{\|q(\alpha u)\|}\big\{\mathbb{E}(X) - \langle \mathbb{E}(X), u\rangle\, u\big\} \;= \tag{6.5}$$

$$=\; \big(b(\alpha) - 1\big)\, u + \sum_{k=1}^{d-1} \frac{\|q(\alpha u)\|\beta_k(\alpha) - \mathbb{E}\langle X, w_k\rangle}{\|q(\alpha u)\|}\, w_k\;.$$

Lemma 6.2 implies that
(6.6)
$$\|q(\alpha u)\| \left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|}, w_k \right\rangle = -\|q(\alpha u)\| \beta_k(\alpha) \;\to\; -\mathbb{E}\langle X, w_k\rangle \quad \text{as } \alpha \uparrow 1$$

for all $k \in \{1, ..., d-1\}$. Besides, let us note that $q(\alpha u)/\|q(\alpha u)\| \in S^{d-1}$ entails

$$b^2(\alpha) + \sum_{k=1}^{d-1} \beta_k^2(\alpha) = 1 . \tag{6.7}$$

Theorem 2.1 shows that $b(\alpha) \to 1$ as $\alpha \uparrow 1$ and thus (6.6) yields:

$$\begin{aligned} \|q(\alpha u)\| \big(1 - b(\alpha)\big) &= \frac{1}{2} \|q(\alpha u)\| \big(1 - b^2(\alpha)\big) \big(1 + \mathrm{o}(1)\big) \\ &= \frac{1}{2} \|q(\alpha u)\| \sum_{k=1}^{d-1} \beta_k^2(\alpha) \big(1 + \mathrm{o}(1)\big) \;\to\; 0 \quad \text{as } \alpha \uparrow 1 . \end{aligned} \tag{6.8}$$

Collecting (6.5), (6.6) and (6.8), we obtain

$$\frac{q(\alpha u)}{\|q(\alpha u)\|} - u - \frac{1}{\|q(\alpha u)\|} \big\{\mathbb{E}(X) - \langle \mathbb{E}(X), u\rangle u\big\} = \mathrm{o}\!\left(\frac{1}{\|q(\alpha u)\|}\right) \quad \text{as } \alpha \uparrow 1$$

which is the first result.

(ii) Recall (6.4) and use Lemma 6.2 to obtain

$$\|q(\alpha u)\| \left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|}, w_k \right\rangle \to -\mathbb{E}\langle X, w_k\rangle \quad \text{as } \alpha \uparrow 1 ,$$

for all $k \in \{1, ..., d-1\}$, leading to

$$\|q(\alpha u)\|^2 \beta_k^2(\alpha) \;\to\; \big[\mathbb{E}\langle X, w_k\rangle\big]^2 \quad \text{as } \alpha \uparrow 1 \tag{6.9}$$

for all $k \in \{1, ..., d-1\}$. Recall (6.7) and use Lemma 6.3 to get

$$\|q(\alpha u)\|^2 \big[\alpha b(\alpha) - 1\big] \;\to\; -\frac{1}{2} \mathbb{E}\big\|X - \langle X, u\rangle u\big\|^2 \quad \text{as } \alpha \uparrow 1 . \tag{6.10}$$

Since $(u, w_1, ..., w_{d-1})$ is an orthonormal basis of $\mathbb{R}^d$, one has the identity

$$\big\|X - \langle X, u\rangle u\big\|^2 = \sum_{k=1}^{d-1} \langle X, w_k\rangle^2 . \tag{6.11}$$

Collecting (6.9), (6.10) and (6.11) leads to

$$\|q(\alpha u)\|^2 \left[1 - \alpha b(\alpha) - \frac{1}{2} \sum_{k=1}^{d-1} \beta_k^2(\alpha)\right] \to \frac{1}{2} \sum_{k=1}^{d-1} \mathrm{Var}\langle X, w_k\rangle \quad \text{as } \alpha \uparrow 1 .$$

Therefore,
(6.12)
$$\|q(\alpha u)\|^2 \left[1 - \alpha b(\alpha) - \frac{1}{2} \big(1 - b^2(\alpha)\big)\right] \to \frac{1}{2} \sum_{k=1}^{d-1} \mathrm{Var}\langle X, w_k\rangle \quad \text{as } \alpha \uparrow 1 ,$$

and easy calculations show that

$$1 - \alpha b(\alpha) - \frac{1}{2}\left(1 - b^2(\alpha)\right) \;=\; \frac{1}{2}\left[(1-\alpha)\,(1+\alpha) + \left(\alpha - b(\alpha)\right)^2\right] . \tag{6.13}$$

Finally, in view of Lemma 6.2,

$$\|q(\alpha u)\| \left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, u \right\rangle \;\to\; 0 \qquad \text{as} \;\; \alpha \uparrow 1$$

which is equivalent to

$$\|q(\alpha u)\|^2 \left(\alpha - b(\alpha)\right)^2 \;\to\; 0 \qquad \text{as} \;\; \alpha \uparrow 1 . \tag{6.14}$$

Collecting (6.12), (6.13) and (6.14), we obtain

$$\|q(\alpha u)\|^2\,(1-\alpha) \;\to\; \frac{1}{2} \sum_{k=1}^{d-1} \mathrm{Var}\langle X, w_k\rangle \qquad \text{as} \;\; \alpha \uparrow 1 .$$

Remarking that, for every orthonormal basis $(e_1, ..., e_d)$ of $\mathbb{R}^d$,

$$\sum_{k=1}^{d} \mathrm{Var}\langle X, e_k\rangle \;=\; \sum_{k=1}^{d} e_k'\,\Sigma e_k \;=\; \mathrm{tr}\,\Sigma \tag{6.15}$$

proves that

$$\|q(\alpha u)\|^2\,(1-\alpha) \;\to\; \frac{1}{2}\left(\mathrm{tr}\,\Sigma - u'\Sigma u\right) \;\geq\; 0 \qquad \text{as} \;\; \alpha \uparrow 1 .$$

Finally, note that if we had $\mathrm{tr}\,\Sigma - u'\Sigma u = 0$ then by (6.15) we would have that $\mathrm{Var}\langle X, w_k\rangle = 0$ for all $k \in \{1, ..., d-1\}$. Thus the projection of X onto the orthogonal complement of $\mathbb{R}u$ would be almost surely constant and X would be contained in a single straight line in $\mathbb{R}^d$, which is a contradiction. This completes the proof of Theorem 2.2. □

Proof of Theorem 3.1: Note that

$$\sqrt{1-\alpha_n}\;\; \widehat{q}_n(\alpha_n u) \;\to\; \left[\frac{1}{2}\left(\mathrm{tr}\,\Sigma - u'\Sigma u\right)\right]^{1/2} u \tag{6.16}$$

almost surely as $n \to \infty$. Moreover, by Theorems 2.1 and 2.2
(6.17)

$$\sqrt{1-\alpha_n}\;\; q(\alpha_n u) \;=\; \sqrt{1-\alpha_n}\;\|q(\alpha_n u)\|\,\frac{q(\alpha_n u)}{\|q(\alpha_n u)\|} \;\to\; \left[\frac{1}{2}\left(\mathrm{tr}\,\Sigma - u'\Sigma u\right)\right]^{1/2} u$$

almost surely as $n \to \infty$. Combining (6.16) and (6.17) completes the proof. □

Proof of Theorem 3.2: Consider the following representation:

$$\sqrt{n(1-\alpha_n)}\,\big(\widehat{q}_n(\alpha_n u) - q(\alpha_n u)\big) \;=\; T_{1,n} + T_{2,n} + T_{3,n}$$

with
$$T_{1,n} \;=\; \sqrt{n}\left(\left[\frac{1}{2}\left\{\mathrm{tr}\,\widehat{\Sigma}_n - u'\widehat{\Sigma}_n u\right\}\right]^{1/2} - \left[\frac{1}{2}\left\{\mathrm{tr}\,\Sigma - u'\Sigma u\right\}\right]^{1/2}\right)\frac{q(\alpha_n u)}{\|q(\alpha_n u)\|}\,,$$
$$T_{2,n} \;=\; \sqrt{n}\left(\left[\frac{1}{2}\left\{\mathrm{tr}\,\Sigma - u'\Sigma u\right\}\right]^{1/2} - \sqrt{1-\alpha_n}\;\|q(\alpha_n u)\|\right)\frac{q(\alpha_n u)}{\|q(\alpha_n u)\|}$$
and
$$T_{3,n} \;=\; -\sqrt{n(1-\alpha_n)}\;\|\widehat{q}_n(\alpha_n u)\|\left(\frac{q(\alpha_n u)}{\|q(\alpha_n u)\|} - u\right).$$

We start by examining the convergence of $T_{1,n}$. Observe first that

$$\begin{aligned} T_{1,n} \;&=\; \sqrt{n}\,\frac{1}{\sqrt{2}}\,\frac{\left\{\mathrm{tr}\,\widehat{\Sigma}_n - u'\widehat{\Sigma}_n u\right\} - \left\{\mathrm{tr}\,\Sigma - u'\Sigma u\right\}}{\left\{\mathrm{tr}\,\widehat{\Sigma}_n - u'\widehat{\Sigma}_n u\right\}^{1/2} + \left\{\mathrm{tr}\,\Sigma - u'\Sigma u\right\}^{1/2}}\,\frac{q(\alpha_n u)}{\|q(\alpha_n u)\|} \\ &=\; \sqrt{n}\,\frac{\left\{\mathrm{tr}\,\widehat{\Sigma}_n - u'\widehat{\Sigma}_n u\right\} - \left\{\mathrm{tr}\,\Sigma - u'\Sigma u\right\}}{2\sqrt{2}\,\left\{\mathrm{tr}\,\Sigma - u'\Sigma u\right\}^{1/2}}\,u\big(1+\mathrm{o}_{\mathbb{P}}(1)\big) \qquad \text{as } n\to\infty \end{aligned}$$

in view of Theorem 2.1(i) and from the consistency of $\widehat{\Sigma}_n$. Denote by M the Gaussian centred limit of $\sqrt{n}(\widehat{\Sigma}_n - \Sigma)$ (see *e.g.* [24]). Since the map $A \mapsto \mathrm{tr}\,A - u'Au$ is linear, it follows that

$$\sqrt{n}\,\frac{\left\{\mathrm{tr}\,\widehat{\Sigma}_n - u'\widehat{\Sigma}_n u\right\} - \left\{\mathrm{tr}\,\Sigma - u'\Sigma u\right\}}{2\sqrt{2}\,\left\{\mathrm{tr}\,\Sigma - u'\Sigma u\right\}^{1/2}} \;\xrightarrow{d}\; Y \qquad \text{as } n\to\infty$$

where Y is a centred Gaussian random variable. Now, clearly $Z := Yu$ is a Gaussian centred random vector and we have

$$T_{1,n} \xrightarrow{d} Z \qquad \text{as } n\to\infty\,. \tag{6.18}$$

The sequence $T_{2,n}$ is controlled in the following way: using Lemmas 6.4 and 6.5 and following the steps of the proof of Theorem 2.2(ii), we obtain

$$\begin{aligned} \|q(\alpha_n u)\|^2\,(1-\alpha_n) \;&=\; \frac{1}{2}\big(\mathrm{tr}\,\Sigma - u'\Sigma u\big) + \mathrm{O}\big(\|q(\alpha_n u)\|^{-1}\big) \\ &=\; \frac{1}{2}\big(\mathrm{tr}\,\Sigma - u'\Sigma u\big) + \mathrm{O}\big(\sqrt{1-\alpha_n}\,\big) \qquad \text{as } n\to\infty\,. \end{aligned}$$

As a consequence

$$\|T_{2,n}\| = \mathrm{O}\left(\sqrt{n(1-\alpha_n)}\right) = \mathrm{o}(1) \qquad \text{as } n\to\infty\,. \tag{6.19}$$

We conclude by controlling $T_{3,n}$. Theorem 2.2 entails

$$\begin{aligned} \|T_{3,n}\| \;&=\; \mathrm{O}_{\mathbb{P}}\left(\sqrt{n(1-\alpha_n)}\;\frac{\|\widehat{q}_n(\alpha_n u)\|}{\|q(\alpha_n u)\|}\right) \\ &=\; \mathrm{O}_{\mathbb{P}}\left(\sqrt{n(1-\alpha_n)}\;\left[\frac{\mathrm{tr}\,\widehat{\Sigma}_n - u'\widehat{\Sigma}_n u}{\mathrm{tr}\,\Sigma - u'\Sigma u}\right]^{1/2}\right) \\ &=\; \mathrm{O}_{\mathbb{P}}\left(\sqrt{n(1-\alpha_n)}\right) \;=\; \mathrm{o}_{\mathbb{P}}(1) \qquad \text{as } n\to\infty \end{aligned} \tag{6.20}$$

by the consistency of $\widehat{\Sigma}_n$. Combining (6.18), (6.19) and (6.20) completes the proof. □

6.3. Proofs of the preliminary results

Proof of Lemma 6.1: The fact that φ is nonnegative and the inequality

$$\forall r \le \|x\| \,, \qquad \varphi(x,r,v) \le 2\,r^2 \tag{6.21}$$

are straightforward consequences of the Cauchy–Schwarz inequality. Furthermore, φ can be rewritten as

$$\varphi(x,r,v) \;=\; r^2 \left[\frac{\big\|x - \langle x,v\rangle\, v\big\|^2}{\|x-rv\| \,\big[\|x-rv\| - \langle x-rv,\, v\rangle\big]} \right].$$

Let us now remark that, if $\|x\| < r$, then, by the Cauchy–Schwarz inequality, $\langle x - rv,\, v\rangle = \langle x,v\rangle - r < 0$ which makes it clear that

$$\varphi(x,r,v)\; 1\!\!1_{\{\|x\|<r\}} \;\le\; r^2\, \frac{\big\|x - \langle x,v\rangle\, v\big\|^2}{\|x-rv\|^2}\; 1\!\!1_{\{\|x\|<r\}} \;=:\; \psi(x,r,v)\; 1\!\!1_{\{\|x\|<r\}}\,. \tag{6.22}$$

Since $\|x-rv\|^2 = \|x\|^2 - 2r\langle x,\, v\rangle + r^2$, the function $\psi(x,\cdot,v)$ is differentiable on $(\|x\|,\, +\infty)$ and some easy computations yield

$$\frac{\partial\psi}{\partial r}(x,r,v) \;=\; 2\,r\,\big[\|x\|^2 - r\langle x,\, v\rangle\big]\, \frac{\big\|x - \langle x,v\rangle\, v\big\|^2}{\|x-rv\|^4}\,.$$

If $\langle x,\, v\rangle \le 0$ then $\psi(x,\cdot,v)$ is increasing on $(\|x\|,\, +\infty)$ and thus

$$\forall r > \|x\|\,, \qquad \psi(x,r,v) \le \lim_{r\to+\infty} \psi(x,r,v) \;=\; \|x - \langle x,v\rangle v\|^2 \;\le\; \|x\|^2\,. \tag{6.23}$$

Otherwise, if $\langle x,v\rangle > 0$ then $\psi(x,\cdot,v)$ reaches its global maximum over $[\|x\|,+\infty)$ at $\|x\|^2/\langle x,\, v\rangle$ and therefore,

$$\forall r > \|x\|\,, \qquad \psi(x,r,v) \le \psi\left(x,\, \frac{\|x\|^2}{\langle x,v\rangle},\, v\right) = \|x\|^2\,. \tag{6.24}$$

Collecting (6.22), (6.23) and (6.24) yields

$$\varphi(x,r,v)\; 1\!\!1_{\{\|x\|<r\}} \;\le\; \|x\|^2\; 1\!\!1_{\{\|x\|<r\}}\,. \tag{6.25}$$

Combining (6.21) and (6.25) shows that $\varphi(x,r,v) \le 2\|x\|^2$ for every $r > 0$ and every $v \in S^{d-1}$ and completes the proof of the result. $\square$

Proof of Lemma 6.2: Let $v \in \mathbb{R}^d$ and $W_\alpha(\cdot,v)\colon \mathbb{R}^d \to \mathbb{R}$ be the function defined by

$$W_\alpha(x,v) \;=\; \left[\left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|}\right\|^{-1} - 1\right] \left\langle \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, v\right\rangle.$$

For α close enough to 1, (2.1) entails

$$\left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, v \right\rangle + \mathbb{E}\big(W_\alpha(X,v)\big) + \frac{1}{\|q(\alpha u)\|}\,\mathbb{E}\langle X, v\rangle \;=\; 0\,. \tag{6.26}$$

It is therefore enough to show that

$$\|q(\alpha u)\|\,\mathbb{E}\big(W_\alpha(X,v)\big) \;\to\; -\langle u,\, v\rangle\,\mathbb{E}\langle X, u\rangle \qquad \text{as} \;\; \alpha \uparrow 1\,. \tag{6.27}$$

Since, for every $x \in \mathbb{R}^d$,

$$\left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\|^2 \;=\; 1 - \frac{2}{\|q(\alpha u)\|}\left\langle x,\, \frac{q(\alpha u)}{\|q(\alpha u)\|}\right\rangle + \frac{\|x\|^2}{\|q(\alpha u)\|^2}\,, \tag{6.28}$$

it follows from a Taylor expansion and Theorem 2.1 that

$$\|q(\alpha u)\|\,W_\alpha(X,v) \;\to\; -\langle u, v\rangle\,\langle X, u\rangle \qquad \text{almost surely as} \;\; \alpha \uparrow 1\,. \tag{6.29}$$

Besides,

$$\begin{aligned}\left| \left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\|^{-1} - 1 \right| \;=\; & \left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\|^{-1} \\ & \times \left[1 + \left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\| \right]^{-1} \\ & \times \left| \frac{2}{\|q(\alpha u)\|}\left\langle x,\, \frac{q(\alpha u)}{\|q(\alpha u)\|}\right\rangle - \frac{\|x\|^2}{\|q(\alpha u)\|^2} \right|,\end{aligned}$$

and the Cauchy–Schwarz inequality yields

$$\left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\|^{-1} \left\langle \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, v \right\rangle \;\le\; \|v\|\,.$$

Thus, using the triangular inequality and the Cauchy–Schwarz inequality, it follows that

$$|W_\alpha(x,v)| \;\le\; \|v\| \left[1 + \left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\| \right]^{-1} \frac{\|x\|}{\|q(\alpha u)\|} \left[2 + \frac{\|x\|}{\|q(\alpha u)\|} \right].$$

Consequently, one has

$$\|q(\alpha u)\|\;|W_\alpha(x,v)|\;\mathbb{1}_{\{\|x\|\le\|q(\alpha u)\|\}} \;\le\; 3\,\|v\|\,\|x\|\,\mathbb{1}_{\{\|x\|\le\|q(\alpha u)\|\}}\,.$$

Furthermore, the reverse triangle inequality entails, for $x \in \mathbb{R}^d$ such that $\|x\| > \|q(\alpha u)\|$:

$$\left[1 + \left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\| \right]^{-1} \le \frac{\|q(\alpha u)\|}{\|x\|}\,,$$

and therefore,

$$\|q(\alpha u)\|\;|W_\alpha(x,v)|\;\mathbb{1}_{\{\|x\|>\|q(\alpha u)\|\}} \;\le\; 3\,\|v\|\,\|x\|\,\mathbb{1}_{\{\|x\|>\|q(\alpha u)\|\}}\,.$$

Finally,

$$\|q(\alpha u)\|\;|W_\alpha(X,v)| \;\le\; 3\,\|v\|\,\|X\|$$

so that the integrand in (6.27) is bounded from above by an integrable random variable. One can now recall (6.29) and apply the dominated convergence theorem to obtain (6.27). The proof is complete. □

Proof of Lemma 6.3: Let $Z_\alpha \colon \mathbb{R}^d \to \mathbb{R}$ be the function defined by

$$Z_\alpha(x) \;=\; 1 + \left\langle \frac{x - q(\alpha u)}{\|x - q(\alpha u)\|},\, \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\rangle .$$

For α close enough to 1, (2.1) yields

$$\left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\rangle + \mathbb{E}\big(Z_\alpha(X)\big) \;=\; 0 \tag{6.30}$$

and it thus remains to prove that

$$\|q(\alpha u)\|^2\, \mathbb{E}\big(Z_\alpha(X)\big) \;\to\; \frac{1}{2}\, \mathbb{E}\big\|X - \langle X, u\rangle\, u\big\|^2 \qquad \text{as } \ \alpha \uparrow 1\,.$$

To this end, rewrite Z_α as

$$Z_\alpha(x) \;=\; 1 - \left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\|^{-1} \left[1 - \frac{1}{\|q(\alpha u)\|} \left\langle x,\, \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\rangle \right]. \tag{6.31}$$

It thus follows from equation (6.28), Theorem 2.1 and a Taylor expansion that

$$Z_\alpha(x) \;=\; \frac{1}{2\,\|q(\alpha u)\|^2} \left\langle x - \left\langle x,\, \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\rangle \frac{q(\alpha u)}{\|q(\alpha u)\|},\, x \right\rangle \big(1 + \mathrm{o}(1)\big)$$

for all $x \in \mathbb{R}^d$. Using Theorem 2.1 again, we then get

$$\|q(\alpha u)\|^2\, Z_\alpha(X) \;\to\; \|X\|^2 - \langle X, u\rangle^2 \;=\; \|X - \langle X, u\rangle\, u\|^2 \qquad \text{almost surely as } \ \alpha \uparrow 1\,. \tag{6.32}$$

To conclude the proof, let $\varphi \colon \mathbb{R}^d \times \mathbb{R}_+ \times S^{d-1} \to \mathbb{R}$ be the function defined by

$$\varphi(x, r, v) \;=\; r^2 \left[1 + \frac{\langle x - rv,\, v\rangle}{\|x - rv\|} \right].$$

Note that $\|q(\alpha u)\|^2 Z_\alpha(x) = \varphi\big(x, \|q(\alpha u)\|, q(\alpha u)/\|q(\alpha u)\|\big)$. By Lemma 6.1:

$$\|q(\alpha u)\|^2\, Z_\alpha(X) \;=\; \varphi\Big(X,\, \|q(\alpha u)\|,\, q(\alpha u)/\|q(\alpha u)\|\Big) \;\le\; 2\,\|X\|^2$$

and the right-hand side is an integrable random variable. Use then (6.32) and the dominated convergence theorem to complete the proof. □

Proof of Lemma 6.4: Let $v \in \mathbb{R}^d$ and recall the notation

$$W_\alpha(x, v) \;=\; \left[\left\| \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\|^{-1} - 1 \right] \left\langle \frac{x}{\|q(\alpha u)\|} - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, v \right\rangle$$

from the proof of Lemma 6.2. From (6.26) there, it is enough to show that

$$\begin{aligned} &\|q(\alpha u)\|\; \mathbb{E}\Big(\|q(\alpha u)\|\, W_\alpha(X, v) + \langle u, v\rangle\, \langle X, u\rangle \Big) \;\to \\ &\quad \to\; \frac{1}{2}\, \langle u, v\rangle\, \mathbb{E}\big\|X - \langle X, u\rangle\, u\big\|^2 - \langle u, v\rangle\, \mathrm{Var}\langle X, u\rangle \\ &\qquad + \mathrm{Cov}\big(\langle X, u\rangle, \langle X, v\rangle\big) - \langle u, v\rangle\, \big\|\mathbb{E}(X - \langle X, u\rangle\, u)\big\|^2 \end{aligned} \tag{6.33}$$

as $\alpha \uparrow 1$. Use now (6.28) in the proof of Lemma 6.2, Theorem 2.2(i) and a Taylor expansion to obtain after some cumbersome computations that

$$\begin{aligned}
&\|q(\alpha u)\| \Big(\|q(\alpha u)\| \, W_\alpha(X,v) + \langle u,v\rangle \, \langle X,u\rangle \Big) = \\
&\quad = \frac{1}{2} \langle u,v\rangle \, \|X - \langle X,u\rangle u\|^2 - \langle u,v\rangle \, \langle X,u\rangle \, \big(\langle X,u\rangle - \mathbb{E}\langle X,u\rangle\big) \\
&\qquad + \langle X,u\rangle \, \big(\langle X,v\rangle - \mathbb{E}\langle X,v\rangle\big) - \langle u,v\rangle \, \big\langle X, \, \mathbb{E}\big(X - \langle X,u\rangle u\big)\big\rangle \\
&\qquad + \sum_{j=0}^{2} \|X\|^j \, \varepsilon_j\big(\alpha, X, q(\alpha u)\big)
\end{aligned}$$

with probability 1, where for all $j \in \{0,1,2\}$, $\varepsilon_j(\alpha,y,z) \to 0$ as $\max(1-\alpha, \|y\|/\|z\|) \downarrow 0$. In particular

$$\begin{aligned}
&\|q(\alpha u)\| \Big(\|q(\alpha u)\| \, W_\alpha(X,v) + \langle u,v\rangle \, \langle X,u\rangle \Big) \to \\
&\quad \to \frac{1}{2} \langle u,v\rangle \, \big\|X - \langle X,u\rangle u\big\|^2 - \langle u,v\rangle \, \langle X,u\rangle \, \big(\langle X,u\rangle - \mathbb{E}\langle X,u\rangle\big) \\
&\qquad - \langle u,v\rangle \, \big\langle X, \, \mathbb{E}\big(X - \langle X,u\rangle u\big)\big\rangle + \langle X,u\rangle \, \big(\langle X,v\rangle - \mathbb{E}\langle X,v\rangle\big) \, ,
\end{aligned} \tag{6.34}$$

almost surely as $\alpha \uparrow 1$. The proof shall be complete provided we can apply the dominated convergence theorem to the left-hand side of (6.34). To this end, let $\delta \in (0,1)$ be such that

$$\alpha \in (1-\delta, 1) \ \text{ and } \ \frac{\|X\|}{\|q(\alpha u)\|} < \delta \implies \max_{0 \le j \le 2} \big|\varepsilon_j(\alpha, X, q(\alpha u))\big| \le 1 \, .$$

Equality (6.34) thus entails for α close enough to 1:

$$\begin{aligned}
&\|q(\alpha u)\| \, \Big| \|q(\alpha u)\| \, W_\alpha(X,v) + \langle u,v\rangle \, \langle X,u\rangle \Big| \, \mathbb{1}_{\{\|X\| < \delta \|q(\alpha u)\|\}} \le \\
&\qquad \le P_1\big(\|X\|\big) \, \mathbb{1}_{\{\|X\| < \delta \|q(\alpha u)\|\}}
\end{aligned}$$

where P_1 is a real polynomial of degree 2. Besides, it is a consequence of the definition of $W_\alpha(X,v)$ and the Cauchy–Schwarz inequality that

$$\begin{aligned}
&\|q(\alpha u)\| \, \Big| \|q(\alpha u)\| \, W_\alpha(X,v) + \langle u,v\rangle \, \langle X,u\rangle \Big| \, \mathbb{1}_{\{\|X\| \ge \delta \|q(\alpha u)\|\}} \le \\
&\qquad \le \frac{2\,(1+\delta)\,\|v\|}{\delta^2} \, \|X\|^2 \, \mathbb{1}_{\{\|X\| \ge \delta \|q(\alpha u)\|\}} \, .
\end{aligned}$$

One can conclude that there exists a real polynomial P_2 of degree 2 such that

$$\|q(\alpha u)\| \, \Big| \|q(\alpha u)\| \, W_\alpha(X,v) + \langle u,v\rangle \, \langle X,u\rangle \Big| \le P_2\big(\|X\|\big)$$

so that the integrand in (6.33) is bounded by an integrable random variable. Recall (6.34) and apply the dominated convergence theorem to complete the proof. □

Proof of Lemma 6.5: The proof is similar to that of Lemma 6.4. Recall from the proof of Lemma 6.3 the notation

$$Z_\alpha(x) \;=\; 1 + \left\langle \frac{x - q(\alpha u)}{\|x - q(\alpha u)\|},\, \frac{q(\alpha u)}{\|q(\alpha u)\|} \right\rangle .$$

From (6.30) there, it is enough to show that

$$\text{(6.35)}\qquad \|q(\alpha u)\|\;\mathbb{E}\Big(\|q(\alpha u)\|^2\, Z_\alpha(X) - \frac{1}{2}\,\mathbb{E}\big\|X - \langle X,u\rangle\, u\big\|^2 \Big) \;\to\; \to\; \mathbb{E}\Big(\langle X,u\rangle \Big[\big\|X - \langle X,u\rangle\, u\big\|^2 - \big\langle X,\, \mathbb{E}\big(X - \langle X,u\rangle u\big)\big\rangle \Big]\Big)$$

as $\alpha \uparrow 1$. We first use (6.28) in the proof of Lemma 6.2, equation (6.31) in the proof of Lemma 6.3, Theorem 2.2(i) and a Taylor expansion to obtain after some burdensome computations that

$$\text{(6.36)}\qquad \|q(\alpha u)\|\,\Big(\|q(\alpha u)\|^2\, Z_\alpha(X) - \frac{1}{2}\,\big\|X - \langle X,u\rangle\, u\big\|^2 \Big) \;=\; =\; \langle X,u\rangle\,\Big(\big\|X - \langle X,u\rangle\, u\big\|^2 - \big\langle X,\, \mathbb{E}\big(X - \langle X,u\rangle u\big)\big\rangle \Big) + \sum_{j=0}^{3} \|X\|^j\, \varepsilon_j\big(\alpha, X, q(\alpha u)\big)$$

with probability 1, where for $j \in \{0,1,2,3\}$, $\varepsilon_j(\alpha, y, z) \to 0$ as $\max(1-\alpha, \|y\|/\|z\|) \downarrow 0$. Especially

$$\text{(6.37)}\qquad \|q(\alpha u)\|\,\Big(\|q(\alpha u)\|^2\, Z_\alpha(X) - \frac{1}{2}\,\big\|X - \langle X,u\rangle\, u\big\|^2 \Big) \;\to\; \to\; \langle X,u\rangle\,\Big(\big\|X - \langle X,u\rangle\, u\big\|^2 - \big\langle X,\, \mathbb{E}\big(X - \langle X,u\rangle\, u\big)\big\rangle \Big)$$

as $\alpha \uparrow 1$. Our aim is now to apply the dominated convergence theorem to the left-hand side of (6.35). To this end, pick $\delta \in (0,1)$ such that

$$\alpha \in (1-\delta, 1) \;\text{ and }\; \frac{\|X\|}{\|q(\alpha u)\|} < \delta \;\Longrightarrow\; \max_{0 \le j \le 3} \big|\varepsilon_j\big(\alpha, X, q(\alpha u)\big)\big| \;\le\; 1\,.$$

Equality (6.36) thus entails for α close enough to 1:

$$\|q(\alpha u)\|\,\Big| \|q(\alpha u)\|^2\, Z_\alpha(X) - \frac{1}{2}\,\big\|X - \langle X,u\rangle\, u\big\|^2 \Big|\; \mathbb{1}_{\{\|X\| < \delta \|q(\alpha u)\|\}} \;\le\; \le\; P_1\big(\|X\|\big)\; \mathbb{1}_{\{\|X\| < \delta \|q(\alpha u)\|\}}$$

where P_1 is a real polynomial of degree 3. Moreover, the Cauchy–Schwarz inequality yields

$$\|q(\alpha u)\|\,\Big| \|q(\alpha u)\|^2\, Z_\alpha(X) - \frac{1}{2}\,\|X - \langle X,u\rangle\, u\|^2 \Big|\; \mathbb{1}_{\{\|X\| \ge \delta \|q(\alpha u)\|\}} \;\le\; \le\; \frac{4 + \delta^2}{2\,\delta^3}\; \|X\|^3\; \mathbb{1}_{\{\|X\| \ge \delta \|q(\alpha u)\|\}}\,.$$

Consequently, there exists a real polynomial P_2 of degree 3 such that

$$\|q(\alpha u)\| \left| \|q(\alpha u)\|^2 Z_\alpha(X) - \frac{1}{2} \left\|X - \langle X, u\rangle u\right\|^2 \right| \leq P_2\big(\|X\|\big) .$$

We conclude that the integrand in (6.35) is bounded by an integrable random variable. Recall (6.37) and apply the dominated convergence theorem to complete the proof. □

REFERENCES

[1] Aggarwal, C. C. (2013). *Outlier analysis*, Springer, New York.

[2] Bingham, N. H.; Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*, Cambridge, U.K.: Cambridge University Press.

[3] Breckling, J. and Chambers, R. (1988). M-quantiles, *Biometrika*, **75**(4), 761–771.

[4] Breckling, J.; Kokic, P. and Lübke, O. (2001). A note on multivariate M-quantiles, *Stat. Probab. Lett.*, **55**, 39–44.

[5] Burridge, P. and Robert Taylor, A. M. (2006). Additive outlier detection via extreme-value theory, *J. Time Series Anal.*, **27**(5), 685–701.

[6] Cadre, B. and Gannoun, A. (2000). Asymptotic normality of consistent estimate of the conditional L_1-median, *Ann. I.S.U.P.*, **44**, 13–35.

[7] Cambanis, S.; Huang, S. and Simons, G. (1981). On the theory of elliptically contoured distributions, *J. Multivariate Anal.*, **11**(3), 368–385.

[8] 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**(1), 18–43.

[9] Chakraborty, B.; Chaudhuri, P. and Oja, H. (1998). Operating transformation retransformation on spatial median and angle test, *Statist. Sinica*, **8**, 767–784.

[10] Chakraborty, B. (2001). On affine equivariant multivariate quantiles, *Ann. Inst. Statist. Math.*, **53**(2), 380–403.

[11] Chakraborty, B. (2003). On multivariate quantile regression, *J. Statist. Plann. Inference*, **110**(2), 109–132.

[12] Chaouch, M. and Goga, C. (2010). Design-based estimation for geometric quantiles with application to outlier detection, *Comput. Statist. Data Anal.*, **54**, 2214–2229.

[13] Chaouch, M. and Laïb, N. (2013). Nonparametric multivariate L_1-median regression estimation with functional covariates, *Electron. J. Stat.*, **7**, 1553–1586.

[14] Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data, *J. Amer. Statist. Assoc.*, **91**(434), 862–872.

[15] Cheng, Y. and De Gooijer, J. G. (2007). On the u-th geometric conditional quantile, *J. Statist. Plann. Inference*, **137**, 1914–1930.

[16] Dhar, S. S.; Chakraborty, B. and Chaudhuri, P. (2014). Comparison of multivariate distributions using quantile-quantile plots and related tests, *Bernoulli*, **20**(3), 1484–1506.

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

[18] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, Springer, New York.

[19] He, Y. and Einmahl, J. H. J. (2016). Estimation of extreme depth-based quantile regions, *Journal of the Royal Statistical Society: Series B*, to appear.

[20] Huberta, M.; Dierckx, G. and Vanpaemela, D. (2012). Detecting influential data points for the Hill estimator in Pareto-type distributions, *Comput. Statist. Data Anal.*, **65**, 13–28.

[21] Kemperman, J. H. B. (1987). *The median of a finite measure on a Banach space.* In "Statistical Data Analysis Based on the L^1-norm and Related Methods" (Y. Dodge, Ed.), 217–230, North Holland, Amsterdam.

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

[23] Koltchinskii, V. I. (1997). M-estimation, convexity and quantiles, *Ann. Statist.*, **25**(2), 435–477.

[24] Neudecker, H. and Wesselman, A. M. (1990). The asymptotic variance matrix of the sample correlation matrix, *Linear Algebra Appl.*, **127**, 589–599.

[25] Serfling, R. (2002). Quantile functions for multivariate analysis: approaches and applications, *Stat. Neerl.*, **56**(2), 214–232.

[26] Serfling, R. (2004). Nonparametric multivariate descriptive measures based on spatial quantiles, *J. Statist. Plann. Inference*, **56**(2), 214–232.

[27] Serfling, R. (2010). Equivariance and invariance properties of multivariate quantile and related functions, and the role of standardization, *J. Nonparametr. Stat.*, **22**, 915–936.

[28] Small, C. G. (1990). A survey of multidimensional medians, *Int. Stat. Rev.*, **58**(3), 263–277.

[29] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *J. Amer. Statist. Assoc.*, **73**, 812–815.