On Directional Multiple-Output Quantile Regression

Davy Paindaveine[*,a,1], Miroslav Šiman[b]

[a]*E.C.A.R.E.S. and Département de Mathématique, Université Libre de Bruxelles, Avenue Roosevelt, 50, CP114, 1050 Brussels, Belgium, Tel: +3226505892, Fax: +3226504475*

[b]*E.C.A.R.E.S., Université Libre de Bruxelles, Avenue Roosevelt, 50, CP114, 1050 Brussels, Belgium*
and
Institute of Information Theory and Automation of the ASCR, Pod Vodárenskou věží 4, CZ-182 08, Prague 8, Czech Republic

---

Abstract

This paper sheds some new light on *projection* quantiles. Contrary to the sophisticated set analysis used in Kong and Mizera (2008), we adopt a more parametric approach and study the subgradient conditions associated with these quantiles. In this setup, we introduce Lagrange multipliers which can be interpreted in various interesting ways, in particular in a portfolio optimization context. The corresponding projection quantile regions were already shown to coincide with the halfspace depth ones in Kong and Mizera (2008), but we provide here an alternative proof (completely based on projection quantiles) that has the advantage of leading to an *exact* computation of halfspace depth regions from projection quantiles. Above all, we systematically consider the regression case, which was barely touched in Kong and Mizera (2008). We show in particular that the regression quantile regions introduced in Hallin, Paindaveine, and Šiman (2010) can also be obtained from projection (regression) quantiles, which may lead to a faster computation of those regions in some particular cases.

Key words: Multivariate quantile, Quantile regression, Multiple-output regression, Halfspace depth, Portfolio optimization, Value-at-Risk
2000 MSC: 62H05, 62J05, 65C60

---

*Corresponding author.
Email address: dpaindav@ulb.ac.be (Davy Paindaveine)
URL: http://homepages.ulb.ac.be/~dpaindav (Davy Paindaveine)

[1]Davy Paindaveine is also member of ECORE, the association between CORE and ECARES.

1. Introduction.

1.1. Multiple-output quantile regression.

Applications of the celebrated (Koenker and Bassett [11]) theory of quantile regression are without number, virtually in all quantitative fields, including economics and econometrics, biomedical studies and clinical trials, biostatistics, and environmental studies; see [9] for an extensive presentation of the topic. Quantile regression techniques have been quickly extended to nonlinear and nonparametric (functional) regression, and modified for handling count, longitudinal, time series or censored data.

On the other hand, their extension to the multiple-output case has been a long-standing statistical challenge. And despite several attempts to define multiple-output regression quantiles (see, e.g., [2], [3], [12]), this theory still remains mostly univariate. In fact, Koenker [9] himself reports multiple-output quantile regression on the list of "problems that fall into the twilight of quantile regression research".

In a world where multivariate data are the rule rather than the exception, this single-output nature of quantile regression clearly constitutes a severe limitation. The main issue is of course the lack of a satisfactory and universally accepted concept of *multivariate quantiles*; we refer to [23] for an excellent survey of the huge literature devoted to multivariate quantiles. The problem is even more delicate if the ultimate goal is to define a concept of multiple-output regression quantile because not every multivariate quantile can be generalized to the regression context.

1.2. Two recent proposals.

Interestingly, two concepts of multivariate quantiles that are potentially useful for multiple-output quantile regression were investigated very recently by Kong and Mizera [13] and by Hallin, Paindaveine, and Šiman [6, 7]—hereafter KM08 and HPŠ10, respectively. Both are of a *directional* nature and define, for distributions on $\mathbb{R}^m$, quantiles that are indexed by an *order* $\tau \in (0,1)$ and a *direction* $\mathbf{u} \in \mathcal{S}^{m-1} := \{\mathbf{y} \in \mathbb{R}^m : \|\mathbf{y}\| = 1\}$, or equivalently, by the vector $\boldsymbol{\tau} = \tau\mathbf{u}$ ranging over the open unit ball (deprived of the origin) $\mathcal{B}^m := \{\mathbf{y} \in \mathbb{R}^m : 0 < \|\mathbf{y}\| < 1\}$.

In the KM approach[2], the $(\tau\mathbf{u})$-quantile of an m-dimensional random vector $\mathbf{Y}$ may simply be defined as the point

$$\mathbf{q}_{\mathrm{KM},\tau\mathbf{u}} := q_\tau(\mathbf{u}'\mathbf{Y})\mathbf{u} \quad (\in \mathbb{R}^m) \tag{1.1}$$

[2]From Section 2 onwards, we will often rather speak of *the projection approach*, mainly because Kong and Mizera [14] reports that the univariate quantiles of projections considered in KM08 were not regarded there as the basis of a multivariate quantile concept.

or as the hyperplane $\pi_{\mathrm{KM},\tau\mathbf{u}}$, orthogonal to $\mathbf{u}$ at $\mathbf{q}_{\mathrm{KM},\tau\mathbf{u}}$, where $q_\tau(X) := \inf\{x \in \mathbb{R} : \mathrm{P}[X \leq x] \geq \tau\}$ stands for the univariate τ-quantile of the random variable X (KM08 also considers other versions of univariate quantiles). The *quantile biplot* contours $B(\tau) := \{\mathbf{q}_{\mathrm{KM},\tau\mathbf{u}} : \mathbf{u} \in \mathcal{S}^{m-1}\}$ (indexed by τ), which are naturally associated with the point-valued quantiles $\mathbf{q}_{\mathrm{KM},\tau\mathbf{u}}$, are hardly satisfactory since they lack any reasonable form of equivariance, heavily depend on the choice of an origin, and moreover exhibit disturbing non-convex shapes with a tendency to self-intersection. However, defining the "upper" quantile halfspaces

$$H^+_{\mathrm{KM},\tau\mathbf{u}} := \big\{\mathbf{y} \in \mathbb{R}^m : \mathbf{u}'\mathbf{y} \geq \mathbf{u}'\mathbf{q}_{\mathrm{KM},\tau\mathbf{u}}\big\}, \tag{1.2}$$

the hyperplane-valued quantiles determine fixed-τ regions

$$R_{\mathrm{KM}}(\tau) := \bigcap_{\mathbf{u}\in\mathcal{S}^{m-1}} \{H^+_{\mathrm{KM},\tau\mathbf{u}}\}, \qquad \tau \in (0,1), \tag{1.3}$$

which happen to coincide (Theorem 3 in KM08) with the celebrated halfspace depth ones—we refer to [17], [21], and [26] for a comprehensive treatment of (location) depth. Multiple-output regression quantiles based on this directional quantile concept are briefly discussed in KM08 too; see also [25].

Of course, sample quantile regions $R^{(n)}_{\mathrm{KM}}(\tau)$ and quantile biplot contours $B^{(n)}(\tau)$ can be defined as the natural empirical analogs of the population objects above, and $R^{(n)}_{\mathrm{KM}}(\tau)$ still coincides with the (sample) halfspace depth region of order τ. However, the construction of any such $R^{(n)}_{\mathrm{KM}}(\tau)$ or $B^{(n)}(\tau)$ via KM quantiles in principle involves computing infinitely many univariate quantiles (one for each $\mathbf{u} \in \mathcal{S}^{m-1}$), which of course is impossible in practice. The competing approach from HPŠ10 (which was inspired by an original idea from [16]) does much better in this respect.

With the same notation as above, the HPŠ10 $\boldsymbol{\tau}$-quantiles are defined as the standard regression τ-quantile hyperplanes $\pi_{\mathrm{HP\check{S}},\tau\mathbf{u}}$ obtained when regressing $\mathbf{Y}_\mathbf{u} := \mathbf{u}'\mathbf{Y}$ on the marginals of $\mathbf{Y}^\perp_\mathbf{u} := \mathbf{\Gamma}'_\mathbf{u}\mathbf{Y}$ and a constant term, where $\mathbf{\Gamma}_\mathbf{u}$ stands for an arbitrary $m \times (m-1)$ matrix such that the columns of $(\mathbf{u} \,\vdots\, \mathbf{\Gamma}_\mathbf{u})$ constitute an orthonormal basis of $\mathbb{R}^m$ (see the location version of Definition 2.3 below); the vector $\mathbf{u}$ therefore indicates the direction of the "vertical" axis in this single-output regression. Denoting the halfspace "above" $\pi_{\mathrm{HP\check{S}},\tau\mathbf{u}}$ by $H^+_{\mathrm{HP\check{S}},\tau\mathbf{u}}$ (where "above" is with respect to the natural orientation of this vertical axis), it turns out that the resulting quantile regions

$$R_{\mathrm{HP\check{S}}}(\tau) := \bigcap_{\mathbf{u}\in\mathcal{S}^{m-1}} \{H^+_{\mathrm{HP\check{S}},\tau\mathbf{u}}\}, \qquad \tau \in (0,1), \tag{1.4}$$

also coincide with the halfspace depth ones. This extends to the sample case, where the regions $R^{(n)}_{\mathrm{HP\check{S}}}(\tau)$ coincide with their KM08 counterparts $R^{(n)}_{\mathrm{KM}}(\tau)$. *Unlike their*

KM counterparts $\pi^{(n)}_{\mathrm{KM},\tau\mathbf{u}}$, however, there is typically only a *finite* collection of HPŠ10 quantile hyperplanes $\pi^{(n)}_{\mathrm{HP\check{S}},\tau\mathbf{u}}$ for fixed τ (these hyperplanes are then piecewise constant functions of $\mathbf{u}$), a collection that includes all prolonged facets of the halfspace depth region $R^{(n)}_{\mathrm{HP\check{S}}}(\tau) = R^{(n)}_{\mathrm{KM}}(\tau)$. From a theoretical point of view, HPŠ quantiles, when compared to KM quantiles, therefore are more directly related to sample halfspace depth regions. From a practical point of view, this of course may be seen as a strong hint that the HPŠ quantiles provide a much better way of computing halfspace depth regions (again, obtaining these regions from KM quantiles in principle requires computing the intersection in the sample version of (1.3), which runs over an infinite collection of quantile halfspaces).

1.3. Our contribution.

In KM08, projection quantiles are thoroughly investigated by means of set analysis in the location case, and may be defined via various different concepts of univariate quantiles. In contrast, the present paper focuses on the standard univariate quantile $q_\tau(.)$ defined above, adopts a more parametric approach, and considers the general regression case throughout (which is hardly touched in KM08). In particular, we study the subgradient conditions associated with (location and regression) projection quantiles, introduce the corresponding Lagrange multipliers and interpret them in various ways, in particular in a portfolio optimization context.

We then turn to projection quantile regions. (i) In the location case, we present an alternative proof (completely based on projection quantiles) that the sample projection quantile regions $R^{(n)}_{\mathrm{KM}}(\tau)$ coincide with the halfspace depth ones. This proof further clarifies the link between projection quantiles and halfspace depth regions, and paves the way to an *exact* computation of sample halfspace depth regions from these quantiles. This significantly improves over KM08, where the only proposed strategy to obtain the regions $R^{(n)}_{\mathrm{KM}}(\tau)$ consists in sampling $\mathcal{S}^{m-1}$ in the sample version of (1.3), which clearly yields approximate halfspace depth regions only. Most importantly, this exact computation surprisingly may be faster than the one based on HPŠ quantiles in some particular cases. (ii) In the regression case, we could not reduce the infinite intersection defining sample projection quantile regions to a finite one, so that exact computation in principle remains infeasible. However, we show that *the HPŠ regression quantile regions can be obtained exactly from projection quantiles* (which is much less trivial than in the location case; see Section 4.2 for details). Parallel to the location case, this may result in a faster computation of HPŠ regions in some particular cases. Our MATLAB implementation of the algorithm for computing (regression) quantile regions from projection quantiles can be freely downloaded from http://homepages.ulb.ac.be/˜dpaindav and is extensively described in the companion paper [18].

The outline of the paper is as follows. Section 2 gives a unified presentation of projection quantiles and HPŠ quantiles in the general regression case. Section 3 derives and interprets the gradient conditions for projection quantiles, and links these quantiles to portfolio optimization. Section 4 turns to quantile regions: we first present (Section 4.1) an alternative proof that sample projection quantile regions coincide with the halfspace depth ones in the location case, and show that this proof leads to an exact computation of halfspace depth regions from projection quantiles. Then we define (Section 4.2) projection quantile regions in the regression case and establish that the HPŠ regression quantile regions can be obtained exactly from projection quantiles. Section 5 briefly discusses computational aspects of projection quantiles and projection quantile regions, leaving the details to [18]. The paper ends with an appendix (collecting technical proofs) and a commented picture gallery.

2. The projection and HPŠ multiple-output regression quantiles.

Consider the multiple-output regression setup in which some m-variate random vector $\mathbf{Y} = (Y_1, \ldots, Y_m)'$ of responses is to be related to a p-variate random vector $\mathbf{X} = (X_1, \ldots, X_p)'$ of regressors, where $X_1 = 1$ a.s. and the other X_j's are random. In the sequel, we let $\mathbf{X} =: (1, \mathbf{W}')'$, which makes of $\{(\mathbf{w}', \mathbf{y}')' : \mathbf{w} \in \mathbb{R}^{p-1}, \mathbf{y} \in \mathbb{R}^m\} = \mathbb{R}^{p-1} \times \mathbb{R}^m$ the natural space for considering fitted regression objects. For $p = 1$, we obtain the important *location case*, in which multiple-output regression quantiles simply reduce to *multivariate quantiles*.

The multiple-output *directional* regression quantiles we introduce below are indexed by vectors $\boldsymbol{\tau}$ ranging over the (open) unit ball (deprived of the origin) $\mathcal{B}^m = \{\mathbf{y} \in \mathbb{R}^m : 0 < \|\mathbf{y}\| < 1\}$ of the response space $\mathbb{R}^m$. To stress their directional nature, we factorize the index $\boldsymbol{\tau}$ into $\boldsymbol{\tau} =: \tau\mathbf{u}$, where $\tau = \|\boldsymbol{\tau}\| \in (0, 1)$ is the *order* of the quantile and $\mathbf{u} \in \mathcal{S}^{m-1}$ is its *direction*. Letting $t \mapsto \rho_\tau(t) := t(\tau - \mathbb{I}_{[t<0]})$ denote the usual τ-quantile check function, we consider the following broad class of directional regression quantiles.

Definition 2.1. *Let $\mathcal{M} = \{\mathcal{M}_\mathbf{u} : \mathbf{u} \in \mathcal{S}^{m-1}\}$ be a family of convex subsets of $\mathbb{R}^{m+p}$ making the mapping $\mathbf{u} \mapsto \mathcal{M}_\mathbf{u}$ injective and reducing to $\mathcal{M} = \big\{\mathcal{M}_u := \{(\mathbf{a}', b)' \in \mathbb{R}^{p+1} : b = u\} : u \in \{-1, 1\}\big\}$ for $m = 1$. Then, for any $\boldsymbol{\tau} = \tau\mathbf{u}$, with $\tau \in (0, 1)$*

and $\mathbf{u} \in \mathcal{S}^{m-1}$*, the* $\mathcal{M}$*-type regression* $\boldsymbol{\tau}$*-quantile of* $\mathbf{Y}$ *with respect to* $\mathbf{X} = (1, \mathbf{W}')'$ *is defined as any element of the collection* $\Pi_{\mathcal{M},\boldsymbol{\tau}}$ *of hyperplanes* $\pi_{\mathcal{M},\boldsymbol{\tau}} := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{b}'_{\mathcal{M},\boldsymbol{\tau}}\mathbf{y} = \mathbf{a}'_{\mathcal{M},\boldsymbol{\tau}}(1, \mathbf{w}')'\}$ *such that*

$$(\mathbf{a}'_{\mathcal{M},\boldsymbol{\tau}}, \mathbf{b}'_{\mathcal{M},\boldsymbol{\tau}})' \in \underset{(\mathbf{a}',\mathbf{b}')' \in \mathcal{M}_{\mathbf{u}}}{\operatorname{argmin}} \Psi_\tau(\mathbf{a}, \mathbf{b}), \qquad \textit{with} \;\; \Psi_\tau(\mathbf{a}, \mathbf{b}) := \mathrm{E}[\rho_\tau(\mathbf{b}'\mathbf{Y} - \mathbf{a}'\mathbf{X})]. \tag{2.1}$$

Each hyperplane $\pi_{\mathcal{M},\boldsymbol{\tau}}$ *characterizes a lower (open) and an upper (closed) regression quantile halfspace* $H^-_{\mathcal{M},\boldsymbol{\tau}} = H^-_{\mathcal{M},\boldsymbol{\tau}}(\mathbf{a}_{\mathcal{M},\boldsymbol{\tau}}, \mathbf{b}_{\mathcal{M},\boldsymbol{\tau}}) := \big\{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{b}'_{\mathcal{M},\boldsymbol{\tau}}\mathbf{y} < \mathbf{a}'_{\mathcal{M},\boldsymbol{\tau}}(1, \mathbf{w}')'\big\}$ *and* $H^+_{\mathcal{M},\boldsymbol{\tau}} = H^+_{\mathcal{M},\boldsymbol{\tau}}(\mathbf{a}_{\mathcal{M},\boldsymbol{\tau}}, \mathbf{b}_{\mathcal{M},\boldsymbol{\tau}}) := \big\{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{b}'_{\mathcal{M},\boldsymbol{\tau}}\mathbf{y} \geq \mathbf{a}'_{\mathcal{M},\boldsymbol{\tau}}(1, \mathbf{w}')'\big\}$*, respectively.*

The convexity assumption on the $\mathcal{M}_{\mathbf{u}}$'s is motivated by theoretical and practical considerations and turns the minimization problem (2.1) into a standard convex optimization exercise. The injectivity assumption ensures that the resulting multiple-output regression quantiles bear a clear directional information. As for the specific form of $\mathcal{M}$ in the single-output case $m = 1$, it guarantees that the quantiles of Definition 2.1 there reduce to the standard Koenker and Bassett [11] quantiles. Finally, we stress that first-order moment assumptions on $\mathbf{Y}$ and $\mathbf{W}$ are tacitly part of Definition 2.1 (and, unless otherwise stated, also of Definitions 2.2-2.3 below).

Just as in the single-output case, the minimization problem (2.1) may have several solutions, yielding distinct hyperplanes $\pi_{\mathcal{M},\boldsymbol{\tau}}$. However, as shown by Proposition 2.1 below, this does not occur under

ASSUMPTION (A). *The distribution of* $\mathbf{Z} := (\mathbf{W}', \mathbf{Y}')'$ *is absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^{m+p-1}$*, with a density that has connected support, and admits finite first-order moments.*

Proposition 2.1. *Under Assumption (A), the minimizer* $(\mathbf{a}'_{\mathcal{M},\boldsymbol{\tau}}, \mathbf{b}'_{\mathcal{M},\boldsymbol{\tau}})'$ *in (2.1), hence also the resulting quantile hyperplane* $\pi_{\mathcal{M},\boldsymbol{\tau}}$*, is unique for any* $\boldsymbol{\tau} \in \mathcal{B}^m$*.*

In this work, the emphasis will mainly be on the collection $\mathcal{M}$ given by

$$\mathcal{M}_{\text{proj}} = \big\{\mathcal{M}_{\text{proj},\mathbf{u}} := \{(\mathbf{a}', \mathbf{b}')' \in \mathbb{R}^{m+p} : \mathbf{b} = \mathbf{u}\} : \mathbf{u} \in \mathcal{S}^{m-1}\big\}, \tag{2.2}$$

which leads to *projection* regression quantiles—reducing, in the location case, to the directional quantiles considered in KM08. Modifying slightly the definition of projection quantiles by subtracting the constant quantity $\Psi_\tau(\mathbf{0}, \mathbf{u})$ from the corresponding objective function in (2.1) does not affect projection quantiles, but allows to avoid any moment assumption on $\mathbf{Y}$—hence any moment assumption at all in the location case $p = 1$. More specifically, we adopt the following definition, which requires finite first-order moments for $\mathbf{W}$ only.

Definition 2.2. *For any $\boldsymbol{\tau} = \tau\mathbf{u}$, with $\tau \in (0,1)$ and $\mathbf{u} \in \mathcal{S}^{m-1}$, the projection regression $\boldsymbol{\tau}$-quantile of $\mathbf{Y}$ with respect to $\mathbf{X} = (1, \mathbf{W}')'$ is defined as any element of the collection $\Pi_{\mathrm{proj},\boldsymbol{\tau}}$ of hyperplanes $\pi_{\mathrm{proj},\boldsymbol{\tau}} := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{u}'\mathbf{y} = \mathbf{a}'_{\mathrm{proj},\boldsymbol{\tau}}(1, \mathbf{w}')'\}$ such that*

$$\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}} \in \underset{\mathbf{a}\in\mathbb{R}^p}{\operatorname{argmin}} \big(\Psi_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}) - \Psi_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{0})\big), \tag{2.3}$$

where $\Psi_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}) := \Psi_\tau(\mathbf{a}, \mathbf{u}) = \mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y} - \mathbf{a}'\mathbf{X})]$. The corresponding lower (open) and upper (closed) regression quantile halfspaces are $H^-_{\mathrm{proj},\boldsymbol{\tau}} = H^-_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}}) := \big\{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{u}'\mathbf{y} < \mathbf{a}'_{\mathrm{proj},\boldsymbol{\tau}}(1, \mathbf{w}')'\big\}$ and $H^+_{\mathrm{proj},\boldsymbol{\tau}} = H^+_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}}) := \big\{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{u}'\mathbf{y} \geq \mathbf{a}'_{\mathrm{proj},\boldsymbol{\tau}}(1, \mathbf{w}')'\big\}$, respectively.

In the location case ($p = 1$), we have $\pi_{\mathrm{proj},\boldsymbol{\tau}} = \pi_{\mathrm{KM},\tau\mathbf{u}}$ and $H^+_{\mathrm{proj},\boldsymbol{\tau}} = H^+_{\mathrm{KM},\tau\mathbf{u}}$; see (1.2). In the general regression case, the quantiles from Definition 2.2 clearly reduce to the ordinary regression quantiles of the projection $\mathbf{u}'\mathbf{Y}$ on the marginals of $\mathbf{W}$ and a constant term, which enlightens many of their features. Clearly, projection quantiles are intrinsically univariate. Nevertheless, the concept of projection quantiles is richer than one would expect at first sight.

Another interesting choice of $\mathcal{M}$ leads to the regression quantiles introduced in Definition 6.1 of HPŠ10. This definition can equivalently be reformulated as follows.

Definition 2.3. *For any $\boldsymbol{\tau} = \tau\mathbf{u}$, with $\tau \in (0,1)$ and $\mathbf{u} \in \mathcal{S}^{m-1}$, the HPŠ regression $\boldsymbol{\tau}$-quantile of $\mathbf{Y}$ with respect to $\mathbf{X} = (1, \mathbf{W}')'$ is defined as any element of the collection $\Pi_{\mathrm{HPŠ},\boldsymbol{\tau}}$ of hyperplanes $\pi_{\mathrm{HPŠ},\boldsymbol{\tau}} := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{u}'\mathbf{y} = \mathbf{d}'_{\boldsymbol{\tau}}\boldsymbol{\Gamma}'_{\mathbf{u}}(\mathbf{w}', \mathbf{y}')' + c_{\boldsymbol{\tau}}\}$ such that*

$$(c_{\boldsymbol{\tau}}, \mathbf{d}'_{\boldsymbol{\tau}})' \in \underset{(c,\mathbf{d}')'\in\mathbb{R}^{m+p-1}}{\operatorname{argmin}} \mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y} - \mathbf{d}'\boldsymbol{\Gamma}'_{\mathbf{u}}(\mathbf{W}', \mathbf{Y}')' - c)], \tag{2.4}$$

where $\boldsymbol{\Gamma}_{\mathbf{u}}$ stands for an arbitrary $(m+p-1) \times (m+p-2)$ matrix such that the matrix $(\dot{\mathbf{u}} \vdots \boldsymbol{\Gamma}_{\mathbf{u}})$ is orthogonal for $\dot{\mathbf{u}} = (\mathbf{0}'_{p-1}, \mathbf{u}')'$.

The HPŠ quantiles are also standard single-output regression quantiles of the same response $\mathbf{u}'\mathbf{Y}$ as in the projection approach, but this time with regressors consisting of the marginals of $\mathbf{\Gamma}'_{\mathbf{u}}(\mathbf{W}', \mathbf{Y}')'$ and a constant term. As shown by the following result, these quantiles also fit in the class of directional quantiles introduced in Definition 2.1 and are associated with

$$\mathcal{M}_{\text{HPŠ}} = \big\{\mathcal{M}_{\text{HPŠ},\mathbf{u}} := \{(\mathbf{a}', \mathbf{b}')' \in \mathbb{R}^{m+p} : \mathbf{b}'\mathbf{u} = 1\} : \mathbf{u} \in \mathcal{S}^{m-1}\big\}. \tag{2.5}$$

Proposition 2.2. *Fix $\boldsymbol{\tau} = \tau\mathbf{u}$, with $\tau \in (0,1)$ and $\mathbf{u} \in \mathcal{S}^{m-1}$, and assume that the underlying distribution has finite first-order moments. Then any $\mathcal{M}_{\text{HPŠ}}$-type regression $\boldsymbol{\tau}$-quantile hyperplane $\pi_{\mathcal{M}_{\text{HPŠ}},\boldsymbol{\tau}}$ of $\mathbf{Y}$ with respect to $\mathbf{X} = (1, \mathbf{W}')'$ is a HPŠ regression $\boldsymbol{\tau}$-quantile hyperplane $\pi_{\text{HPŠ},\boldsymbol{\tau}}$ from Definition 2.3, and vice versa.*

In line with the notation already introduced in Definition 2.3, the quantities $\pi_{\mathcal{M}_{\text{HPŠ}},\boldsymbol{\tau}}$, $\mathbf{a}_{\mathcal{M}_{\text{HPŠ}},\boldsymbol{\tau}}$, $\mathbf{b}_{\mathcal{M}_{\text{HPŠ}},\boldsymbol{\tau}}$, $H^{-}_{\mathcal{M}_{\text{HPŠ}},\boldsymbol{\tau}}$, and $H^{+}_{\mathcal{M}_{\text{HPŠ}},\boldsymbol{\tau}}$ from Definition 2.1 will simply be denoted as $\pi_{\text{HPŠ},\boldsymbol{\tau}}$, $\mathbf{a}_{\text{HPŠ},\boldsymbol{\tau}}$, $\mathbf{b}_{\text{HPŠ},\boldsymbol{\tau}}$, $H^{-}_{\text{HPŠ},\boldsymbol{\tau}}$, and $H^{+}_{\text{HPŠ},\boldsymbol{\tau}}$, respectively. For HPŠ quantiles, it is not possible to get rid of the first-order moment conditions, even in the location case; see the comment above Assumption (A) in HPŠ10. More importantly, note that the projection quantiles from Definition 2.2 are actually constrained (with a $\mathbf{y}$-space projection orthogonal to $\mathbf{u}$) versions of the HPŠ quantiles.

In the sequel, we focus on $\mathcal{M}_{\text{proj}}$ and $\mathcal{M}_{\text{HPŠ}}$. Still, there may be other interesting choices of $\mathcal{M} = \{\mathcal{M}_{\mathbf{u}}\}$, leading to original concepts of multiple-output regression quantiles. In this context, we thank an anonymous referee for pointing out [24] as a possible source of inspiration for defining such $\mathcal{M}$. The definition of alternative collections $\mathcal{M}$ of interest and the investigation of the resulting regression quantiles are, however, beyond the scope of the paper. Here, we only point out that it seems desirable to restrict to sets $\mathcal{M}_{\mathbf{u}}$ that (i) are defined by means of equality constraints that are linear—or at least linearizable—in $(\mathbf{a}', \mathbf{b}')'$ (to keep the computational burden as light as possible) and that (ii) keep $\mathbf{b}$ away from the zero vector of $\mathbb{R}^m$ (because otherwise we could always get $(\mathbf{a}'_{\mathcal{M},\boldsymbol{\tau}}, \mathbf{b}'_{\mathcal{M},\boldsymbol{\tau}})' = \mathbf{0}$). Clearly, both $\mathcal{M}_{\text{proj}}$ and $\mathcal{M}_{\text{HPŠ}}$ meet these properties.

Finally, we consider the empirical case. Assume that a sample of n observations $\mathbf{Z}_i := (\mathbf{W}'_i, \mathbf{Y}'_i)'$, $i = 1, \ldots, n$, is available (in the sequel, we often use this simple notation instead of $(\mathbf{X}'_i, \mathbf{Y}'_i)' := (1, \mathbf{W}'_i, \mathbf{Y}'_i)'$, $i = 1, \ldots, n$). Empirical regression quantiles can then simply be obtained as the natural sample analogs of the population concepts in Definition 2.1. To be more specific, we define the *sample $\mathcal{M}$-type regression $\boldsymbol{\tau}$-quantile* of the $\mathbf{Y}_i$'s with respect to the $\mathbf{X}_i$'s, $i = 1, \ldots, n$, as any element of the collection $\Pi^{(n)}_{\mathcal{M},\boldsymbol{\tau}}$ of hyperplanes $\pi^{(n)}_{\mathcal{M},\boldsymbol{\tau}} := \{(\mathbf{w}',\mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{b}^{(n)\prime}_{\mathcal{M},\boldsymbol{\tau}}\mathbf{y} = \mathbf{a}^{(n)\prime}_{\mathcal{M},\boldsymbol{\tau}}(1,\mathbf{w}')'\}$ such that

$$(\mathbf{a}^{(n)\prime}_{\mathcal{M},\boldsymbol{\tau}}, \mathbf{b}^{(n)\prime}_{\mathcal{M},\boldsymbol{\tau}})' \in \underset{(\mathbf{a}',\mathbf{b}')'\in\mathcal{M}_{\mathbf{u}}}{\operatorname{argmin}} \Psi^{(n)}_{\boldsymbol{\tau}}(\mathbf{a},\mathbf{b}), \quad \text{with} \quad \Psi^{(n)}_{\boldsymbol{\tau}}(\mathbf{a},\mathbf{b}) := \frac{1}{n}\sum_{i=1}^{n} \rho_\tau(\mathbf{b}'\mathbf{Y}_i - \mathbf{a}'\mathbf{X}_i). \tag{2.6}$$

These empirical quantiles allow to define in an obvious way the sample analogs $H^{(n)-}_{\mathcal{M},\boldsymbol{\tau}}$ and $H^{(n)+}_{\mathcal{M},\boldsymbol{\tau}}$ of the lower and upper quantile halfspaces $H^{-}_{\mathcal{M},\boldsymbol{\tau}}$ and $H^{+}_{\mathcal{M},\boldsymbol{\tau}}$ of Definition 2.1. Corresponding sample quantities will be denoted by symbols used for the population ones and equipped with $^{(n)}$.

Of course, empirical distributions are inherently discrete, so that sample $\boldsymbol{\tau}$-quantiles and halfspaces are not uniquely defined in general. However, the set of minimizers of (2.6) must be connected and convex for any given $\boldsymbol{\tau}$, which readily follows from the convexity of minimized objective functions.

3. Fixed-u analysis of projection regression quantiles.

In this section, we derive and discuss the subgradient conditions associated with projection regression quantiles. For the sake of comparison, we also extend the HPŠ quantile subgradient conditions to the regression setup (in HPŠ10, they are explicitly stated in the location case only). Finally, we link projection quantiles with portfolio optimization.

3.1. Subgradient conditions.

Under Assumption (A), the objective function $\mathbf{a} \mapsto \Psi_{\text{proj},\boldsymbol{\tau}}(\mathbf{a}) - \Psi_{\text{proj},\boldsymbol{\tau}}(\mathbf{0}) = \mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y} - \mathbf{a}'\mathbf{X})] - \mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y})]$ appearing in Definition 2.2 is convex and continuously differentiable on $\mathbb{R}^p$, so that projection regression quantiles can equivalently be defined as the collection of hyperplanes $\pi_{\text{proj},\boldsymbol{\tau}}$ associated with the solutions $\mathbf{a}_{\text{proj},\boldsymbol{\tau}}$ of the system of equations

$$\operatorname{grad}_{\mathbf{a}}\Psi_{\text{proj},\boldsymbol{\tau}}(\mathbf{a}) = -\mathrm{E}[\mathbf{X}(\tau - \mathbb{I}_{[\mathbf{u}'\mathbf{Y}-\mathbf{a}'\mathbf{X}<0]})] = \mathbf{0}. \tag{3.1}$$

Alternatively, recalling the constrained optimization problem in Definition 2.1, one may consider the optimization problem with Lagrangian function $L_{\text{proj},\boldsymbol{\tau}}(\mathbf{a},\mathbf{b},\boldsymbol{\lambda}) := \Psi_\tau(\mathbf{a},\mathbf{b}) - \boldsymbol{\lambda}'(\mathbf{b}-\mathbf{u})$, which yields the gradient conditions

$$(\operatorname{grad}_{(\mathbf{a},\mathbf{b},\boldsymbol{\lambda})} L_{\text{proj},\boldsymbol{\tau}}(\mathbf{a},\mathbf{b},\boldsymbol{\lambda}))_{(\mathbf{a}_{\text{proj},\boldsymbol{\tau}},\mathbf{b}_{\text{proj},\boldsymbol{\tau}},\boldsymbol{\lambda}_{\text{proj},\boldsymbol{\tau}})} = \mathbf{0} \tag{3.2}$$

(the only points in $\mathbb{R}^{p+2m}$ where $(\mathbf{a}, \mathbf{b}, \boldsymbol{\lambda}) \mapsto L_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}, \mathbf{b}, \boldsymbol{\lambda})$ is not continuously differentiable are of the form $(\mathbf{0}', \mathbf{0}', \boldsymbol{\lambda}')'$, and therefore cannot be associated with a minimum of (2.3)). Letting again $\mathbf{Z} := (\mathbf{W}', \mathbf{Y}')'$, (3.2) can be rewritten as

$$\mathbf{0} = (\mathrm{grad}_{\mathbf{a}} L_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}, \mathbf{b}, \boldsymbol{\lambda}))_{(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}}, \mathbf{b}_{\mathrm{proj},\boldsymbol{\tau}}, \boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}})} \tag{3.3a}$$
$$= -\tau \, \mathrm{E}[\mathbf{X}] + \mathrm{E}[\mathbf{X} \, \mathbb{I}_{[\mathbf{Z} \in H^-_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}})]}]$$
$$\mathbf{0} = (\mathrm{grad}_{\mathbf{b}} L_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}, \mathbf{b}, \boldsymbol{\lambda}))_{(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}}, \mathbf{b}_{\mathrm{proj},\boldsymbol{\tau}}, \boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}})} \tag{3.3b}$$
$$= \tau \, \mathrm{E}[\mathbf{Y}] - \mathrm{E}[\mathbf{Y} \, \mathbb{I}_{[\mathbf{Z} \in H^-_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}})]}] - \boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}$$
$$\mathbf{0} = (\mathrm{grad}_{\boldsymbol{\lambda}} L_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}, \mathbf{b}, \boldsymbol{\lambda}))_{(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}}, \mathbf{b}_{\mathrm{proj},\boldsymbol{\tau}}, \boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}})} = -(\mathbf{b}_{\mathrm{proj},\boldsymbol{\tau}} - \mathbf{u}). \tag{3.3c}$$

For such a constrained optimization problem, gradient conditions in general are necessary but not sufficient. In this case, however, sufficiency is clearly achieved since the necessary condition (3.3a) is equivalent to the sufficient one in (3.1) (whereas (3.3b) may be viewed as defining $\boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}$ only). To interpret these gradient conditions, note that (3.3a)-(3.3b) are equivalent to (with $H^-_{\mathrm{proj},\boldsymbol{\tau}} := H^-_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}})$)

$$\mathrm{P}[\mathbf{Z} \in H^-_{\mathrm{proj},\boldsymbol{\tau}}] = \tau \tag{3.4a}$$
$$\frac{1}{1-\tau} \, \mathrm{E}[\mathbf{W} \, \mathbb{I}_{[\mathbf{Z} \in H^+_{\mathrm{proj},\boldsymbol{\tau}}]}] - \frac{1}{\tau} \, \mathrm{E}[\mathbf{W} \, \mathbb{I}_{[\mathbf{Z} \in H^-_{\mathrm{proj},\boldsymbol{\tau}}]}] = \mathbf{0} \tag{3.4b}$$
$$\frac{1}{1-\tau} \, \mathrm{E}[\mathbf{Y} \, \mathbb{I}_{[\mathbf{Z} \in H^+_{\mathrm{proj},\boldsymbol{\tau}}]}] - \frac{1}{\tau} \, \mathrm{E}[\mathbf{Y} \, \mathbb{I}_{[\mathbf{Z} \in H^-_{\mathrm{proj},\boldsymbol{\tau}}]}] = \frac{1}{\tau(1-\tau)} \, \boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}. \tag{3.4c}$$

Clearly, (3.4a) provides projection regression $\boldsymbol{\tau}$-quantiles with a natural probabilistic interpretation, as it keeps the probability of their lower halfspaces equal to $\tau (= \|\boldsymbol{\tau}\|)$. As for (3.4b)-(3.4c), they show (combined with (3.4a)) that the line segment joining the probability mass centers $\frac{1}{\tau} \, \mathrm{E}[\mathbf{Z} \, \mathbb{I}_{[\mathbf{Z} \in H^-_{\mathrm{proj},\boldsymbol{\tau}}]}]$ and $\frac{1}{1-\tau} \, \mathrm{E}[\mathbf{Z} \, \mathbb{I}_{[\mathbf{Z} \in H^+_{\mathrm{proj},\boldsymbol{\tau}}]}]$ of the lower and upper $\boldsymbol{\tau}$-quantile halfspaces is parallel to the vector $(\mathbf{0}', \boldsymbol{\lambda}'_{\mathrm{proj},\boldsymbol{\tau}})'$. In particular, both mass centers share the first $p-1$ coordinates.

The reason why we consider the constrained version of the optimization problem defining projection regression quantiles is that the gradient conditions (3.2) are actually richer than the original ones in (3.1), as the latter do not say anything about $\mathbf{y}$-space projections of these two probability mass centers.

This relative location in the response space (see (3.4c)) actually clarifies the role of the Lagrange multiplier $\boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}$. In general, such a multiplier only measures the impact of the boundary constraint (in this case, the constraint $\mathbf{b}_{\mathrm{proj},\boldsymbol{\tau}} = \mathbf{u}$), but here appears as a functional that is potentially useful for measuring directional outlyingness and tail behavior or for testing (spherical or central) symmetry of the underlying distribution;

see Figure 7 for an illustration. Besides, if we premultiply (3.3a) with $\mathbf{a}'_{\mathrm{proj},\boldsymbol{\tau}}$ and (3.3b) with $\mathbf{b}'_{\mathrm{proj},\boldsymbol{\tau}}$, add both resulting equations and then apply (3.3c), we obtain

$$\Psi_{\mathrm{proj},\boldsymbol{\tau}}(\mathbf{a}_{\mathrm{proj},\boldsymbol{\tau}}) = \mathbf{u}'\boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}, \tag{3.5}$$

so that we can easily extract the minimum achieved in (2.3) from $\boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}$ for any given $\boldsymbol{\tau}$.

If HPŠ quantiles are considered, then the Lagrangian function is $L_{\mathrm{HP\check{S}},\boldsymbol{\tau}}(\mathbf{a},\mathbf{b},\boldsymbol{\lambda}) := \Psi_\tau(\mathbf{a},\mathbf{b}) - \lambda(\mathbf{b}'\mathbf{u}-1)$ and similar arguments as above show that the resulting quantiles can equivalently be defined by

$$\mathrm{P}[\mathbf{Z} \in H^-_{\mathrm{HP\check{S}},\boldsymbol{\tau}}] = \tau \tag{3.6a}$$

$$\frac{1}{1-\tau}\,\mathrm{E}[\mathbf{W}\,\mathbb{I}_{[\mathbf{Z}\in H^+_{\mathrm{HP\check{S}},\boldsymbol{\tau}}]}] - \frac{1}{\tau}\,\mathrm{E}[\mathbf{W}\,\mathbb{I}_{[\mathbf{Z}\in H^-_{\mathrm{HP\check{S}},\boldsymbol{\tau}}]}] = \mathbf{0} \tag{3.6b}$$

$$\frac{1}{1-\tau}\,\mathrm{E}[\mathbf{Y}\,\mathbb{I}_{[\mathbf{Z}\in H^+_{\mathrm{HP\check{S}},\boldsymbol{\tau}}]}] - \frac{1}{\tau}\,\mathrm{E}[\mathbf{Y}\,\mathbb{I}_{[\mathbf{Z}\in H^-_{\mathrm{HP\check{S}},\boldsymbol{\tau}}]}] = \frac{1}{\tau(1-\tau)}\,\lambda_{\mathrm{HP\check{S}},\boldsymbol{\tau}}\mathbf{u}, \tag{3.6c}$$

where we let $H^-_{\mathrm{HP\check{S}},\boldsymbol{\tau}} = H^-_{\mathrm{HP\check{S}},\boldsymbol{\tau}}(\mathbf{a}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}, \mathbf{b}_{\mathrm{HP\check{S}},\boldsymbol{\tau}})$ and $\mathbf{b}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}$ must satisfy the boundary constraint $\mathbf{b}'_{\mathrm{HP\check{S}},\boldsymbol{\tau}}\mathbf{u} = 1$.

These subgradient conditions can clearly be interpreted in the same way as those for projection quantiles, and indicate that both types of quantiles are equally rich. Still, one might argue that projection quantiles are linked in a simpler way to the direction $\mathbf{u}$ in which they are computed, since $\mathbf{u}$ always provides the normal direction to (the $\mathbf{y}$-space projection of) projection $(\tau\mathbf{u})$-quantile hyperplanes whereas the corresponding normal direction for HPŠ $(\tau\mathbf{u})$-quantile hyperplanes is the one bearing the vector $\mathbf{b}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}$ (that depends on $\mathbf{u}$ in a more complicated way). The simple relation between projection quantiles and the corresponding direction $\mathbf{u}$ is, however, just a corollary of the intrinsically univariate nature of projection quantiles.

Let us now turn to the sample case, and let us focus on projection quantiles again. The sample objective function in the projection (i.e., $\mathcal{M} = \mathcal{M}_{\mathrm{proj}}$) version of (2.6) is not continuously differentiable, but still has directional derivatives in all directions, which can be used to formulate fixed-$\mathbf{u}$ subgradient conditions for the sample $\boldsymbol{\tau}$-quantiles. It is easy to check that the coefficients $(\mathbf{a}^{(n)\prime}_{\mathrm{proj},\boldsymbol{\tau}}, \mathbf{b}^{(n)\prime}_{\mathrm{proj},\boldsymbol{\tau}})'$ and the corresponding Lagrange multiplier $\boldsymbol{\lambda}^{(n)}_{\mathrm{proj},\boldsymbol{\tau}}$ of any sample projection regression $\boldsymbol{\tau}$-quantile

$\pi^{(n)}_{\text{proj},\boldsymbol{\tau}} := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1} : \mathbf{u}'\mathbf{y} = \mathbf{a}^{(n)\prime}_{\text{proj},\boldsymbol{\tau}}(1, \mathbf{w}')'\}$ must satisfy

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{[r^{(n)}_{\text{proj},i\boldsymbol{\tau}}<0]} \leq \tau \leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{[r^{(n)}_{\text{proj},i\boldsymbol{\tau}}\leq 0]} \tag{3.7a}$$

$$-\frac{1}{n}\sum_{i=1}^{n} \mathbf{W}_i^{-}\mathbb{I}_{[r^{(n)}_{\text{proj},i\boldsymbol{\tau}}=0]} \leq \tau\left[\frac{1}{n}\sum_{i=1}^{n}\mathbf{W}_i\right] - \left[\frac{1}{n}\sum_{i=1}^{n}\mathbf{W}_i\mathbb{I}_{[r^{(n)}_{\text{proj},i\boldsymbol{\tau}}<0]}\right] \tag{3.7b}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} \mathbf{W}_i^{+}\mathbb{I}_{[r^{(n)}_{\text{proj},i\boldsymbol{\tau}}=0]}$$

$$-\frac{1}{n}\sum_{i=1}^{n} \mathbf{Y}_i^{-}\mathbb{I}_{[r^{(n)}_{\text{proj},i\boldsymbol{\tau}}=0]} \leq \tau\left[\frac{1}{n}\sum_{i=1}^{n}\mathbf{Y}_i\right] - \left[\frac{1}{n}\sum_{i=1}^{n}\mathbf{Y}_i\mathbb{I}_{[r^{(n)}_{\text{proj},i\boldsymbol{\tau}}<0]}\right] - \boldsymbol{\lambda}^{(n)}_{\text{proj},\boldsymbol{\tau}} \tag{3.7c}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} \mathbf{Y}_i^{+}\mathbb{I}_{[r^{(n)}_{\text{proj},i\boldsymbol{\tau}}=0]},$$

where we let $r^{(n)}_{\text{proj},i\boldsymbol{\tau}} := \mathbf{u}'\mathbf{Y}_i - \mathbf{a}^{(n)\prime}_{\text{proj},\boldsymbol{\tau}}(1, \mathbf{W}_i')'$ and write $\mathbf{z}^+ := (\max(z_1, 0), \ldots, \max(z_k, 0))'$ and $\mathbf{z}^- := (-\min(z_1, 0), \ldots, -\min(z_k, 0))'$ for any $\mathbf{z} \in \mathbb{R}^k$. These necessary conditions are obtained by requiring all $2(m + p)$ derivatives of the Lagrangian function in the $\mathbf{a}$ and $\mathbf{b}$ semiaxial directions to be nonnegative. The inequalities in (3.7a), (3.7b), and (3.7c) must be strict if the sample regression $\boldsymbol{\tau}$-quantile is uniquely defined.

Note that (3.7a) indicates that

$$\frac{N}{n} \leq \tau \leq \frac{N+Z}{n}, \quad \text{equivalently} \quad \frac{P}{n} \leq 1 - \tau \leq \frac{P+Z}{n}, \tag{3.8}$$

where N, P, and Z are the numbers of negative, positive, and zero values, respectively, in the residual series $r^{(n)}_{\text{proj},i\boldsymbol{\tau}}$, $i = 1, \ldots, n$. This implies that, for non-integer values of $n\tau$, projection $\boldsymbol{\tau}$-quantile hyperplanes have to contain some of the $\mathbf{Z}_i = (\mathbf{W}_i', \mathbf{Y}_i')'$'s. Actually, if $\mathbf{u}$ is such that the "$\mathbf{u}$-projected" observations $(\mathbf{W}_i', \mathbf{u}'\mathbf{Y}_i)' \in \mathbb{R}^p$ are in general position, then there exists a projection $\boldsymbol{\tau}$-quantile hyperplane $\pi^{(n)}_{\text{proj},\boldsymbol{\tau}}$ which fits exactly p observations, and (3.8) holds with $Z = p$; see Sections 2.2.1 and 2.2.2 of [9]. However, as we will see in the sequel, *the exceptions to this rule, namely the directions* $\mathbf{u}$ *for which degeneracies occur in* $(\mathbf{W}_i', \mathbf{u}'\mathbf{Y}_i)'$, $i = 1, \ldots, n$, *play a crucial role in the computation of quantile regions*. We also stress that necessary subgradient conditions for sample HPŠ quantiles can be derived in the general regression case analogously (the location case was already treated in HPŠ10). In fact, parallel to the population case, these necessary conditions are obtained from (3.7a)-(3.7c) by replacing $\boldsymbol{\lambda}^{(n)}_{\text{proj},\boldsymbol{\tau}}$ with $\lambda^{(n)}_{\text{HPŠ},\boldsymbol{\tau}}\mathbf{u}$ and $r^{(n)}_{\text{proj},i\boldsymbol{\tau}}$ with $r^{(n)}_{\text{HPŠ},i\boldsymbol{\tau}} := \mathbf{b}^{(n)\prime}_{\text{HPŠ},\boldsymbol{\tau}}\mathbf{Y}_i - \mathbf{a}^{(n)\prime}_{\text{HPŠ},\boldsymbol{\tau}}(1, \mathbf{W}_i')'$.

As already mentioned, the sample gradient conditions just discussed are only necessary. The necessary and sufficient ones, for sample projection quantiles or HPŠ quantiles, would directly follow from Theorem 2.1 in [9] thanks to their representation as single-output regression quantiles (see the comments below Definitions 2.2 and 2.3, respectively).

We conclude this section with the remark that all the conditions for sample projection and HPŠ quantiles do not require any finite moments, independence, continuity, or unimodality; actually, they do not require any assumption at all. Still, the number N of negative residuals is always under control and we suggest that this proportion of negative residuals should be preferred to probability of outlyingness measured by $\mathrm{P}[\mathbf{Z} \in H^-_{\mathrm{proj},\boldsymbol{\tau}}]$ and $\mathrm{P}[\mathbf{Z} \in H^-_{\mathrm{HP\check{S}},\boldsymbol{\tau}}]$. In fact, this suggestion is already adopted by various technical norms that prescribe maximum frequency of failures no matter how much these bad cases are intercorrelated. With this said, it is clear that, if the observations $(\mathbf{W}_i', \mathbf{Y}_i')'$, $i = 1, \ldots, n \gg m$, are i.i.d. with a common distribution satisfying Assumption (A), then the standard asymptotic theory can be applied and (3.7a)-(3.7c) may essentially be interpreted as if their population analogs (3.4a)-(3.4c) were almost satisfied, which would imply roughly the same consequences.

3.2. Projection quantiles and portfolio optimization.

A natural field of application for directional quantiles is portfolio optimization. To explain this, assume that the m-dimensional random vector $\mathbf{Y}$ collects m asset returns, and consider the portfolio $\mathbf{Y}_{\boldsymbol{\omega}} := \boldsymbol{\omega}'\mathbf{Y}$, where the vector of portfolio weights $\boldsymbol{\omega} = \omega\mathbf{u}$ (with $\mathbf{u} \in \mathcal{S}^{m-1}$) is a fixed non-zero m-vector.

Portfolio risk behavior can be measured by Value-at-Risk (VaR), tail conditional expectation (TailVaR) or shortfall (s); see Bertsimas et al. [1]. We adopt their definitions but replace the weak inequalities there with strict ones (which clearly makes no difference under Assumption (A)), that is,

$$\mathrm{VaR}_\tau(\boldsymbol{\omega}) := \mathrm{E}[\boldsymbol{\omega}'\mathbf{Y}] - q_\tau(\boldsymbol{\omega}'\mathbf{Y}) \tag{3.9}$$

$$\mathrm{TailVaR}_\tau(\boldsymbol{\omega}) := -\mathrm{E}[\boldsymbol{\omega}'\mathbf{Y}|\boldsymbol{\omega}'\mathbf{Y} < q_\tau(\boldsymbol{\omega}'\mathbf{Y})] \tag{3.10}$$

$$s_\tau(\boldsymbol{\omega}) := \mathrm{E}[\boldsymbol{\omega}'\mathbf{Y}] - \mathrm{E}[\boldsymbol{\omega}'\mathbf{Y}\,|\,\boldsymbol{\omega}'\mathbf{Y} < q_\tau(\boldsymbol{\omega}'\mathbf{Y})], \tag{3.11}$$

where $\tau \in (0,1)$ denotes the level of risk and $q_\tau(.)$ is the same τ-quantile as in (1.1).

Since $\mathrm{E}[(\tau - \mathbb{I}_{[\boldsymbol{\omega}'\mathbf{Y} - q_\tau(\boldsymbol{\omega}'\mathbf{Y})<0]})q_\tau(\boldsymbol{\omega}'\mathbf{Y})] = 0$, we obtain that

$$s_\tau(\boldsymbol{\omega}) = \frac{1}{\tau}\,\mathrm{E}[\rho_\tau(\boldsymbol{\omega}'\mathbf{Y} - q_\tau(\boldsymbol{\omega}'\mathbf{Y}))] = \frac{\omega}{\tau}\,\mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y} - q_\tau(\mathbf{u}'\mathbf{Y}))] = \frac{\omega}{\tau}\,\Psi_{\mathrm{proj},\boldsymbol{\tau}}(a_{\mathrm{proj},\boldsymbol{\tau}}), \tag{3.12}$$

which clearly relates shortfall (and indirectly also TailVaR) to (location) projection quantiles and shows that $\mathrm{E}[\boldsymbol{\omega}'\mathbf{Y}] - \omega a_{\mathrm{proj},\boldsymbol{\tau}}$ equals Value-at-Risk. Note further that (3.5) and (3.12) yield

$$s_\tau(\boldsymbol{\omega}) = \frac{\boldsymbol{\omega}'\boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}}{\tau},$$

which shows that the scaled Lagrange multiplier $\boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}/\tau$ can also be interpreted in this portfolio optimization setup, namely as a vector of individual asset contributions to the overall portfolio risk measured by shortfall.

Clearly, the relation (3.12) between shortfall and projection quantiles has two types of corollaries. First, it allows to infer properties of the projection quantile quantities from the many results on $s_\tau(\boldsymbol{\omega})$ already available in the literature. For instance, it follows from [1] that, if $\mathbf{Y}$ is multinormal with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, then

$$\Psi_{\mathrm{proj},\boldsymbol{\tau}}(a_{\mathrm{proj},\boldsymbol{\tau}}) = \big(\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u}\big)^{1/2}\,\phi(\Phi^{-1}(\tau)),$$

where ϕ and Φ stand for the density and distribution function of the standard normal distribution, respectively. More generally, if the distribution of $\mathbf{Y}$ is elliptically symmetric with mean $\boldsymbol{\mu}$ and finite covariance matrix $\boldsymbol{\Sigma}$, then we have that $\Psi_{\mathrm{proj},\boldsymbol{\tau}}(a_{\mathrm{proj},\boldsymbol{\tau}}) = \big(\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u}\big)^{\frac{1}{2}}\,c(\tau)$, where $c(\tau)$ depends on the specific form of the elliptical distribution and can for instance be obtained from [22]. Other properties state that $\Psi_{\mathrm{proj},\boldsymbol{\tau}}(a_{\mathrm{proj},\boldsymbol{\tau}}) = \tau\mathbf{u}'\mathrm{E}[\mathbf{Y}] - \int_0^\tau a_{t\mathbf{u}}dt$, that both $\Psi_{\mathrm{proj},\boldsymbol{\tau}}(a_{\mathrm{proj},\boldsymbol{\tau}})/\tau$ and $-a_{\mathrm{proj},\boldsymbol{\tau}}$ are non-increasing functions of τ, or that

$$\max\big\{\tau(\mathbf{u}'\mathrm{E}[\mathbf{Y}] - a_{\mathrm{proj},\boldsymbol{\tau}}), (\tau-1)(\mathbf{u}'\mathrm{E}[\mathbf{Y}] - a_{\mathrm{proj},\boldsymbol{\tau}})\big\} \\ \leq \Psi_{\mathrm{proj},\boldsymbol{\tau}}(a_{\mathrm{proj},\boldsymbol{\tau}}) \leq (\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u})^{1/2}(\tau(1-\tau))^{1/2},$$

where $\boldsymbol{\Sigma}$ stands for the finite covariance matrix of $\mathbf{Y}$.

Second, we can bring to shortfall and portfolio optimization all the results regarding projection quantiles, including regression generalization and efficient computation procedures; see Sections 4.1 and 5. Note that both the definition of shortfall and the interpretation of $\boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}$ can indeed be easily generalized to the regression case ($p > 1$) through

$$s_\tau^{\mathrm{regr}}(\boldsymbol{\omega}) := \frac{\omega}{\tau}\min_{\mathbf{a}\in\mathbb{R}^p}\mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y} - \mathbf{a}'\mathbf{X})] = \frac{\boldsymbol{\omega}'\boldsymbol{\lambda}_{\mathrm{proj},\boldsymbol{\tau}}}{\tau}, \qquad \boldsymbol{\omega} = \omega\mathbf{u}.$$

Clearly, TailVaR and VaR (as well as some of the results given above) can also be extended in the same spirit to this generalized regression context. Since the portfolio return and risk are likely to depend on many economic factors (or other covariates), represented by $\mathbf{X}$ here, this extension to the regression setup makes good sense and might have priceless consequences in finance and related areas.

Although we focused above on projection quantiles, it is important to stress that such portfolio interpretation can similarly be derived for HPŠ quantiles if one restricts (as is natural) to weights $\boldsymbol{\omega}$ that are optimal among those satisfying $\boldsymbol{\omega}'\mathbf{u} = \omega$ for given $\mathbf{u}$ and $\omega > 0$. *The important advantage of HPŠ quantiles here is their ability to find easily such optimal portfolio weights* (we simply have $\boldsymbol{\omega}_{\text{opt}} = \omega \mathbf{b}_{\text{HPŠ},\tau\mathbf{u}}$); in comparison, projection quantiles do not offer any possibility to optimize weights without considering them all. Besides, all formulae already derived for projection quantiles can be translated into formulae for HPŠ quantiles by considering the former in direction $\mathbf{u}_1 := \mathbf{b}_{\text{HPŠ},\tau\mathbf{u}}/\|\mathbf{b}_{\text{HPŠ},\tau\mathbf{u}}\|$, with $a_{\text{proj},\tau\mathbf{u}_1} = a_{\text{HPŠ},\tau\mathbf{u}}/\|\mathbf{b}_{\text{HPŠ},\tau\mathbf{u}}\|$ and $\boldsymbol{\lambda}_{\text{proj},\tau\mathbf{u}_1} = \lambda_{\text{HPŠ},\tau\mathbf{u}}\mathbf{u}$ (and by further substituting all projection quantities with HPŠ ones).

4. Quantile regions.

In this section, we first focus on the location case and define there the quantile regions associated with the directional quantiles from Definition 2.1. As explained in the Introduction, the projection quantile regions coincide with the halfspace depth ones, but this identification unfortunately does not provide any way to compute the latter exactly by means of projection quantiles. However, as we show here, another proof of this identification leads to exact computation of these regions. Most importantly, we also consider quantile regions in the general regression case and show that the sample quantile regions defined in HPŠ10 can also be obtained exactly from projection regression quantiles.

4.1. The location case.

In the location case, τ-quantile regions, for any fixed $\tau(=\|\boldsymbol{\tau}\|) \in (0,1)$, can be obtained by taking the "upper envelope" of the corresponding ($\tau\mathbf{u}$)-quantile hyperplanes $\pi_{\mathcal{M},\tau\mathbf{u}}$ from Definition 2.1. More precisely, we define the $\mathcal{M}$-type τ-quantile region as

$$R_{\mathcal{M}}(\tau) := \textstyle\bigcap_{\mathbf{u}\in\mathcal{S}^{m-1}} \bigcap \{H^+_{\mathcal{M},\tau\mathbf{u}}\}, \qquad (4.1)$$

where $\bigcap\{H^+_{\mathcal{M},\tau\mathbf{u}}\}$ stands for the intersection of the collection $\{H^+_{\mathcal{M},\tau\mathbf{u}}\}$ of all (closed) upper ($\tau\mathbf{u}$)-quantile halfspaces associated with the minimum in (2.1). If a sample of data points $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ is available, empirical versions $R^{(n)}(\tau)$ result from (4.1) by replacing the population quantile halfspaces $H^+_{\mathcal{M},\tau\mathbf{u}}$ with their sample counterparts $H^{(n)+}_{\mathcal{M},\tau\mathbf{u}}$ (with the intersection over all the halfspaces associated with the minimimum in (2.6)). The corresponding τ-quantile contours are naturally defined as the boundaries $\partial R_{\mathcal{M}}(\tau)$ and $\partial R^{(n)}_{\mathcal{M}}(\tau)$ of $R_{\mathcal{M}}(\tau)$ and $R^{(n)}_{\mathcal{M}}(\tau)$, respectively.

These τ-quantile regions are closed and convex since they are obtained by intersecting closed halfspaces. For a general $\mathcal{M}$, there is no guarantee that they are nested (in the sense that $R_{\mathcal{M}}(\tau_1) \subset R_{\mathcal{M}}(\tau_2)$ and $R^{(n)}_{\mathcal{M}}(\tau_1) \subset R^{(n)}_{\mathcal{M}}(\tau_2)$ if $\tau_1 \geq \tau_2$). Of course, one can always obtain nested regions by defining the intersection quantile regions $R^{\cap}_{\mathcal{M}}(\tau) := \bigcap_{t\in(0,\tau]} R_{\mathcal{M}}(t)$ and $R^{\cap(n)}_{\mathcal{M}}(\tau) := \bigcap_{t\in(0,\tau]} R^{(n)}_{\mathcal{M}}(t)$. But there is no need to consider such intersection regions for the projection and HPŠ regions $R_{\text{proj/HPŠ}}(\tau)$ and $R^{(n)}_{\text{proj/HPŠ}}(\tau)$ (with obvious notation), since these are almost surely nested under Assumption (A), which readily follows from their strong connection with halfspace depth regions in Theorem 4.1 below.

To state this theorem, let us recall that the order-τ halfspace depth region associated with the probability distribution P is defined as $D(\tau) = D(\tau, \mathrm{P}) := \{\mathbf{y} \in \mathbb{R}^m : HD(\mathbf{y}, \mathrm{P}) \geq \tau\}$, where

$$HD(\mathbf{y}) = HD(\mathbf{y}, \mathrm{P}) := \inf\{\mathrm{P}[H] : H \text{ is a closed halfspace containing } \mathbf{y}\} \tag{4.2}$$

is the halfspace depth of $\mathbf{y}$ with respect to P. It can be shown that

$$D(\tau) = \bigcap \{H : H \text{ is a closed halfspace with } \mathrm{P}[H] > 1 - \tau\}, \tag{4.3}$$

for any $\tau > 0$; see Proposition 6 in [21]. Sample versions of $HD(\mathbf{y})$ and $D(\tau)$ are simply given by $HD^{(n)}(\mathbf{y}) := HD(\mathbf{y}, \mathrm{P}_n)$ and $D^{(n)}(\tau) := D(\tau, \mathrm{P}_n)$, respectively, where P_n stands for the empirical measure associated with the observed n-tuple $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ at hand. Clearly, there are at most n compact sample halfspace depth regions

$$D^{(n)}(\tfrac{\ell}{n}) = \bigcap \{H : H \text{ is a closed halfspace with } n\mathrm{P}_n[H] \geq n - \ell + 1\}, \quad \ell = 1, \ldots, n; \tag{4.4}$$

see (4.3). We then have the following result.

Theorem 4.1. *(i) Under Assumption (A), $R_{\text{proj}}(\tau) = D(\tau) = R_{\text{HPŠ}}(\tau)$ for all $\tau \in (0, 1)$. (ii) Assume that the data points $\mathbf{Y}_i$, $i = 1, \ldots, n(\geq m + 1)$, are in general position. Then, for any $\ell \in \{1, 2, \ldots, n - m\}$ such that $D^{(n)}(\frac{\ell}{n})$ has a non-empty interior, we have that $R^{(n)}_{\text{proj}}(\tau) = D^{(n)}(\frac{\ell}{n}) = R^{(n)}_{\text{HPŠ}}(\tau)$ for any positive τ in $[\frac{\ell-1}{n}, \frac{\ell}{n})$.*

The equality between HPŠ regions and halfspace depth regions was obtained in Theorems 4.1 and 4.2 of HPŠ10, whereas the corresponding result for projection regions was proved in Theorem 3 of KM08. A direct corollary of Theorem 4.1 is that the regions $R_{\text{proj/HPŠ}}(\tau)$ and $R^{(n)}_{\text{proj/HPŠ}}(\tau)$ are affine-equivariant, nested, and compact; see [21]. They are also always convex; we refer to [8] for an extension of halfspace depth regions yielding possibly non-convex shapes.

When it comes to computing the sample halfspace depth regions $D^{(n)}(\tau)$ for some fixed τ on the basis of Theorem 4.1(ii), it appears that the link of $D^{(n)}(\tau)$ to HPŠ quantiles is much more exploitable than that to projection quantiles, as is argued in HPŠ10. Contrary to the strictly finite collection $\{\pi^{(n)}_{\text{HPŠ},\tau\mathbf{u}} : \mathbf{u} \in \mathcal{S}^{m-1}\}$, its projection counterpart $\{\pi^{(n)}_{\text{proj},\tau\mathbf{u}} : \mathbf{u} \in \mathcal{S}^{m-1}\}$ indeed contains uncountably many hyperplanes (one for each $\mathbf{u}$). Therefore the intersection defining $R^{(n)}_{\text{proj}}(\tau)$ runs over an infinite number of upper halfspaces $H^{(n)+}_{\text{proj},\tau\mathbf{u}}$, which seems impossible to compute in practice. Consequently, it is proposed in KM08 to sample the unit sphere $\mathcal{S}^{m-1}$, a strategy that can clearly lead to approximate regions $D^{(n)}(\tau)$ only.

On the other hand, the "directional" decomposition of polyhedral halfspace depth regions $D^{(n)}(\tau)$ into projection upper quantile halfspaces nicely provides their faces with a neat and interesting quantile interpretation. Indeed, each face of $D^{(n)}(\tau)$ is included in the projection quantile hyperplane $\pi^{(n)}_{\text{proj},\tau\mathbf{u}_0}$, where $\mathbf{u}_0$ stands for the unit vector orthogonal to that face and pointing to the interior of $D^{(n)}(\tau)$.

To sum up, projection quantiles are helpful for the quantile interpretation of the halfspace depth regions $D^{(n)}(\tau)$, but appear less useful for their exact computation than their HPŠ counterparts. However, as we show below, an alternative proof of the identity $R^{(n)}_{\text{proj}}(\tau) = D^{(n)}(\tau)$ reveals that $D^{(n)}(\tau)$ can also be computed efficiently from projection quantiles. First, we need a couple of preliminary lemmas, proved in the Appendix.

Lemma 4.1. *Assume that the observations are in general position and fix* $\ell \in \{1, 2, \ldots, n-m\}$ *such that* $D^{(n)}(\frac{\ell}{n})$ *has a nonempty interior. Then* $D^{(n)}(\frac{\ell}{n}) = \bigcap \mathcal{H}^{(n)}_{\ell-1}$, *where* $\mathcal{H}^{(n)}_k := \{H : H$ *is a closed halfspace with* $n\mathrm{P}_n[\partial H] = m$ *and* $n\mathrm{P}_n[H^c] = k\}$.

To state the second lemma, we define the collection of τ-*critical directions* as

$$C^{(n)}_\tau := \{\mathbf{u} \in \mathcal{S}^{m-1} : \text{there exists at least one } H^{(n)+}_{\text{proj},\tau\mathbf{u}} \text{ with } n\mathrm{P}_n[\partial H^{(n)+}_{\text{proj},\tau\mathbf{u}}] = m\}.$$

If τ is such that $n\tau$ is not an integer, then the projection $(\tau\mathbf{u})$-quantile hyperplane is unique for each direction $\mathbf{u}$ and thus $C^{(n)}_\tau = \{\mathbf{u} \in \mathcal{S}^{m-1} : n\mathrm{P}_n[\partial H^{(n)+}_{\text{proj},\tau\mathbf{u}}] = m\}$. For such a value of τ, most directions $\mathbf{u}$ yield a halfspace $H^{(n)+}_{\text{proj},\tau\mathbf{u}}$ with exactly one data point on its boundary.

Lemma 4.2. *Fix $\ell \in \{1, 2, \ldots, n-m\}$ and $\tau \in [\frac{\ell-1}{n}, \frac{\ell}{n})$. Then*

$$\mathcal{H}^{(n)}_{\ell-j} = \bigcup_{\mathbf{u}\in C^{(n)}_\tau} \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m \text{ and } n\mathrm{P}_n[(H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}})^c] = \ell - j\} \quad (4.5)$$

for any $j = 1, 2, \ldots, \min(m + \delta_{n\tau,\ell-1}, \ell)$, where $\delta_{r,s}$ is equal to one if $r = s$ and zero otherwise.

Note that the constraint $n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m$ can be removed from the right-hand side of (4.5) without any loss of generality if $\tau \in (\frac{\ell-1}{n}, \frac{\ell}{n})$, since projection quantile hyperplanes (hence also upper quantile halfspaces) are uniquely defined for such values of τ and since only critical directions are considered. We then have the following result, which is proved in the Appendix.

Theorem 4.2. *Assume that the data points $\mathbf{Y}_i$, $i = 1, \ldots, n(\geq m+1)$, are in general position and fix $\ell \in \{1, 2, \ldots, n-m\}$ such that $D^{(n)}(\frac{\ell}{n})$ has a nonempty interior. Then, for any positive $\tau \in [\frac{\ell-1}{n}, \frac{\ell}{n})$, we have that (i) $R^{(n)}_{\mathrm{proj}}(\tau) = D^{(n)}(\frac{\ell}{n})$ and (ii) $D^{(n)}(\frac{\ell}{n}) = \bigcap_{\mathbf{u}\in C^{(n)}_\tau} \bigcap \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m\}$.*

Assume that the conditions of this theorem are fulfilled for some $\tau \in (\frac{\ell-1}{n}, \frac{\ell}{n})$. While Part (i) of Theorem 4.2 simply restates the projection result of Theorem 4.1(ii), Part (ii) has a number of crucial implications for the computation of halfspace depth regions. Clearly, it implies that the problem of computing $R^{(n)}_{\mathrm{proj}}(\tau) = D^{(n)}(\frac{\ell}{n})$ reduces to that of determining the set of τ-critical directions $C^{(n)}_\tau$. But not only that: applying successively Lemmas 4.1 and 4.2 leads to

$$\begin{aligned} D^{(n)}(\tfrac{\ell+1-j}{n}) &= \bigcap \mathcal{H}^{(n)}_{\ell-j} \\ &= \bigcap \bigcup_{\mathbf{u}\in C^{(n)}_\tau} \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m \text{ and } n\mathrm{P}_n[(H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}})^c] = \ell - j\} \\ &= \bigcap \bigcup_{\mathbf{u}\in C^{(n)}_\tau} \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[(H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}})^c] = \ell - j\} \end{aligned}$$

for any $j = 1, \ldots, \min(m, \ell)$, where the last equality follows from the remark below Lemma 4.2. In other words, $D^{(n)}(\frac{\ell+1-j}{n})$ can be obtained for such j's by intersecting all the upper quantile halfspaces $H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}$ that are associated with τ-critical directions and cut off exactly $\ell - j$ observations. This means that *one* $C^{(n)}_\tau$ allows to compute simultaneously *several* (typically m) halfspace depth regions in an exact way.

The methodology can be used in practice because it is possible to determine $C^{(n)}_\tau$ efficiently by means of parametric programming. Clearly, critical directions are among those directional vectors $\mathbf{u}$ where the optimal basis of the associated linear program can change with $\mathbf{u}$ in the maximum number of ways (that is to say that each of the exactly fitted observations may leave the corresponding projection ($\tau\mathbf{u}$)-quantile hyperplane with a tiny change of $\mathbf{u}$ in a suitable direction). We make this more precise in Section 5.2 below.

Of course, one can obtain Theorem 4.2 more directly from Theorem 4.1(ii). Yet our derivation allows to derive this result on projection quantiles by using projection quantiles only (whereas the proof based on Theorem 4.1(ii) requires considering HPŠ quantiles, too). We stress that, contrary to the location case, no result available in the literature—to the best of our knowledge—would allow to establish easily the link between HPŠ regression quantile regions and regression projection quantiles we provide below.

4.2. The general regression case.

In the regression setup $p \geq 2$, quantile regions $R_{\mathcal{M}}(\tau)/R^{(n)}_{\mathcal{M}}(\tau)$, parallel to the location case, can be defined through

$$R_{\mathcal{M}}(\tau) := \textstyle\bigcap_{\mathbf{u}\in\mathcal{S}^{m-1}} \bigcap \{H^{+}_{\mathcal{M},\tau\mathbf{u}}\} \tag{4.6}$$

and

$$R^{(n)}_{\mathcal{M}}(\tau) := \textstyle\bigcap_{\mathbf{u}\in\mathcal{S}^{m-1}} \bigcap \{H^{(n)+}_{\mathcal{M},\tau\mathbf{u}}\}, \tag{4.7}$$

where the second intersection in (4.6) (resp., (4.7)) is over all upper ($\tau\mathbf{u}$)-quantile halfspaces associated with the minimum in (2.1) (resp., (2.6)), and this still produces regions that are connected and convex. As in the previous section, we are mainly interested in the projection and HPŠ regions $R_{\text{proj}}(\tau)/R^{(n)}_{\text{proj}}(\tau)$ and $R_{\text{HPŠ}}(\tau)/R^{(n)}_{\text{HPŠ}}(\tau)$, associated with $\mathcal{M}_{\text{proj}}$ and $\mathcal{M}_{\text{HPŠ}}$, respectively.

4.2.1. Identification of projection and HPŠ regression quantile regions.

Since Theorem 4.1 shows that projection and HPŠ quantile regions coincide in the location case, a natural question is whether this extends to the regression case or not. As shown by the following result, the answer is positive for population regions.

Theorem 4.3. *Consider the regression setup with $p \geq 2$. Then, under Assumption (A), $R_{\text{proj}}(\tau) = R_{\text{HPŠ}}(\tau)$ for all $\tau \in (0,1)$.*

The proof of this result (see the Appendix) is based on the fact that the subgradient conditions (3.4a)-(3.4c) and (3.6a)-(3.6c) directly reveal that (i) any projection regression quantile halfspace $H^{+}_{\text{proj},\tau\mathbf{u}}$ is also a HPŠ regression quantile halfspace $H^{+}_{\text{HPŠ},\tau\mathbf{v}}$ for a possibly different direction $\mathbf{v} \in \mathcal{S}^{m-1}$ (which establishes that $R_{\text{HPŠ}}(\tau) \subset R_{\text{proj}}(\tau)$) and that, conversely, (ii) any HPŠ regression quantile halfspace $H^{+}_{\text{HPŠ},\tau\mathbf{u}}$ is also a projection regression quantile halfspace $H^{+}_{\text{proj},\tau\mathbf{v}}$, still for a possibly different $\mathbf{v} \in \mathcal{S}^{m-1}$ (which establishes that $R_{\text{proj}}(\tau) \subset R_{\text{HPŠ}}(\tau)$).

Now, translating this into the sample case turns out to be unexpectedly complicated, as it is by no means obvious that a projection regression quantile halfspace $H^{(n)+} = H^{(n)+}_{\text{proj},\tau\mathbf{u}}$ *whose boundary hyperplane contains* $m+p-1$ *data points* (this restriction is needed since HPŠ regression quantile hyperplanes typically interpolate $m+p-1$ data points) can always be identified with some sample HPŠ regression quantile halfspace $H^{(n)+}_{\text{HPŠ},\tau\mathbf{v}}$: while the boundary hyperplane of $H^{(n)+}$ is clearly a Koenker and Bassett regression quantile hyperplane when the "vertical direction" of this single-output regression is along the vector $(\mathbf{v}'_{\mathbf{w}}, \mathbf{v}'_{\mathbf{y}})' \in \mathcal{S}^{m+p-1}$ linking the probability mass centers of $H^{(n)+}$ and $\mathbb{R}^{m+p-1} \setminus H^{(n)+}$ (see—the sample version of—(3.6) in HPŠ10), this direction in general does not belong to the response space of the considered multiple-output regression problem (i.e., is not of the form $(\mathbf{0}'_{p-1}, \mathbf{v}'_{\mathbf{y}})'$), hence is not admissible for a HPŠ regression quantile hyperplane. Showing that there always exists an admissible direction in which the Koenker and Bassett quantile upper halfspace remains $H^{(n)+}$ is extremely delicate yet possible, and is the most important step to establish the following result.

Theorem 4.4. *Consider the regression setup with* $p \geq 2$*, and assume that the data points* $(\mathbf{W}'_i, \mathbf{Y}'_i)'$*,* $i = 1, \ldots, n (\geq m+p)$*, are in general position. Then, for all* $\tau \in (0,1)$*,*

$$\begin{aligned} R^{(n)\text{prac}}_{\text{HPŠ}}(\tau) &:= \textstyle\bigcap_{\mathbf{u}\in\mathcal{S}^{m-1}} \bigcap \{H^{(n)+}_{\text{HPŠ},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\text{HPŠ},\tau\mathbf{u}}] = m+p-1\} \qquad (4.8) \\ &= \textstyle\bigcap_{\mathbf{u}\in C^{(n)}_{\tau}} \bigcap \{H^{(n)+}_{\text{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\text{proj},\tau\mathbf{u}}] = m+p-1\} =: R^{(n)\text{crit}}_{\text{proj}}(\tau), \end{aligned}$$

where $C^{(n)}_{\tau}$ *denotes the collection of all* τ-critical directions*, that is, the collection of directions* $\mathbf{u} \in \mathbb{R}^m$ *for which there exists a projection* $(\tau\mathbf{u})$*-quantile hyperplane* $\pi^{(n)}_{\text{proj},\tau\mathbf{u}}$ *fitting* $m+p-1$ *observations (and not only* p *as in most directions* $\mathbf{u}$*).*

Note that any direction $\mathbf{u} \in \mathcal{S}^{m-1}$ gives raise (for any $\tau \in (0,1)$) to (at least) one HPŠ quantile hyperplane $\pi^{(n)}_{\text{HPŠ},\tau\mathbf{u}}$ containing $m+p-1$ data points, which explains that the regions $R^{(n)\text{prac}}_{\text{HPŠ}}(\tau)$ in (4.8) may be non-empty. Theorem 4.4 then shows that, if, for some fixed τ, all HPŠ $(\tau\mathbf{u})$-quantiles ($\mathbf{u} \in \mathcal{S}^{m-1}$) are uniquely defined, then the resulting HPŠ regression quantile regions $R^{(n)}_{\text{HPŠ}}(\tau) = R^{(n)\text{prac}}_{\text{HPŠ}}(\tau)$ and the "critical" projection regions $R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau)$ do coincide. This identification in the empirical case is, as explained above, highly non-trivial, and it allows to compute HPŠ regression quantile regions from projection quantiles, which may be computationally advantageous if m is very small ($m = 2$); we refer to the companion paper [18] for extensive details.

The HPŠ quantile regions $R^{(n)}_{\mathrm{HP\check{S}}}(\tau)$ can be computed in practice only when all HPŠ ($\tau\mathbf{u}$)-quantiles ($\mathbf{u} \in \mathcal{S}^{m-1}$) are uniquely defined, and we then have $R^{(n)}_{\mathrm{HP\check{S}}}(\tau) = R^{(n)\mathrm{prac}}_{\mathrm{HP\check{S}}}(\tau)$. This explains the notation $R^{(n)\mathrm{prac}}_{\mathrm{HP\check{S}}}(\tau)$ and shows that the fact that Theorem 4.4 involves $R^{(n)\mathrm{prac}}_{\mathrm{HP\check{S}}}(\tau)$—and not $R^{(n)}_{\mathrm{HP\check{S}}}(\tau)$—is not a limitation from the computational point of view.

For $p \geq 2$, there is still an open question whether $R^{(n)}_{\mathrm{proj}}(\tau) = R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau)$ or not. Of course, it is always possible—at least for (very) small dimensions m—to compute the regions $R^{(n)}_{\mathrm{proj}}(\tau)$ approximately by sampling the unit sphere $\mathcal{S}^{m-1}$, and experiments of this type lead us to conjecture that, under the same conditions as in Theorem 4.4, $R^{(n)}_{\mathrm{proj}}(\tau) = R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau)$. As long as this remains a conjecture, however, the original projection regions $R^{(n)}_{\mathrm{proj}}(\tau)$ cannot be computed exactly for $p \geq 2$, which makes the HPŠ approach superior in this respect.

4.2.2. Towards a point regression depth.

Note that, due to the *quantile crossing phenomenon*, projection and HPŠ regression quantile regions need not have the same nesting property as in the location case, which is especially apparent in the single-output setup known from the standard quantile regression theory. This implies that we must turn to projection and HPŠ intersection (regression) quantile regions $R^{\cap}_{\mathrm{proj/HP\check{S}}}(\tau) := \bigcap_{t\in(0,\tau]} R_{\mathrm{proj/HP\check{S}}}(t)$ and $R^{\cap(n)}_{\mathrm{proj/HP\check{S}}}(\tau) := \bigcap_{t\in(0,\tau]} R^{(n)}_{\mathrm{proj/HP\check{S}}}(t)$, if nestedness is required.

The regions $R^{\cap}_{\mathrm{proj/HP\check{S}}}(\tau)$ and $R^{\cap(n)}_{\mathrm{proj/HP\check{S}}}(\tau)$ are nested, connected, and convex, and therefore implicitly define a (population and sample, respectively) regression depth measure via

$$RD^{\cap}_{\mathrm{proj/HP\check{S}}}(\mathbf{w},\mathbf{y}) := \sup\left\{\tau \in \big(0,1\big) : (\mathbf{w}',\mathbf{y}')' \in R^{\cap}_{\mathrm{proj/HP\check{S}}}(\tau)\right\} \tag{4.9}$$

and

$$RD^{\cap(n)}_{\mathrm{proj/HP\check{S}}}(\mathbf{w},\mathbf{y}) := \sup\left\{\tau \in \big(0,1\big) : (\mathbf{w}',\mathbf{y}')' \in R^{\cap(n)}_{\mathrm{proj/HP\check{S}}}(\tau)\right\}, \tag{4.10}$$

with $\sup\emptyset := 0$. In view of Theorem 4.1, this regression depth naturally extends the halfspace depth in (4.2) to the regression context, and this construction clearly makes it possible to define a concept of regression depth in all settings where quantile regression works. Note that (4.9) and (4.10) define the depth of a *point* of the regression space, and not the depth of a *regression hyperplane* as is the classical *regression depth* of [20]. To stress the difference, we will use the term *point regression halfspace depth* for $RD^{\cap}_{\mathrm{proj/HP\check{S}}}(\mathbf{w},\mathbf{y})$ and $RD^{\cap(n)}_{\mathrm{proj/HP\check{S}}}(\mathbf{w},\mathbf{y})$.

The regions $R^{\cap}_{\mathrm{proj/HP\check{S}}}(\tau)$ and $R^{\cap(n)}_{\mathrm{proj/HP\check{S}}}(\tau)$ are important especially from the theoretical point of view because of the induced depth measures in (4.9)-(4.10). Unfortunately, the intersection defining $R^{\cap(n)}_{\mathrm{proj/HP\check{S}}}(\tau)$ is virtually impossible to compute—unless, of course, for $p = 1$ or $m = 1$. However, we might base our sample depth measure from (4.10) on the $R^{(n)}_{\mathrm{proj/HP\check{S}}}(\tau)$'s and define a *point regression halfspace pseudo-depth* $RD^{(n)}_{\mathrm{proj/HP\check{S}}}(\mathbf{w},\mathbf{y})$, say. The term *pseudo-depth* stresses the possible lack of monotonicity of this depth measure, which is the penalty for the possible non-nestedness of the $R^{(n)}_{\mathrm{proj/HP\check{S}}}(\tau)$'s. Note that both point regression halfspace depth and pseudo-depth reduce to standard halfspace depth in the location case, and hence can be regarded as extensions of the latter to the regression setup. Also, the regression regions $R^{(n)}_{\mathrm{proj/HP\check{S}}}(\tau)$ and $R^{\cap(n)}_{\mathrm{proj/HP\check{S}}}(\tau)$ must contain the location halfspace depth regions computed from the data points $(\mathbf{W}'_i,\mathbf{Y}'_i)'$, $i = 1,\ldots,n$, and hence are non-empty for $\tau \le 1/(m+p)$; see Proposition 9 in [21].

5. Computational aspects for projection quantiles.

Here we discuss computation of the projection (regression) quantiles and projection (regression) quantile regions introduced in the previous sections. First, we only briefly comment on the evaluation of fixed-$\mathbf{u}$ projection regression quantiles since they are of a standard single-output regression nature. Then we describe how parametric programming allows for computing quantile contours extremely efficiently. Finally, we define and interpret projection regression rank scores.

5.1. *Fixed-$\mathbf{u}$ projection regression quantiles.*

Let $(1,\mathbf{W}'_i,\mathbf{Y}'_i)' = (\mathbf{X}'_i,\mathbf{Y}'_i)'$, $i = 1,\ldots,n$, be n observations from the regression model considered in the previous sections. Define

$$\mathbb{U}^y_{n\times 2m} := (\mathbf{Y}^c_1, -\mathbf{Y}^c_1, \ldots, \mathbf{Y}^c_m, -\mathbf{Y}^c_m) \quad\text{and}\quad \mathbb{V}^x_{n\times 2p} := (\mathbf{X}^c_1, -\mathbf{X}^c_1, \ldots, \mathbf{X}^c_p, -\mathbf{X}^c_p),$$

where $\mathbf{Y}^c_j$, $j = 1,\ldots,m$, and $\mathbf{X}^c_j$, $j = 1,\ldots,p$, stand for the jth column of the response data matrix $\mathbb{Y} := (\mathbf{Y}_1,\ldots,\mathbf{Y}_n)'$ and design data matrix $\mathbb{X} := (\mathbf{X}_1,\ldots,\mathbf{X}_n)'$, respectively. For any $\tau \in (0,1)$ and any $\mathbf{u} \in \mathcal{S}^{m-1}$, the resulting sample projection regression $(\tau\mathbf{u})$-quantile then results from the linear programming problem (P) given by

$$\min_{\mathbf{z}} \mathbf{c}_{\mathrm{P}}'\mathbf{z} \quad \text{subject to} \quad \mathbb{A}_{\mathrm{P}}\mathbf{z} = \mathbf{b}_{\mathrm{P}}, \quad \mathbf{z} \geq \mathbf{0},$$

with

$$\mathbf{z} = (b_{1+}, b_{1-}, \ldots, b_{m+}, b_{m-}, a_{1+}, a_{1-}, \ldots, a_{p+}, a_{p-}, \mathbf{r}_+', \mathbf{r}_-')' \in \mathbb{R}^{2m+2p+2n},$$

$$\mathbf{c}_{\mathrm{P}} = (\mathbf{0}_{2m+2p}', \tau\mathbf{1}_n', (1-\tau)\mathbf{1}_n')' \in \mathbb{R}^{2m+2p+2n}, \quad \mathbf{b}_{\mathrm{P}} = (\mathbf{u}', \mathbf{0}_n')' \in \mathbb{R}^{m+n},$$

and

$$\mathbb{A}_{\mathrm{P}} = \begin{pmatrix} \mathbb{A}^1_{\mathrm{P}(m\times(2m+2p+2n))} \\ \mathbb{A}^2_{\mathrm{P}(n\times(2m+2p+2n))} \end{pmatrix} = \begin{pmatrix} \mathbb{M}_{m\times 2m} & \mathbb{O}_{m\times 2p} & \mathbb{O}_{m\times n} & \mathbb{O}_{m\times n} \\ \mathbb{U}^y_{n\times 2m} & -\mathbb{V}^x_{n\times 2p} & -\mathbb{I}_{n\times n} & \mathbb{I}_{n\times n} \end{pmatrix},$$

where $\mathbf{1}_\ell \in (1,\ldots,1)' \in \mathbb{R}^\ell$, $\mathbf{0}_\ell \in (0,\ldots,0)' \in \mathbb{R}^\ell$, and $\mathbb{M} = (m_{i,j})$ is the $(m \times 2m)$ matrix defined by $m_{i,2i-1} = 1$, $m_{i,2i} = -1$, and $m_{i,j} = 0$ otherwise. The dual twin brother (D) of (P) is of the form

$$\max_{\boldsymbol{\mu}_{\mathrm{P}} = (\boldsymbol{\mu}^{b\prime}, \boldsymbol{\mu}_{\mathrm{P}}^{r\,\prime})'} \mathbf{u}'\boldsymbol{\mu}^b$$

subject to

$$\boldsymbol{\mu}^b = -\mathbb{Y}'\boldsymbol{\mu}^r_{\mathrm{P}}, \quad \mathbb{X}'\boldsymbol{\mu}^r_{\mathrm{P}} = \mathbf{0}_p, \quad \text{and} \quad -\tau\mathbf{1}_n \leq \boldsymbol{\mu}^r_{\mathrm{P}} \leq (1-\tau)\mathbf{1}_n.$$

Both (P) and (D) have at least one feasible solution, and hence also an optimal one. Although (P) may have more distinct optimal solutions, we need not be too worried about that under Assumption (A), since the asymptotic theory for single-output sample regression quantiles then ensures that any sequence of such solutions converges almost surely to the unique population regression τ-quantile as $n \to \infty$. Besides, $\mathbb{A}_{\mathrm{P}}$ has full rank $n+m$ and therefore each optimal solution to (P) can be expressed as a linear combination of basic solutions that have at most $n+m$ positive coordinates.

The optimal Lagrange multiplier vector $\boldsymbol{\mu}_{\mathrm{P}}^{(n)}$ corresponds to the equality constraints from (P). Therefore,

$$\boldsymbol{\mu}^{b(n)} = n\,\boldsymbol{\lambda}^{(n)}_{\mathrm{proj},\tau\mathbf{u}}.$$

The Strong Duality Theorem then guarantees that

$$\Psi^{(n)}_{\mathrm{proj},\tau\mathbf{u}}(\mathbf{a}^{(n)}_{\mathrm{proj},\tau\mathbf{u}}) = \mathbf{u}'\boldsymbol{\lambda}^{(n)}_{\mathrm{proj},\tau\mathbf{u}},$$

which generalizes (3.5) to the sample case.

For any fixed $\mathbf{u}$, one can compute the sample projection regression quantiles with the aid of standard quantile regression solvers. In particular, there is an excellent package for advanced quantile regression analysis in R (see [10]) and the key function for computing quantile regression estimates is also available for MATLAB from Roger Koenker's homepage.

5.2. Projection regression quantile contours.

The previous subsection shows that fixed-$\mathbf{u}$ computation of projection regression quantiles is essentially straightforward. The real challenge is to solve (P) efficiently *for all directions* $\mathbf{u} \in \mathcal{S}^{m-1}$ and for any given $\tau \in (0,1) \setminus \{\frac{1}{n}, \ldots, \frac{n-1}{n}\}$, say. As mentioned in Section 4, we propose a solution that relies on *parametric programming.*

Under Assumption (A), it turns out that the whole space $\mathbb{R}^m$ may almost surely be segmented into a finite number of non-degenerate cones $\mathcal{C}_q$, $q = 1, 2, \ldots, N_C$, in such a way that

(i) the signs of all coordinates of $\mathbf{a}^{(n)}_{\text{proj},\tau\mathbf{u}}$, $\mathbf{u}$, and $\mathbf{r}^{(n)}_{(\tau\mathbf{u})} = \mathbf{r}^{(n)}_{+} - \mathbf{r}^{(n)}_{-}$ are constant in the interior of any $\mathcal{C}_q$, and that

(ii) there exists a p-element row index set $h_q := \{k_1, \ldots, k_p\} \subset \{1, \ldots, n\}$ such that it holds, for any $\mathbf{u} \in \mathcal{C}_q$, $\mathbb{A}(h_q) := (\mathbb{X}(h_q))^{-1}\mathbb{Y}(h_q)$, $f^{h_q}_{\boldsymbol{\tau}}(\mathbf{X}_i, \mathbf{Y}_i) := \tau - \mathbb{I}_{[\mathbf{u}'\mathbf{Y}_i - \mathbf{u}'\mathbb{A}'(h_q)\mathbf{X}_i < 0]}$ and $\boldsymbol{\mu}^{r(n)}_{\text{P}(\tau\mathbf{u})}(h_q) := ((\mu^{r(n)}_{\text{P}(\tau\mathbf{u})})_{k_1}, \ldots, (\mu^{r(n)}_{\text{P}(\tau\mathbf{u})})_{k_p})'$, that

$$\begin{aligned}
\mathbf{a}^{(n)}_{\text{proj},\tau\mathbf{u}} &= \mathbb{A}(h_q)\mathbf{u}, \\
\boldsymbol{\lambda}^{(n)}_{\text{proj},\tau\mathbf{u}} &= -\frac{1}{n}\mathbb{Y}'\boldsymbol{\mu}^{r(n)}_{\text{P}(\tau\mathbf{u})} = \frac{1}{n}\sum_{i \notin h_q} f^{h_q}_{\boldsymbol{\tau}}(\mathbf{X}_i, \mathbf{Y}_i)(\mathbf{Y}_i - \mathbb{A}'(h_q)\mathbf{X}_i), \\
\boldsymbol{\mu}^{r(n)}_{\text{P}(\tau\mathbf{u})}(h_q) &= \sum_{i \notin h_q} f^{h_q}_{\boldsymbol{\tau}}(\mathbf{X}_i, \mathbf{Y}_i)(\mathbb{X}'(h_q))^{-1}\mathbf{X}_i \in [-\tau, 1-\tau]^p,
\end{aligned}$$

and

$$(\mu^{r(n)}_{\text{P}(\tau\mathbf{u})})_j = \begin{cases} (1-\tau) & \text{if } r^{(n)}_j < 0 \\ -\tau & \text{if } r^{(n)}_j > 0 \end{cases}$$

for any $j \in \{1, \ldots, n\}$, $j \notin h_q$.

We see that both $\boldsymbol{\lambda}^{(n)}_{\text{proj},\tau\mathbf{u}}$ and $\boldsymbol{\mu}^{r(n)}_{\text{P}(\tau\mathbf{u})}$ do not depend on $\mathbf{u}$ in any $\mathcal{C}_q$, while $\mathbf{a}^{(n)}_{\text{proj},\tau\mathbf{u}}$ does depend on $\mathbf{u}$ there, in a linear way. The set h_q determines the observations fitted by the projection regression $(\tau\mathbf{u})$-quantile hyperplane and each cone $\mathcal{C}_q$ corresponds to one optimal basis of (P). Clearly, faces of these cones are associated with those vectors $\mathbf{u}$ for which one coordinate of $\mathbf{r}^{(n)}$, $\mathbf{a}^{(n)}_{\mathrm{proj},\tau\mathbf{u}}$, or $\mathbf{u}$ changes its sign. We are mainly interested in the vertices of these cones because the corresponding unit vectors $\mathbf{u}$ comprise all the directions that are called τ-*critical* in Section 4.

In the small-sample location case, τ-critical directions can usually be identified visually as the pin points (or change points) of the corresponding quantile biplot $B^{(n)}(\tau)$; see the Introduction. This is because $a^{(n)}_{\mathrm{proj},\tau\mathbf{u}}$ is then always equal to the projection quantile of a data point and this observation in action changes in these directions in the maximum number of ways.

The problem (P) falls into the category of linear programs with parametric right-hand side. They are quite common in practice, their theory is well developed, and a general MATLAB toolbox for them has also been written; see [15]. Surprisingly, the task can be simplified substantially in our special case, which gives rise to a relatively fast, simple and reliable solver presented and evaluated in [18]. This only confirms the trend that applications of parametric programming in computational geometry still grow in number; see [19] for another recent paper on this topic.

Finally, we define projection regression rank scores $\boldsymbol{\nu}^{(n)}_{\boldsymbol{\tau}} \in [0,1]^n$ as

$$\boldsymbol{\nu}^{(n)}_{\boldsymbol{\tau}} = \tau\mathbf{1}_n + \boldsymbol{\mu}^{(n)}_{\boldsymbol{\tau}}, \quad \boldsymbol{\nu}^{(n)}_{\mathbf{0}} = \mathbf{0}_n.$$

If all projection $\boldsymbol{\tau}$-quantiles are uniquely defined, then $(\boldsymbol{\nu}^{(n)}_{\boldsymbol{\tau}})_i < 1$ if and only if the ith observation lies in the upper $\boldsymbol{\tau}$-quantile halfspace, $i = 1, 2, \ldots, n$. Consequently, the point regression halfspace pseudo-depth (see Section 4.2) of the ith data point might then be expressed as

$$\delta_i := \sup\{\tau \in (0,1) : \sup\{(\boldsymbol{\nu}^{(n)}_{\tau\mathbf{u}})_i : \mathbf{u} \in \mathcal{S}^{m-1}\} < 1\}. \tag{5.1}$$

This would exactly define the sample halfspace depth in the location case (and could be extended to the HPŠ approach analogously).

A. Appendix.

PROOF OF PROPOSITION 2.1. Let us show that the minimizer in (2.1) is unique under Assumption (A). For any fixed $(\mathbf{a}_0', \mathbf{b}_0')' \in \mathcal{M}_{\mathbf{u}}$, the function from $\mathcal{M}_{\mathbf{u}}$ to $\mathbb{R}$ mapping $(\mathbf{a}', \mathbf{b}')'$ to $\mathrm{E}[(\mathbf{b}_0'\mathbf{Y} - \mathbf{a}_0'\mathbf{X})\mathbb{I}_{[\mathbf{b}'\mathbf{Y}-\mathbf{a}'\mathbf{X}<0]}]$ is minimal at $(\mathbf{a}_0', \mathbf{b}_0')'$, and Assumption (A) ensures uniqueness of this minimum. For all $(\mathbf{a}', \mathbf{b}')' = t(\mathbf{a}_1', \mathbf{b}_1')' + (1-t)(\mathbf{a}_2', \mathbf{b}_2')'$, with $t \in (0,1)$ and $(\mathbf{a}_1', \mathbf{b}_1')' \neq (\mathbf{a}_2', \mathbf{b}_2')'$ (both in the *convex* set $\mathcal{M}_{\mathbf{u}}$), we then have

$$\begin{aligned}\Psi_\tau(\mathbf{a}, \mathbf{b}) &= \mathrm{E}[(\mathbf{b}'\mathbf{Y} - \mathbf{a}'\mathbf{X})(\tau - \mathbb{I}_{[\mathbf{b}'\mathbf{Y}-\mathbf{a}'\mathbf{X}<0]})] \\ &= t\,\mathrm{E}[(\mathbf{b}_1'\mathbf{Y} - \mathbf{a}_1'\mathbf{X})(\tau - \mathbb{I}_{[\mathbf{b}'\mathbf{Y}-\mathbf{a}'\mathbf{X}<0]})] + (1-t)\,\mathrm{E}[(\mathbf{b}_2'\mathbf{Y} - \mathbf{a}_2'\mathbf{X})(\tau - \mathbb{I}_{[\mathbf{b}'\mathbf{Y}-\mathbf{a}'\mathbf{X}<0]})] \\ &< t\,\mathrm{E}[(\mathbf{b}_1'\mathbf{Y} - \mathbf{a}_1'\mathbf{X})(\tau - \mathbb{I}_{[\mathbf{b}_1'\mathbf{Y}-\mathbf{a}_1'\mathbf{X}<0]})] + (1-t)\,\mathrm{E}[(\mathbf{b}_2'\mathbf{Y} - \mathbf{a}_2'\mathbf{X})(\tau - \mathbb{I}_{[\mathbf{b}_2'\mathbf{Y}-\mathbf{a}_2'\mathbf{X}<0]})] \\ &= t\Psi_\tau(\mathbf{a}_1, \mathbf{b}_1) + (1-t)\Psi_\tau(\mathbf{a}_2, \mathbf{b}_2).\end{aligned}$$

This shows that Ψ_τ is strictly convex on $\mathcal{M}_{\mathbf{u}}$, and hence the $\boldsymbol{\tau}$-quantiles defined through (2.1) are unique. □

Proof of Proposition 2.2. In this proof, we always take (without any loss of generality) $\boldsymbol{\Gamma}_{\mathbf{u}}$ of the form

$$\boldsymbol{\Gamma}_{\mathbf{u}} = \begin{pmatrix} \mathbf{I}_{p-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}_{\mathbf{u}}^{\mathbf{y}} \end{pmatrix}$$

in Definition 2.3. Now, writing $\mathbf{d} = ((\mathbf{d}^{\mathbf{w}})', (\mathbf{d}^{\mathbf{y}})')'$, we have

$$\mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y} - \mathbf{d}'\boldsymbol{\Gamma}_{\mathbf{u}}'(\mathbf{W}', \mathbf{Y}')' - c)] = \mathrm{E}[\rho_\tau([\mathbf{u} - \boldsymbol{\Gamma}_{\mathbf{u}}^{\mathbf{y}}\mathbf{d}^{\mathbf{y}}]'\mathbf{Y} - (c, (\mathbf{d}^{\mathbf{w}})')\mathbf{X})],$$

where $[\mathbf{u} - \boldsymbol{\Gamma}_{\mathbf{u}}^{\mathbf{y}}\mathbf{d}^{\mathbf{y}}]'\mathbf{u} = 1$. Therefore, if $(c_{\boldsymbol{\tau}}, \mathbf{d}_{\boldsymbol{\tau}}')'$ is a minimizer of (2.4), then $(\mathbf{a}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}', \mathbf{b}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}')' := ((c_{\boldsymbol{\tau}}, (\mathbf{d}_{\boldsymbol{\tau}}^{\mathbf{w}})'), (\mathbf{u} - \boldsymbol{\Gamma}_{\mathbf{u}}^{\mathbf{y}}\mathbf{d}_{\boldsymbol{\tau}}^{\mathbf{y}})')'$ minimizes the $\mathcal{M}_{\mathrm{HP\check{S}}}$-based version of (2.1). On the other hand, since $\mathbf{u}\mathbf{u}' + \boldsymbol{\Gamma}_{\mathbf{u}}^{\mathbf{y}}(\boldsymbol{\Gamma}_{\mathbf{u}}^{\mathbf{y}})' = \mathbf{I}_m$, we can write

$$\mathrm{E}[\rho_\tau(\mathbf{b}'\mathbf{Y} - \mathbf{a}'\mathbf{X})] = \mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y} - ((\mathbf{a}^{\mathbf{w}})', -\mathbf{b}'\boldsymbol{\Gamma}_{\mathbf{u}}^{\mathbf{y}})\boldsymbol{\Gamma}_{\mathbf{u}}'(\mathbf{W}', \mathbf{Y}')' - a^1)]$$

for any $\mathbf{b}$ satisfying $\mathbf{b}'\mathbf{u} = 1$, where $\mathbf{a} =: (a^1, (\mathbf{a}^{\mathbf{w}})')'$. Hence, for any minimizer $(\mathbf{a}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}', \mathbf{b}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}')'$ of the $\mathcal{M}_{\mathrm{HP\check{S}}}$-based version of (2.1), $(c_{\boldsymbol{\tau}}, \mathbf{d}_{\boldsymbol{\tau}}')' := (c_{\boldsymbol{\tau}}, (\mathbf{d}_{\boldsymbol{\tau}}^{\mathbf{w}})', (\mathbf{d}_{\boldsymbol{\tau}}^{\mathbf{y}})')' := (a_{\mathrm{HP\check{S}},\boldsymbol{\tau}}^1, (\mathbf{a}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}^{\mathbf{w}})', -\mathbf{b}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}'\boldsymbol{\Gamma}_{\mathbf{u}}^{\mathbf{y}})'$ minimizes the objective function in (2.4). The two families of $\boldsymbol{\tau}$-quantile hyperplanes $\Pi_{\mathrm{HP\check{S}},\boldsymbol{\tau}}$ and $\Pi_{\mathcal{M}_{\mathrm{HP\check{S}}},\boldsymbol{\tau}}$ thus coincide. □

Proof of Lemma 4.1. Recall that (4.4) states that $D^{(n)}(\frac{\ell}{n})$ coincides with the intersection of all closed halfspaces containing at least $n-\ell+1$ observations. Actually, one can restrict to closed halfspaces containing exactly $n-\ell+1$ observations; see [4], page 1805. It can also be shown (see [5]) that $D^{(n)}(\frac{\ell}{n})$—provided that its interior is not empty—is bounded by hyperplanes containing at least m points that span an $(m-1)$-dimensional subspace of $\mathbb{R}^m$. This establishes the result since we assume that the observations are in general position and that $D^{(n)}(\frac{\ell}{n})$ has a nonempty interior. □

Proof of Lemma 4.2. (i) Fix ℓ, τ, and j as in the statement of the lemma. Clearly,

$$\bigcup_{\mathbf{u}\in C_\tau^{(n)}} \{H_{\mathrm{proj},\tau\mathbf{u}}^{(n)+} : n\mathrm{P}_n[\partial H_{\mathrm{proj},\tau\mathbf{u}}^{(n)+}] = m \text{ and } n\mathrm{P}_n[(H_{\mathrm{proj},\tau\mathbf{u}}^{(n)+})^c] = \ell - j\} \subset \mathcal{H}_{\ell-j}^{(n)}.$$

Now, fix $H \in \mathcal{H}^{(n)}_{\ell-j}$ and let $\mathbf{u}$ be the unit vector that is orthogonal to ∂H and points to the interior of H. Then H is a projection upper $(\tau\mathbf{u})$-quantile halfspace, since H satisfies the necessary and sufficient conditions equivalent to (3.8) : $n\tau \in [\ell - j, (\ell - j) + m] = [N, N + Z]$ (for $\tau > \frac{\ell-1}{n}$, this holds with strict inequalities so that H is then the unique such upper halfspace). This shows that we indeed have $\mathcal{H}^{(n)}_{\ell-j} \subset \bigcup_{\mathbf{u}\in C^{(n)}_\tau} \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m \text{ and } n\mathrm{P}_n[(H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}})^c] = \ell - j\}$, which establishes the result. □

Proof of Theorem 4.2. Fix ℓ and τ as in the statement of the theorem, and define $\delta_{n\tau,\ell-1}$ as in Lemma 4.2. The definition of $R^{(n)}_{\mathrm{proj}}(\tau)$ (see (4.1) with $\mathcal{M} = \mathcal{M}_{\mathrm{proj}}$) directly yields

$$\begin{aligned} R^{(n)}_{\mathrm{proj}}(\tau) &\subset \bigcap_{\mathbf{u}\in C^{(n)}_\tau} \bigcap \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m\} \\ &= \bigcap \bigcup_{\mathbf{u}\in C^{(n)}_\tau} \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m\}. \end{aligned} \tag{A.1}$$

Now, note that any upper halfspace $H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}$ (irrespective of the number of data points contained in its boundary) satisfies

$$n\mathrm{P}_n[(H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}})^c] \in \{\ell - m - \delta_{n\tau,\ell-1}, \ell - m - \delta_{n\tau,\ell-1} + 1, \ldots, \ell - 1\}, \tag{A.2}$$

since (3.8) implies that $N = n\mathrm{P}_n[(H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}})^c] \in [\lceil n\tau\rceil - m, \lfloor n\tau\rfloor] \subset [\ell - m - \delta_{n\tau,\ell-1}, \ell - 1]$. Using (A.2) jointly with Lemma 4.2 and Lemma 4.1 yields

$$\bigcap \bigcup_{\mathbf{u}\in C^{(n)}_\tau} \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m\} = \bigcap \bigcup_{j=1}^r \mathcal{H}^{(n)}_{\ell-j} \subset \bigcap_{j=1}^r D^{(n)}\big(\tfrac{\ell-j+1}{n}\big) = D^{(n)}\big(\tfrac{\ell}{n}\big), \tag{A.3}$$

where we set $r := \min(m + \delta_{n\tau,\ell-1}, \ell)$. Eventually, (A.2) shows that any $H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}$ contains at least $n - \ell + 1$ data points, so that (4.4) proves that $D^{(n)}(\frac{\ell}{n}) \subset R^{(n)}_{\mathrm{proj}}(\tau)$. This, jointly with (A.1) and (A.3), establishes the result. □

Proof of Theorem 4.3. As mentioned in Section 4, the result follows quite easily from the comparison of the projection and HPŠ subgradient conditions (3.4a)-(3.4c) and (3.6a)-(3.6c). These two sets of subgradient conditions indeed imply that

$$H^+_{\mathrm{proj},\tau\mathbf{u}} = H^+_{\mathrm{HP\check{S}},\tau\boldsymbol{\lambda}_{\mathrm{proj},\tau\mathbf{u}}/\|\boldsymbol{\lambda}_{\mathrm{proj},\tau\mathbf{u}}\|} \quad \text{and} \quad H^+_{\mathrm{HP\check{S}},\tau\mathbf{u}} = H^+_{\mathrm{proj},\tau\mathbf{b}_{\mathrm{HP\check{S}},\tau\mathbf{u}}/\|\mathbf{b}_{\mathrm{HP\check{S}},\tau\mathbf{u}}\|}$$

for any $\tau \in (0, 1)$ and $\mathbf{u} \in \mathcal{S}^{m-1}$, which shows both $R_{\mathrm{HP\check{S}}}(\tau) \subset R_{\mathrm{proj}}(\tau)$ and $R_{\mathrm{proj}}(\tau) \subset R_{\mathrm{HP\check{S}}}(\tau)$. □

Proof of Theorem 4.4. We start with the proof of $R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau) \subset R^{(n)\mathrm{prac}}_{\mathrm{HP\check{S}}}(\tau)$ for $\tau \in (0, 1)$. Recall that $R^{(n)\mathrm{prac}}_{\mathrm{HP\check{S}}}(\tau) := \bigcap_{\mathbf{u}\in\mathcal{S}^{m-1}} \bigcap \{H^{(n)+}_{\mathrm{HP\check{S}},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{HP\check{S}},\tau\mathbf{u}}] = m + p - 1\}$

and fix a halfspace $H^{(n)+}_{\mathrm{HP\check{S}},\tau\mathbf{u}} = H^{(n)+}_{\mathrm{HP\check{S}},\tau\mathbf{u}}(\mathbf{a}^{(n)}_{\mathrm{HP\check{S}},\tau\mathbf{u}}, \mathbf{b}^{(n)}_{\mathrm{HP\check{S}},\tau\mathbf{u}})$ in this intersection. By definition, this means that $(\mathbf{a}, \mathbf{b}) \mapsto \Psi^{(n)}_\tau(\mathbf{a}, \mathbf{b}) = \frac{1}{n}\sum_{i=1}^n \rho_\tau(\mathbf{b}'\mathbf{Y}_i - \mathbf{a}'\mathbf{X}_i)$ achieves its minimum over $\mathcal{M}_{\mathrm{HP\check{S}},\mathbf{u}} := \{(\mathbf{a}', \mathbf{b}')' \in \mathbb{R}^{m+p} : \mathbf{b}'\mathbf{u} = 1\}$ at $(\mathbf{a}^{(n)}_{\mathrm{HP\check{S}},\tau\mathbf{u}}, \mathbf{b}^{(n)}_{\mathrm{HP\check{S}},\tau\mathbf{u}})$. By using convexity of the mapping $(\mathbf{a}, \mathbf{b}) \mapsto \Psi^{(n)}_\tau(\mathbf{a}, \mathbf{b})$ and the fact that $\Psi^{(n)}_\tau(\gamma\mathbf{a}, \gamma\mathbf{b}) = \gamma\Psi^{(n)}_\tau(\mathbf{a}, \mathbf{b})$ for any positive real value γ, we obtain that $\mathbf{a}^{(n)}_{\mathrm{HP\check{S}},\tau\mathbf{u}}/\|\mathbf{b}^{(n)}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}\|$ coincides with the minimizer $\mathbf{a}^{(n)}_{\mathrm{proj},\tau\mathbf{u}}$ of the same mapping over $\mathcal{M}_{\mathrm{proj},\mathbf{u}_0} := \{(\mathbf{a}', \mathbf{b}')' \in \mathbb{R}^{m+p} : \mathbf{b} = \mathbf{u}_0\}$, with $\mathbf{u}_0 := \mathbf{b}^{(n)}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}/\|\mathbf{b}^{(n)}_{\mathrm{HP\check{S}},\boldsymbol{\tau}}\|$. Therefore $H^{(n)+}_{\mathrm{HP\check{S}},\tau\mathbf{u}} = H^{(n)+}_{\mathrm{HP\check{S}},\tau\mathbf{u}}(\mathbf{a}^{(n)}_{\mathrm{HP\check{S}},\tau\mathbf{u}}, \mathbf{b}^{(n)}_{\mathrm{HP\check{S}},\tau\mathbf{u}}) = H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}_0}(\mathbf{a}^{(n)}_{\mathrm{proj},\tau\mathbf{u}_0}) = H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}_0}$. We conclude that $R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau) := \bigcap_{\mathbf{u}\in C^{(n)}_\tau} \bigcap \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m + p - 1\} \subset R^{(n)\mathrm{prac}}_{\mathrm{HP\check{S}}}(\tau)$.

We then turn to the proof that $R^{(n)\mathrm{prac}}_{\mathrm{HP\check{S}}}(\tau) \subset R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau)$ for $\tau \in (0, 1)$, which is much more difficult. Fix a halfspace $H = H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}_{\mathrm{proj}}}$ in the intersection $R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau) := \bigcap_{\mathbf{u}\in C^{(n)}_\tau} \bigcap \{H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}} : n\mathrm{P}_n[\partial H^{(n)+}_{\mathrm{proj},\tau\mathbf{u}}] = m + p - 1\}$. Denote by $h := (i_1, \ldots, i_k)$, $1 \le i_1 < \ldots < i_k \le n$, the index set of those $k := m+p-1$ data points $\mathbf{Z}_i = (\mathbf{W}'_i, \mathbf{Y}'_i)'$ that are contained in ∂H. Let us also write $\mathbf{X}_i := (1, \mathbf{W}'_i)'$, $\mathbb{W}(h) := (\mathbf{W}_{i_1}, \ldots, \mathbf{W}_{i_k})$, $\mathbb{X}(h) := (\mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_k})$, $\mathbb{Y}(h) := (\mathbf{Y}_{i_1}, \ldots, \mathbf{Y}_{i_k})$, and $\mathbb{Z}(h) := (\mathbf{Z}_{i_1}, \ldots, \mathbf{Z}_{i_k})$. The arrays $\mathbb{W}(s)$, $\mathbb{W}(t)$, $\mathbb{X}(s)$, etc. will be defined accordingly on the basis of the index sets $s := \{i_1, \ldots, i_p\}$ and $t := \{i_{p+1}, \ldots, i_k\}$.

It follows from Proposition 2.2 that the HPŠ version of the problem in (2.6) is equivalent to the standard single-output quantile regression of responses $\mathbf{u}'\mathbf{Y}_i$, $i = 1, \ldots, n$, with respect to $\boldsymbol{\Gamma}'_\mathbf{u}\mathbf{Z}_i$, $i = 1, \ldots, n$, and a constant term. Consequently, Theorem 2.1 of [9] shows that $H = H^{(n)+}_{\mathrm{HP\check{S}},\tau\mathbf{u}}$ if and only if

$$-\tau\mathbf{1}_k \le \boldsymbol{\xi}_{\boldsymbol{\tau}}(h) := (\mathbb{X}'_\mathbf{u}(h))^{-1} \sum_{i\notin h} \left(\tau - \mathbb{I}_{[\mathbf{Z}_i \notin H]}\right) \begin{pmatrix} 1 \\ \boldsymbol{\Gamma}'_\mathbf{u}\mathbf{Z}_i \end{pmatrix} \le (1-\tau)\mathbf{1}_k \tag{A.4}$$

where $\mathbb{X}_\mathbf{u}(h) := (\mathbf{1}_k \vdots \mathbb{Z}'(h)\boldsymbol{\Gamma}_\mathbf{u})$ and the inequalities are considered coordinatewise.

It is clear (see Definition 2.3) that $\boldsymbol{\Gamma}_\mathbf{u}$ can always be taken as

$$\boldsymbol{\Gamma}'_\mathbf{u} = \begin{pmatrix} \mathbf{I}_{p-1} & \mathbf{0} \\ \mathbf{0} & \mathbb{C} \end{pmatrix}$$

for some $(m-1) \times m$ matrix $\mathbb{C}$, which allows to decompose $\mathbb{X}'_\mathbf{u}(h)$ into

$$\mathbb{X}'_\mathbf{u}(h) = \begin{pmatrix} \mathbb{X}(h) \\ \mathbb{C}\mathbb{Y}(h) \end{pmatrix} = \begin{pmatrix} \mathbb{X}(s) & \mathbb{X}(t) \\ \mathbb{C}\mathbb{Y}(s) & \mathbb{C}\mathbb{Y}(t) \end{pmatrix}.$$

Inverting this block matrix yields

$$(\mathbb{X}'_\mathbf{u}(h))^{-1} = \begin{pmatrix} \mathbb{X}(s)^{-1} + \mathbb{X}(s)^{-1}\mathbb{X}(t)\mathbb{S}\mathbb{C}\mathbb{Y}(s)\mathbb{X}(s)^{-1} & -\mathbb{X}(s)^{-1}\mathbb{X}(t)\mathbb{S} \\ -\mathbb{S}\mathbb{C}\mathbb{Y}(s)\mathbb{X}(s)^{-1} & \mathbb{S} \end{pmatrix}, \tag{A.5}$$

where

$$\mathbb{S}^{-1} := \mathbb{C}[\mathbb{Y}(t) - \mathbb{Y}(s)\mathbb{X}(s)^{-1}\mathbb{X}(t)]. \tag{A.6}$$

Now, defining the p-vector $\mathbf{T}_D^X$ and the m-vector $\mathbf{T}_D^Y$ through

$$\begin{pmatrix} \mathbf{T}_D^X \\ \mathbf{T}_D^Y \end{pmatrix} := \begin{pmatrix} \mathbf{T}_D^1 \\ \mathbf{T}_D^Z \end{pmatrix} := \tau \sum_{\mathbf{Z}_i \in H \setminus \partial H} \begin{pmatrix} 1 \\ \mathbf{Z}_i \end{pmatrix} + (\tau - 1) \sum_{\mathbf{Z}_i \notin H} \begin{pmatrix} 1 \\ \mathbf{Z}_i \end{pmatrix},$$

let us consider

$$\mathbf{u} = \mathbf{u}_{\mathbf{w}} = \frac{\mathbf{T}_D^Y - \mathbf{v}_{\mathbf{w}}}{\|\mathbf{T}_D^Y - \mathbf{v}_{\mathbf{w}}\|},$$

where $\mathbf{v}_{\mathbf{w}} := \mathbb{Y}(s)\mathbb{X}(s)^{-1}\mathbf{T}_D^X - [\mathbb{Y}(t) - \mathbb{Y}(s)\mathbb{X}(s)^{-1}\mathbb{X}(t)]\mathbf{w}$ for some fixed $(m-1)$-vector $\mathbf{w}$. Definition 2.3 requires the orthogonality condition $\mathbb{C}\mathbf{u} = \mathbf{0}$, which implies

$$\begin{aligned} \sum_{i \notin h} \left(\tau - \mathbb{I}_{[\mathbf{Z}_i \notin H]}\right) \begin{pmatrix} 1 \\ \mathbf{\Gamma}_{\mathbf{u}}' \mathbf{Z}_i \end{pmatrix} &= \tau \sum_{\mathbf{Z}_i \in H \setminus \partial H} \begin{pmatrix} 1 \\ \mathbf{\Gamma}_{\mathbf{u}}' \mathbf{Z}_i \end{pmatrix} + (\tau - 1) \sum_{\mathbf{Z}_i \notin H} \begin{pmatrix} 1 \\ \mathbf{\Gamma}_{\mathbf{u}}' \mathbf{Z}_i \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{T}_D^X \\ \mathbb{C}\mathbf{T}_D^Y \end{pmatrix} = \begin{pmatrix} \mathbf{T}_D^X \\ \mathbb{C}\mathbf{v}_{\mathbf{w}} \end{pmatrix}. \end{aligned}$$

Hence, in view of (A.5), we obtain that

$$\begin{aligned} \boldsymbol{\xi}_{\boldsymbol{\tau}}(h) &= (\mathbb{X}_{\mathbf{u}}'(h))^{-1} \sum_{i \notin h} \left(\tau - \mathbb{I}_{[\mathbf{Z}_i \notin H]}\right) \begin{pmatrix} 1 \\ \mathbf{\Gamma}_{\mathbf{u}}' \mathbf{Z}_i \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{X}(s)^{-1}\mathbf{T}_D^X + \mathbb{X}(s)^{-1}\mathbb{X}(t)\mathbb{S}\mathbb{C}\mathbb{Y}(s)\mathbb{X}(s)^{-1}\mathbf{T}_D^X - \mathbb{X}(s)^{-1}\mathbb{X}(t)\mathbb{S}\mathbb{C}\mathbf{v}_{\mathbf{w}} \\ -\mathbb{S}\mathbb{C}\mathbb{Y}(s)\mathbb{X}(s)^{-1}\mathbf{T}_D^X + \mathbb{S}\mathbb{C}\mathbf{v}_{\mathbf{w}} \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{X}(s)^{-1}\mathbf{T}_D^X + \mathbb{X}(s)^{-1}\mathbb{X}(t)\mathbb{S}\mathbb{C}[\mathbb{Y}(s)\mathbb{X}(s)^{-1}\mathbf{T}_D^X - \mathbf{v}_{\mathbf{w}}] \\ -\mathbb{S}\mathbb{C}[\mathbb{Y}(s)\mathbb{X}(s)^{-1}\mathbf{T}_D^X - \mathbf{v}_{\mathbf{w}}] \end{pmatrix}. \end{aligned}$$

The definition of $\mathbf{v}_{\mathbf{w}}$ and (A.6) lead to

$$\boldsymbol{\xi}_{\boldsymbol{\tau}}(h) = \begin{pmatrix} \mathbb{X}(s)^{-1}\mathbf{T}_D^X + \mathbb{X}(s)^{-1}\mathbb{X}(t)\mathbf{w} \\ -\mathbf{w} \end{pmatrix}. \tag{A.7}$$

Since the observations are assumed to be in general position, the directional vector $\mathbf{u}_{\text{proj}}$ (for which $H = H^{(n)+}_{\text{proj},\tau\mathbf{u}_{\text{proj}}}$) can always be changed a little into a unit vector $\bar{\mathbf{u}}_{\text{proj}}$ such that $\partial H^{(n)+}_{\text{proj},\tau\bar{\mathbf{u}}_{\text{proj}}}$ contains exactly p data points from the $m+p-1$ indexed by h; see Section 5.2. If the observations are reindexed so that these p data points are indexed by s, then the necessary and sufficient subgradient conditions for the projection ($\tau\bar{\mathbf{u}}_{\rm proj}$)-quantile (provided by Theorem 2.1 of [9]) state that, for a certain $\mathbf{w}_0 \in \{\tau, \tau-1\}^{m-1}$, we have

$$-\tau\mathbf{1}_p \leq \mathbb{X}(s)^{-1}\mathbf{T}_D^X + \mathbb{X}(s)^{-1}\mathbb{X}(t)\mathbf{w}_0 \leq (1-\tau)\mathbf{1}_p,$$

where $\mathbb{X}(s)$, $\mathbb{X}(t)$, and $\mathbf{T}_D^X$ are the same as above. Hence, in view of (A.7), all inequalities in (A.4) hold with $\boldsymbol{\tau} = \tau\mathbf{u}_{\mathbf{w}_0}$, which proves that $H = H^{(n)+}_{{\rm proj},\tau\mathbf{u}_{\rm proj}} = H^{(n)+}_{{\rm HP\check{S}},\tau\mathbf{u}_{\mathbf{w}_0}}$. Therefore $R^{(n)\rm prac}_{\rm HP\check{S}}(\tau) \subset R^{(n)\rm crit}_{\rm proj}(\tau)$. □

Acknowledgement

The research work of Davy Paindaveine was supported by a Mandat d'Impulsion Scientifique of the Fonds National de la Recherche Scientifique, Communauté française de Belgique. That of Miroslav Šiman was supported partly by a *chercheur postdoctoral temporaire* contract of the Fonds National de la Recherche Scientifique, Communauté française de Belgique, and partly by Project 1M06047 of the Ministry of Education, Youth and Sports of the Czech Republic. The authors would like to thank Ivan Mizera for kindly providing the Matlab code for computation of bivariate halfspace depth contours he coauthored with David Eppstein.

References

[1] D. Bertsimas, G.J. Lauprete, A. Samarov, Shortfall as a risk measure: properties, optimization and applications, J. Econ. Dyn. Control 28 (2004) 1353-1381.

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

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

[4] D.L. Donoho, M. Gasko, Breakdown properties of location estimates based on halfspace depth and projected outlyingness, Ann. Statist. 20 (1992) 1803-1827.

[5] K. Fukuda, V. Rosta, Data depth and maximum feasible subsystems, in D. Avis, A. Hertz, O. Marcotte (eds.), Graph Theory and Combinatorial Optimization, Springer, 2005, pp. 37-67.

[6] M. Hallin, D. Paindaveine, M. Šiman, Multivariate quantiles and multiple-output regression quantiles: From L_1 optimization to halfspace depth, Ann. Statist. 38 (2010) 635-669 (with discussion).

[7] M. Hallin, D. Paindaveine, M. Šiman, Rejoinder: Multivariate quantiles and multiple-output regression quantiles, Ann. Statist. 38 (2010) 694-703.

[8] D. Hlubinka, L. Kotík, O. Vencálek, Weighted halfspace depth, Kybernetika 46 (2010) 125-148.

[9] R. Koenker, Quantile Regression, 1st edition, Cambridge University Press, New York, 2005.

[10] R. Koenker, Quantile regression in R: a vignette, downloadable from CRAN: *http://cran.r-project.org*, 2007.

[11] R. Koenker, G.J. Bassett, Regression quantiles, Econometrica 46 (1978) 33-50.

[12] V. Koltchinskii, M-estimation, convexity and quantiles, Ann. Statist. 25 (1997) 435-477.

[13] L. Kong, I. Mizera, Quantile tomography: using quantiles with multivariate data, submitted (2008), arXiv:0805.0056.

[14] L. Kong, I. Mizera, 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 (2010) 685-693.

[15] M. Kvasnica, P. Grieder, M. Baotić, Multi-Parametric Toolbox (MPT), downloadable from *http://control.ee.ethz.ch/˜mpt/*, 2004.

[16] B. Laine, Depth contours as multivariate quantiles: A directional approach. Unpublished Master thesis (advisor M. Hallin). Univ. Libre de Bruxelles, Brussels, 2001.

[17] R.Y. Liu, J.M. Parelius, K. Singh, Multivariate analysis by data depth: descriptive statistics, graphics and inference, Ann. Statist. 27 (1999) 783-840.

[18] D. Paindaveine, M. Šiman, Computing multiple-output regression quantile regions from projection quantiles, submitted.

[19] S.V. Raković, P. Grieder, C. Jones, Computation of Voronoi diagrams and Delaunay triangulation via parametric linear programming, ETH Technical Report AUT04-03 (2004).

[20] P.J. Rousseeuw, M. Hubert, Regression depth, J. Amer. Statist. Assoc. 94 (1999) 388-402.

[21] P.J. Rousseeuw, I. Ruts, The depth function of a population distribution, Metrika 49 (1999) 213-244.

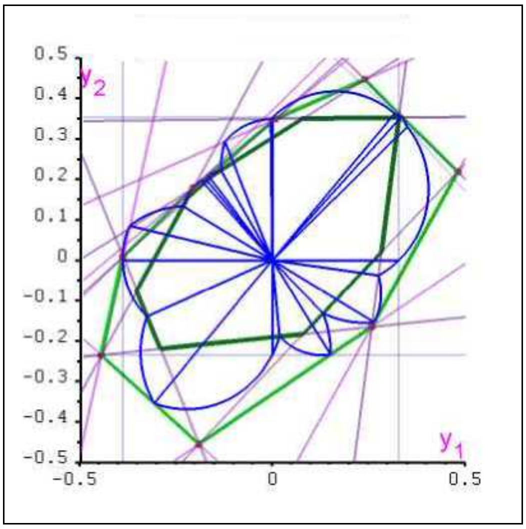
[22] J. Sadefo Kamdem, Value-at-Risk and expected shortfall for linear portfolios with elliptically distributed risk factors, Int. J. Theor. Appl. Finance 8 (2005) 537-551.

[23] R. Serfling, Quantile functions for multivariate analysis: approaches and applications, Statist. Neerlandica 56 (2002) 214-232.

[24] C.G. Small, Measures of centrality for multivariate and directional distributions, Canad. J. Statist. 56 (1987) 31-39.

[25] Y. Wei, An approach to multivariate covariate-dependent quantile contours with application to bivariate conditional growth charts, J. Amer. Statist. Assoc. 103 (2008) 397-409.

[26] Y. Zuo, R. Serfling, General notions of statistical depth function, Ann. Statist. 28 (2000) 461-482.



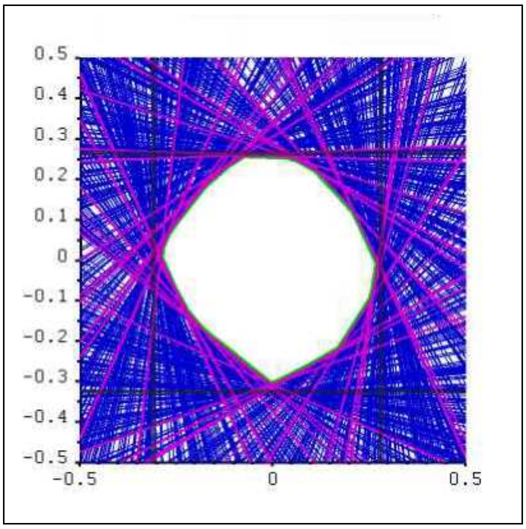


Figure 1: The left plot contains $n = 9$ (red) points drawn from the bivariate distribution with independent $U([-.5, .5])$ (uniform over $(-.5, .5)$) marginals, and shows how two successive halfspace depth contours, $D^{(n)}(\tau_1 = .1)$ (light green) and $D^{(n)}(\tau_2 = .2)$ (dark green), can be inferred from the τ_2-quantile biplot (blue contour). All the directions found by parametric programming (indicated by the blue segments leading from the origin) are also faces of all the cones C_q and comprise not only all τ_2-critical directions but also some others (such as the horizontal and vertical directions in this case). All these directions are orthogonal to the corresponding projection τ_2-quantile hyperplanes (gray or magenta) that include all the prolonged faces (magenta) of the two halfspace depth contours displayed. Note that even the redundant τ_2-quantile hyperplanes (gray) border the inner halfspace depth contour from the right side.

Figure 2: The right plot shows all the projection τ-quantile hyperplanes (blue, black or magenta) corresponding to the directions obtained from parametric programming, for $\tau = .2$ and $n = 499$ points from the bivariate distribution with independent $U([-.5, .5])$ marginals. The projection τ-quantile hyperplanes really determining the shape of the halfspace depth contour $D^{(n)}(\tau)$ (green) are drawn in magenta and those corresponding to semiaxial directions are plotted in black.

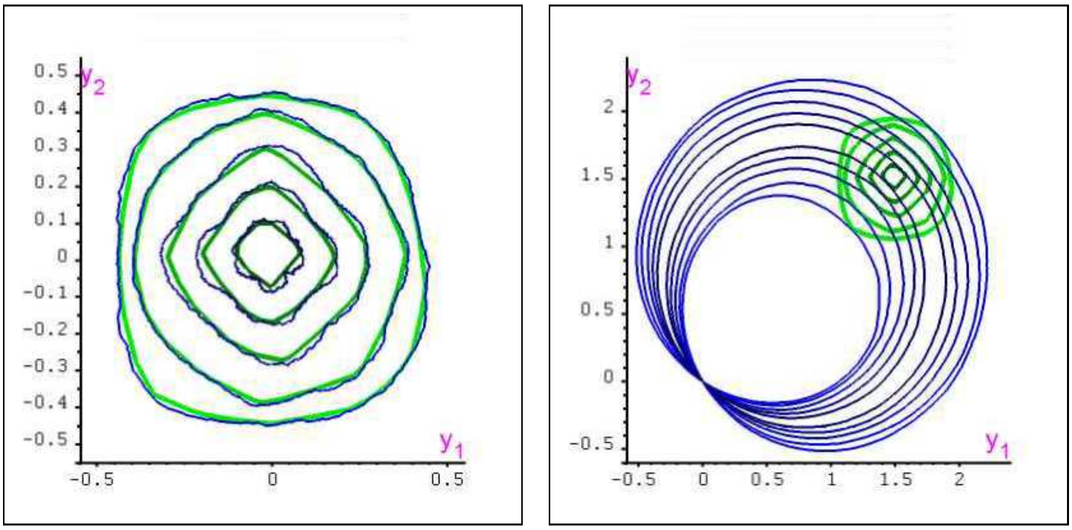


Figure 3: Halfspace depth contours $D^{(n)}(\tau)$ (green) and quantile biplot contours $B^{(n)}(\tau)$ (blue), for $n = 999$ and $\tau \in \{.05, .10, .20, .30, .40\}$, obtained from the bivariate distribution with independent $U([-.5, .5])$ (left) and $U([1, 2])$ (right) marginals. Quantile biplots are not even shift equivariant, but always contain halfspace depth contours at the same level (which is trivial to prove).

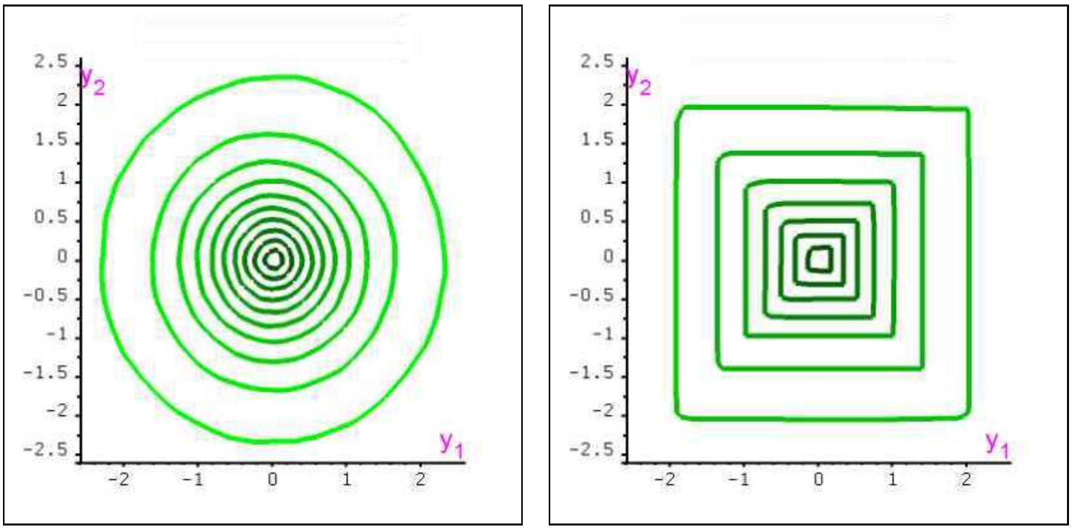


Figure 4: Halfspace depth contours $D^{(n)}(\tau)$, for $n = 9\ 999$, obtained from the bivariate standard normal distribution and for $\tau \in \{.01, .05, .10, .15, \ldots, .45\}$ (left), and from the bivariate distribution with independent standard Cauchy marginals and for $\tau \in \{.15, .20, \ldots, .45\}$ (right). They match their theoretical counterparts (circles and squares) very well.

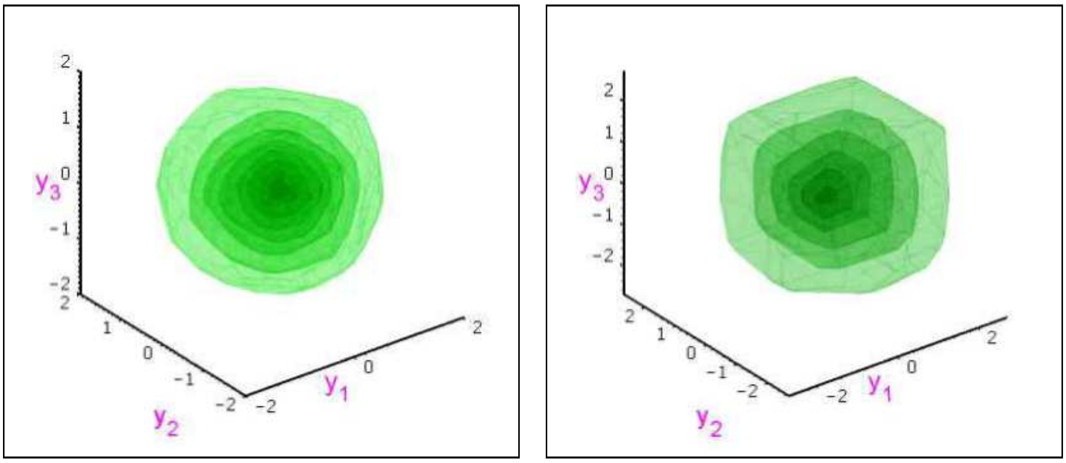


Figure 5: Halfspace depth contours $D^{(n)}(\tau)$, for $n = 199$, obtained from the trivariate standard normal distribution and for $\tau \in \{.05, .10, .15, \ldots, .40\}$ (left), and from the trivariate distribution with independent standard Cauchy marginals and for $\tau \in \{.15, .20, \ldots, .40\}$ (right). They match their theoretical counterparts (spheres and cubes) quite well, even for such a small sample size.

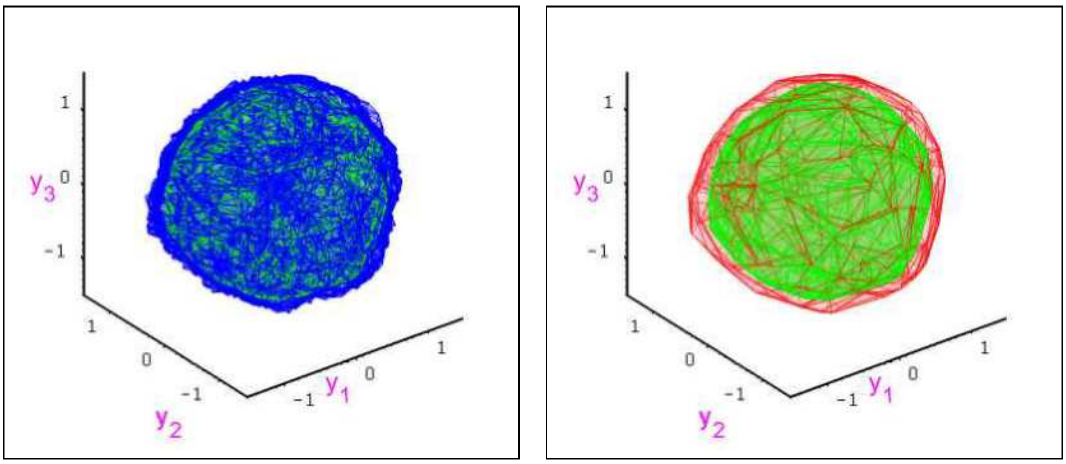


Figure 6: Raw piecewise linear (blue) and convex (red) approximations of quantile biplots $B^{(n)}(\tau)$ of order $\tau = .10$ and the resulting halfspace depth contour $D^{(n)}(\tau)$ (green), with $n = 199$, from the trivariate standard normal distribution.

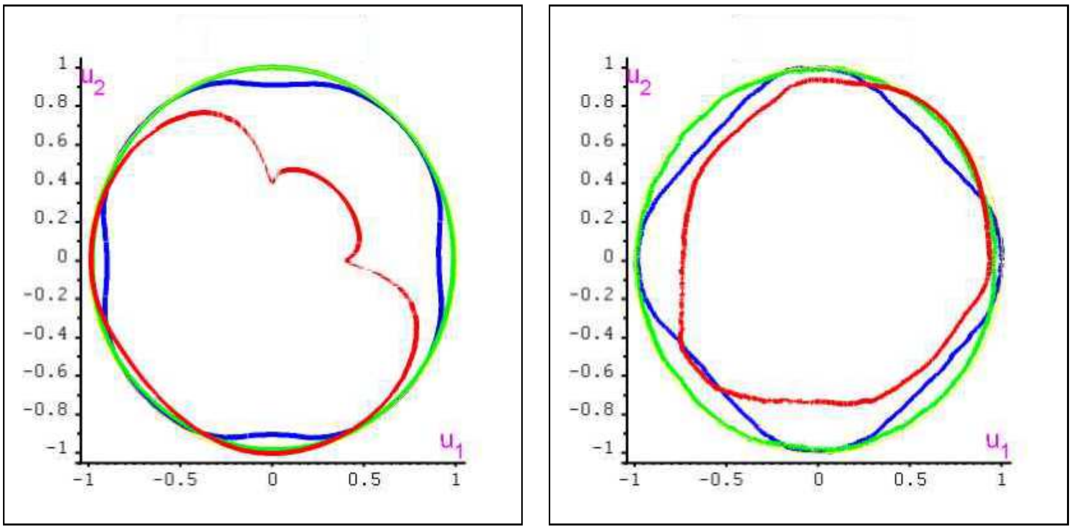


Figure 7: Polar plots of the mappings $\mathbf{u} \in \mathcal{S}^1 \mapsto \|\boldsymbol{\lambda}^{(n)}_{\mathrm{proj},\tau\mathbf{u}}\|\mathbf{u}/(\sup_{\mathbf{v}\in\mathcal{S}^1} \|\boldsymbol{\lambda}^{(n)}_{\mathrm{proj},\tau\mathbf{v}}\|)$ (left) and $\mathbf{u} \mapsto |a^{(n)}_{\mathrm{proj},\tau\mathbf{u}}|\mathbf{u}/(\sup_{\mathbf{v}\in\mathcal{S}^1} |a^{(n)}_{\mathrm{proj},\tau\mathbf{v}}|)$ (right), for $\tau = 0.1$ and $n = 9\ 999$ points drawn from the bivariate standard normal distribution (green), the distribution with independent $U([-.5, .5])$ marginals (blue), and the distribution with independent $Exp(1)-1$ marginals (red), respectively. The resulting shapes clearly reflect the axes of symmetry of the underlying distributions.

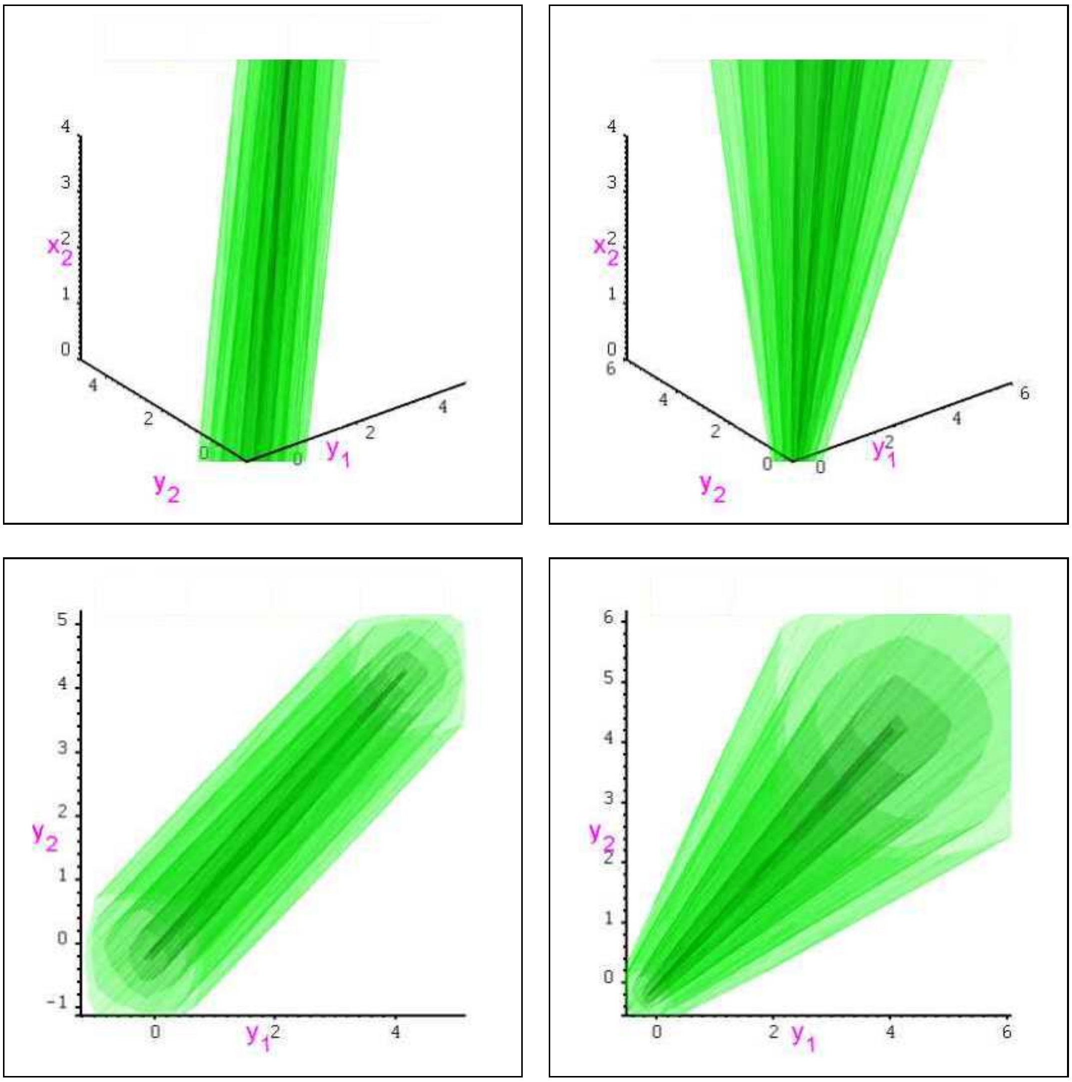


Figure 8: Two different views on the regression τ-quantile contours $R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau)$, with $\tau \in \{.01, .05, .15, .30, .45\}$, from 999 data points in a homoscedastic ($(Y_1, Y_2)' = 4(X_2, X_2)' + 2\varepsilon$; left) and a heteroscedastic ($(Y_1, Y_2)' = 4(X_2, X_2)' + 2\sqrt{X_2}\varepsilon$; right) bivariate-output regression setting, where $X_2 \sim U([0, 1])$ and ε has independent $U([-.5, .5])$ marginals.

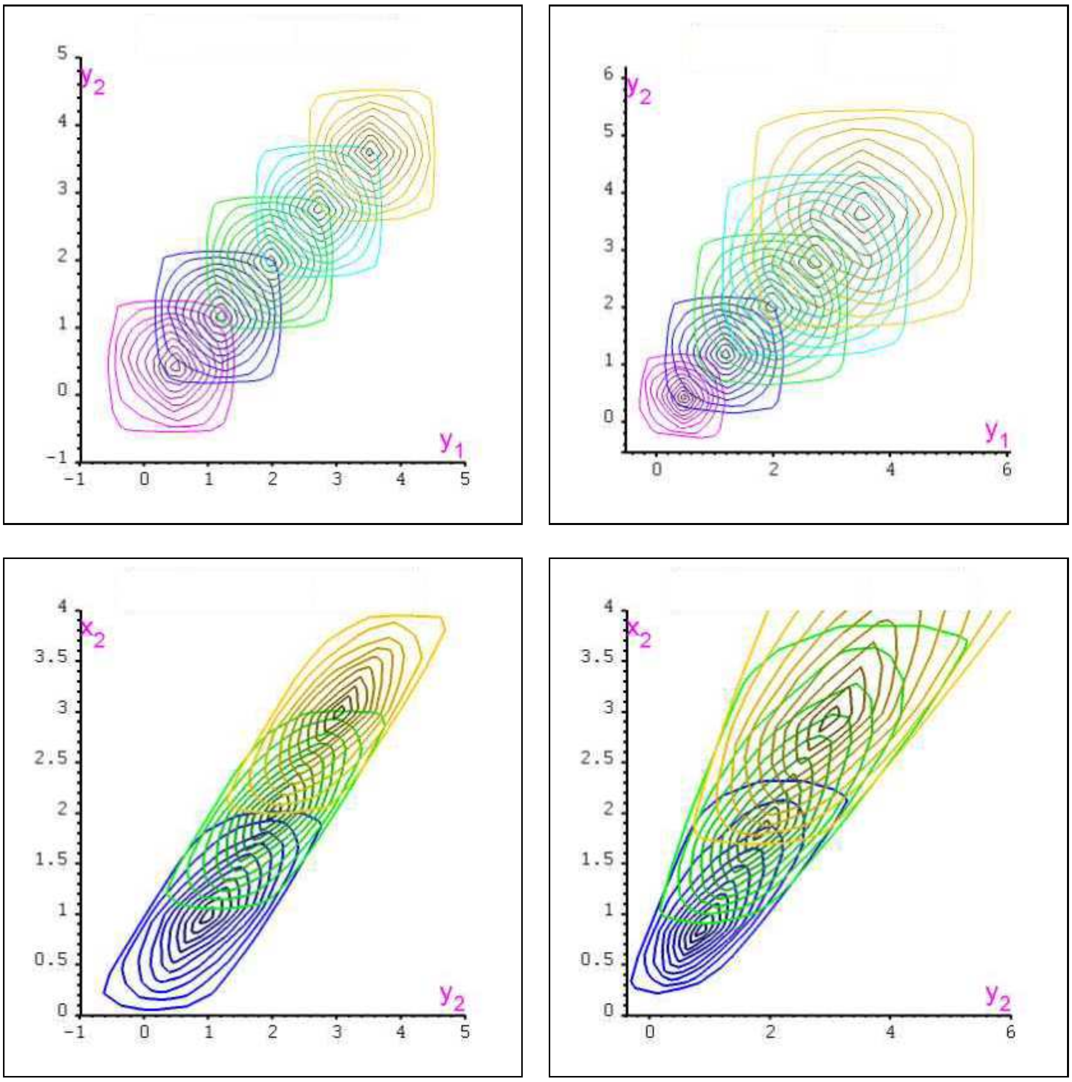


Figure 9: Various cuts of the regression τ-quantile regions $R^{(n)\mathrm{crit}}_{\mathrm{proj}}(\tau)$ from the same two models (left and right, respectively) as in Figure 8 with $n = 999$ observations. The top plots provide regression τ-quantile cuts, $\tau \in \{.01, .05, .10, .15, .20, \ldots, .45\}$, through 10% (magenta), 30% (blue), 50% (green), 70% (cyan) and 90% (yellow) empirical quantiles of X_2; the bottom ones show regression τ-quantile cuts for the same τ values, and through 25% (blue), 50% (green) and 75% (yellow) empirical quantiles of Y_1. Their centers provide information about trend, and their shapes and sizes shed light on variability.