



N° 1022

February 2023

## “Multivariate Expectiles, Expectile Depth and Multiple-Output Expectile Regression”

Abdelaati Daouia and Davy Paindaveine

# Multivariate Expectiles, Expectile Depth and Multiple-Output Expectile Regression

Abdelaati Daouia\* and Davy Paindaveine†

\*<sup>,</sup>†Toulouse School of Economics, Université Toulouse 1 Capitole

†ECARES and Département de Mathématique, Université Libre de Bruxelles

## Abstract

Despite the importance of expectiles in fields such as econometrics, risk management, and extreme value theory, expectile regression unfortunately so far remains limited to single-output problems. To improve on this, we define hyperplane-valued multivariate expectiles that show strong advantages over their point-valued competitors. Our expectiles are directional in nature and provide centrality regions when all directions are considered. These regions define a new statistical depth, the *halfspace expectile depth*, that is an  $L_2$  version of the celebrated ( $L_1$ ) Tukey halfspace depth. We study thoroughly the proposed expectiles, the expectile depth, and the corresponding regions. When compared to their  $L_1$  counterparts, these concepts enjoy distinctive properties that will be of primary interest to practitioners. In particular, expectile depth is maximized at the mean vector, is smoother than the halfspace depth, and exhibits surprising monotonicity properties that are key for computational purposes. Finally, the proposed multivariate expectiles allow us to define multiple-output expectile regression methods, that, in risk-oriented applications in particular, dominate their analogs based on standard quantiles.

*Keywords:* Centrality regions; Multivariate expectiles; Multivariate quantiles; Multiple-output regression; Statistical depth

---

\*Abdelaati Daouia is Maître de Conférences, Université Toulouse 1 Capitole TSE, 1, Esplanade de l’Université, 31080 Toulouse Cedex 06, France (E-mail: [abdelati.daouia@tse-fr.eu](mailto:abdelati.daouia@tse-fr.eu)). Davy Paindaveine is Professor of Statistics, Université libre de Bruxelles, ECARES and Département de Mathématique, Avenue F. D. Roosevelt, 50, CP 114/04, B-1050 Bruxelles, Belgium (E-mail: [dpaindav@ulb.ac.be](mailto:dpaindav@ulb.ac.be)).

# 1 Introduction

Whenever one wants to assess the impact of a vector of covariates  $\mathbf{X}$  on a scalar response  $Y$ , mean regression, in its various forms (linear, nonlinear, or nonparametric), remains by far the most popular method. Mean regression, however, only captures the conditional mean

$$\mu(\mathbf{x}) := \mathbb{E}[Y|\mathbf{X} = \mathbf{x}] = \arg \min_{y \in \mathbb{R}} \mathbb{E}[(Y - y)^2 | \mathbf{X} = \mathbf{x}]$$

of the response, hence fails to describe thoroughly the conditional distribution of  $Y$  given  $\mathbf{X}$ . This was the main motivation to introduce the [Newey and Powell \(1987\)](#) *expectile regression*, that considers the conditional expectiles

$$e_\alpha(\mathbf{x}) := \arg \min_{y \in \mathbb{R}} \mathbb{E}[\rho_\alpha(Y - y) - \rho_\alpha(Y) | \mathbf{X} = \mathbf{x}], \quad \alpha \in (0, 1), \quad (1)$$

where  $\rho_\alpha(t) := \{(1 - \alpha)\mathbb{I}[t < 0] + \alpha\mathbb{I}[t > 0]\}t^2$  is an asymmetric quadratic loss function (throughout,  $\mathbb{I}[A]$  stands for the indicator function of  $A$ ). Like the quantile regression from [Koenker and Bassett \(1978\)](#), that is obtained by substituting the absolute function  $|t|$  for the quadratic one  $t^2$  in  $\rho_\alpha(t)$ , expectile regression fully characterizes the conditional distribution of the response, but it nicely includes the conditional mean  $\mu(\mathbf{x})$  as a particular case. Sample conditional expectiles, unlike their quantile counterparts, are sensitive to extreme observations, which may actually be an asset in some applications; in financial risk management, for instance, quantiles are often criticized as too liberal or optimistic (due to their insensitivity to extreme losses) and expectiles are therefore favoured in any prudent and reactive risk analysis ([Daouia et al., 2018](#)).

Expectile regression shows other advantages over quantile regression, of which we mention only a few here. First, inference on quantiles requires estimating nonparametrically the conditional density of the response at the considered quantiles, which is notoriously difficult. In contrast, inference on expectiles can be performed without resorting to any

smoothing, bootstrap or Bayesian technique, which makes it easy, e.g., to test for homoscedasticity or for conditional symmetry in linear regression models (Newey and Powell, 1987). Second, since expectile regression encompasses classical mean regression, it is closer to the least squares notion of explained variance and, in parametric cases, expectile regression coefficients can be interpreted with respect to variance heteroscedasticity. This is of particular relevance in complex regression specifications including nonlinear, random or spatial effects (Sobotka and Kneib, 2012). Third, expectile smoothing techniques, based on kernel smoothing (Yao and Tong, 1996) or penalized splines (Schnabel and Eilers, 2009), show better smoothness and stability than their quantile counterparts and also make expectile crossings far more rare than quantile crossings; see Schnabel and Eilers (2009), Eilers (2013) and Schulze Waltrup et al. (2015). These points explain why expectiles recently regained much interest in econometrics; see, e.g., Kuan et al. (2009), De Rossi and Harvey (2009), and Embrechts and Hofert (2014).

Despite these nice properties, expectile regression still suffers from an important drawback, namely its limitation to single-output problems. In contrast, many works developed multiple-output *quantile* regression methods. We refer, e.g., to Chakraborty (2003), Cheng and De Gooijer (2007), Hallin et al. (2010), Cousin and Di Bernardino (2013), Waldmann and Kneib (2015), Hallin et al. (2015), Carlier et al. (2016, 2017), and Chavas (2018), with applications in particular in the analysis of growth trajectories; see, e.g., Wei (2008) or McKeague et al. (2011). This is in line with the fact that defining a satisfactory concept of multivariate quantile is a classical problem that has attracted much attention in the literature (we refer to Serfling (2002) and to the references therein), whereas the literature on multivariate expectiles is much sparser. Some early efforts to define multivariate expectiles can be found in Koltchinski (1997), Breckling et al. (2001) and Kokic et al. (2002), that all define more generally multivariate versions of the *M-quantiles* from Breckling and Cham-

bers (1988); a first concept of multivariate M-quantile was actually already discussed in Breckling and Chambers (1988) itself. Recently, there has been a renewed interest in defining multivariate expectiles; see, e.g., Cousin and Di Bernardino (2014), Maume-Deschamps et al. (2017a,b), and Herrmann et al. (2018). Multivariate risk handling in finance and actuarial sciences is mostly behind this growing interest.

This paper introduces a concept of multivariate expectiles that enjoys many desirable properties, particularly in terms of affine equivariance (while this equivariance property is a standard requirement in the companion problem of defining multivariate quantiles, the available concepts of multivariate expectiles are at best orthogonal-equivariant). Like their competitors, our multivariate expectiles are directional quantities, but they are hyperplane-valued rather than point-valued. Despite this different nature, they still generate centrality regions when all directions are considered. While this has not been discussed in the multivariate expectile literature, this defines an expectile concept of statistical depth. The resulting *halfspace expectile depth* is the  $L_2$  version of the Tukey (1975) halfspace depth and satisfies the desirable properties of depth from Zuo and Serfling (2000). Remarkably, this new depth can alternatively be obtained by replacing, in the halfspace depth, standard quantile outlyingness with expectile outlyingness. This is a key result that allows us to study expectile depth. This new depth offers many properties that, in comparison with halfspace depth, should be most appealing to practitioners. In particular, it is maximized at the mean vector, it is smooth, and it shows a surprising monotonicity that is key for its computation. Finally, our multivariate expectiles allow us to define multiple-output expectile regression methods, that, in risk-oriented applications in particular, will dominate their analogs based on standard quantiles.

The outline of the paper is as follows. In Section 2, we briefly review the concept of univariate expectile. In Section 3, we introduce our concept of multivariate expectiles and

compare the resulting expectile regions with those associated with alternative expectile concepts. In Section 4, we define the expectile depth and investigate its properties. In Section 5, we treat several real data examples, which gives us the opportunity to show that the proposed concept allows performing multiple-output expectile regression. Supplementary materials provide the following further contributions: for the sake of completeness, we describe there some of the main competing multivariate expectile concepts (Section S.1). We compute expectile depth and expectile depth regions in several multivariate examples (Section S.2). We state asymptotic results for the proposed expectile depth (Section S.3). We illustrate on simulated data the proposed multiple-output expectile regression methods and show that these dominate the corresponding quantile-based methods in terms of crossings (Section S.4). We discuss the relation between multivariate expectiles and risk measures, and we show that our expectiles satisfy the coherency axioms of multivariate risk measures (Section S.5). Finally, we prove all results of the paper (Section S.6).

## 2 Univariate expectiles and expectile depth

The expectiles from Newey and Powell (1987) arise as the solution of an asymmetric least squares problem. More precisely, an order- $\alpha$  expectile of  $P$  (a probability measure over  $\mathbb{R}$  with finite first moment) is

$$e_\alpha(P) = \arg \min_{z \in \mathbb{R}} \mathbb{E}[\rho_\alpha(Z - z) - \rho_\alpha(Z)], \quad (2)$$

where  $\rho_\alpha$  is the asymmetric quadratic loss function in (1) and where the random variable  $Z$  has distribution  $P$ . The mean  $\mu(P) = \mathbb{E}[Z]$  is obtained for  $\alpha = 1/2$ , and the interexpectile intervals  $[e_\alpha(P), e_{1-\alpha}(P)]$  (with  $\alpha < 1/2$ ) form a family of nested regions that all contain the mean. Other functionals of interest can be built, such as the “interpexpectile

range”  $\sigma(P) := e_{3/4}(P) - e_{1/4}(P)$ , which is a scale measure. Since the minimizer in (2) may be non-unique, we will rather define the order- $\alpha$  expectile of  $P$  as

$$e_\alpha(P) := \inf\{z \in \mathbb{R} : G(z) \geq \alpha\}, \quad \text{with } G(z) := \frac{\mathbb{E}[|Z - z|\mathbb{I}[Z \leq z]]}{\mathbb{E}[|Z - z|]}, \quad (3)$$

which is indeed a particular minimizer; see the theorem in [Jones \(1994\)](#) or its generalization in [Theorem S.4](#), that also state that  $G$  is a (continuous) cumulative distribution function. For later purposes, it is important to note that the larger  $G(z)(\leq 1/2)$  (resp.,  $1 - G(z)(\leq 1/2)$ ), the less  $z$  is outlying below (resp., above) the central location  $\mu(P)$ . Therefore,  $D(z, P) := \min(G(z), 1 - G(z))$  measures the centrality—as opposed to outlyingness—of  $z$  with respect to  $P$ . In other words,  $D(z, P)$  defines a measure of *statistical depth* over  $\mathbb{R}$ ; see [Zuo and Serfling \(2000\)](#). To the best of our knowledge, this *expectile depth* has not been considered in the literature. The corresponding depth regions  $R_\alpha(P) := \{z \in \mathbb{R} : D(z, P) \geq \alpha\}$  coincide with the interexpectile intervals  $[e_\alpha(P), e_{1-\alpha}(P)]$  considered above. The deepest point is the mean  $\mu(P)$ , as it is for the only other  $L_2$ -flavoured depth in the literature, namely the zonoid depth from [Koshevoy and Mosler \(1997\)](#).

It is useful to compare with the corresponding  $L_1$  quantities that are obtained by substituting  $|t|$  for  $t^2$  in the asymmetric  $L_2$  loss  $\rho_\alpha$  above. This leads to the usual quantiles  $q_\alpha(P)$  and to the underlying distribution function  $G(z) = \mathbb{P}[Z \leq z]$  for which the resulting depth  $D(z, P) := \min(G(z), 1 - G(z - 0))$  is the celebrated [Tukey \(1975\)](#) halfspace depth; throughout,  $\mathbb{P}$  refers to the probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  on which all random variables and random vectors are defined, and  $f(z - 0)$  stands for the limit of  $f(y)$  as  $y$  converges to  $z$  from below. The depth regions are therefore the interquantile intervals  $[q_\alpha(P), q_{1-\alpha}(P)]$  (with  $\alpha < 1/2$ ) that all contain the deepest point  $q_{1/2}(P)$ , the usual median. Obviously, the correspondance between univariate expectiles and expectile depth is as strong as the well-known one between usual quantiles and halfspace depth.

Now, far from restricting to  $\mathbb{R}$ , the halfspace depth was introduced as a device to order multivariate data points in  $\mathbb{R}^d$  and actually finds most of its applications there, such as, for instance, in the context of supervised classification; see, e.g., [Li et al., 2012](#), and the many papers that built on this work. It is therefore desirable to extend expectile depth to  $\mathbb{R}^d$  and to introduce a concept of multivariate expectiles bearing a connection with multivariate expectile depth that is as strong as in the univariate case. This is one of the objectives of this paper. We start by introducing the proposed multivariate expectiles.

### 3 Our multivariate expectiles

The first multivariate expectiles were defined in [Breckling and Chambers \(1988\)](#) as special cases of more general multivariate M-quantiles (the quantile and expectile particular cases were investigated in [Chaudhuri \(1996\)](#) and [Herrmann et al. \(2018\)](#), where the quantiles/expectiles are called “geometric”). Since then, several concepts of multivariate expectiles have been proposed. For the sake of completeness, we describe the multivariate expectiles above, as well as those resulting from [Breckling et al. \(2001\)](#) and [Kokic et al. \(2002\)](#), in Section S.1 of the supplement. For now, it is only important to mention that, possibly after an unimportant reparametrization, all aforementioned multivariate expectiles can be written as functionals  $P \mapsto \mathbf{e}_{\alpha,\mathbf{u}}(P)$  that take values in  $\mathbb{R}^d$  and are indexed by a scalar order  $\alpha \in (0, 1)$  and a direction  $\mathbf{u} \in \mathcal{S}^{d-1} := \{\mathbf{z} \in \mathbb{R}^d : \|\mathbf{z}\|^2 := \mathbf{z}'\mathbf{z} = 1\}$ ; here,  $P$  is a probability measure over  $\mathbb{R}^d$ . Typically,  $\mathbf{e}_{\alpha,\mathbf{u}}(P)$  does not depend on  $\mathbf{u}$  for  $\alpha = 1/2$ , and the resulting common location is the center of  $P$ .

In the univariate case, it is often important to know whether some test statistic takes a value below or above a given quantile, that is used as a critical value; in the multivariate case, point-valued quantiles cannot be used as critical values with vector-valued test statis-

tics, which suggests favouring *hyperplane-valued* multivariate quantiles. Since expectiles are also quantiles (of the transformation  $G$  of the original distribution; see (3)), this provides a motivation to propose multivariate expectiles that, unlike their competitors above, are hyperplane-valued rather than point-valued. To define the new multivariate expectiles, we introduce the collection  $\mathcal{P}_d$  of probability measures  $P$  over  $\mathbb{R}^d$  that have finite first moments and that do not attribute probability one to any hyperplane of  $\mathbb{R}^d$ . We then adopt the following definition.

**Definition 1.** Fix  $P \in \mathcal{P}_d$  and let  $\mathbf{Z}$  be a random  $d$ -vector with distribution  $P$ . For any  $\mathbf{u} \in \mathcal{S}^{d-1}$ , denote as  $P_{\mathbf{u}}$  the distribution of  $\mathbf{u}'\mathbf{Z}$ . Then, for any  $\alpha \in (0, 1)$  and  $\mathbf{u} \in \mathcal{S}^{d-1}$ , the order- $\alpha$  expectile of  $P$  in direction  $\mathbf{u}$  is the hyperplane

$$\pi_{\alpha, \mathbf{u}}(P) = \left\{ \mathbf{z} \in \mathbb{R}^d : \mathbf{u}'\mathbf{z} = e_{\alpha}(P_{\mathbf{u}}) \right\},$$

where  $e_{\alpha}(P_{\mathbf{u}})$  is the order- $\alpha$  expectile of  $P_{\mathbf{u}}$ ; see (3). The upper-halfspace  $H_{\alpha, \mathbf{u}}(P) = \left\{ \mathbf{z} \in \mathbb{R}^d : \mathbf{u}'\mathbf{z} \geq e_{\alpha}(P_{\mathbf{u}}) \right\}$  will then be called order- $\alpha$  expectile halfspace of  $P$  in direction  $\mathbf{u}$ .

These expectile hyperplanes  $\pi_{\alpha, \mathbf{u}}$  are linked in a straightforward way to the direction  $\mathbf{u}$ , as they are simply orthogonal to  $\mathbf{u}$ . In contrast, the point-valued competitors mentioned above typically depend on  $\mathbf{u}$  in an intricate way, and in particular  $\mathbf{e}_{\alpha, \mathbf{u}}(P)$  usually does not belong to the halffine with direction  $\mathbf{u}$  originating from the corresponding center (see above). Note that the “intercepts” of our hyperplanes are the univariate expectiles  $e_{\alpha}(P_{\mathbf{u}})$  of the projection  $\mathbf{u}'\mathbf{Z}$  of  $\mathbf{Z}$  onto  $\mathbf{u}$ , hence also allow for a direct interpretation.

Unlike their competitors, our multivariate expectiles are equivariant under affine transformations. We have the following result.

**Theorem 1.** Fix  $P \in \mathcal{P}_d$ . Let  $\mathbf{A}$  be an invertible  $d \times d$  matrix and  $\mathbf{b}$  be a  $d$ -vector. Then, for any  $\alpha \in (0, 1)$  and  $\mathbf{u} \in \mathcal{S}^{d-1}$ ,

$$\pi_{\alpha, \mathbf{u}_{\mathbf{A}}}(P_{\mathbf{A}, \mathbf{b}}) = \mathbf{A}\pi_{\alpha, \mathbf{u}}(P) + \mathbf{b} \quad \text{and} \quad H_{\alpha, \mathbf{u}_{\mathbf{A}}}(P_{\mathbf{A}, \mathbf{b}}) = \mathbf{A}H_{\alpha, \mathbf{u}}(P) + \mathbf{b},$$

where  $\mathbf{u}_\mathbf{A} := (\mathbf{A}^{-1})'\mathbf{u}/\|(\mathbf{A}^{-1})'\mathbf{u}\|$  and where  $P_{\mathbf{A}, \mathbf{b}}$  is the distribution of  $\mathbf{A}\mathbf{Z} + \mathbf{b}$  when  $\mathbf{Z}$  is a random  $d$ -vector with distribution  $P$ .

At first sight, a possible advantage of point-valued expectiles  $\mathbf{e}_{\alpha, \mathbf{u}}(P)$  is that they naturally generate contours and regions. More precisely, they allow considering, for any  $\alpha \in (0, \frac{1}{2}]$ , the order- $\alpha$  expectile contour  $\{\mathbf{e}_{\alpha, \mathbf{u}}(P) : \mathbf{u} \in \mathcal{S}^{d-1}\}$ , the interior part of which is then the corresponding order- $\alpha$  expectile region. Our hyperplane-valued expectiles, however, also provide centrality regions, hence corresponding contours.

**Definition 2.** Fix  $P \in \mathcal{P}_d$ . For any  $\alpha \in (0, \frac{1}{2}]$ , the order- $\alpha$  expectile region of  $P$  is  $R_\alpha(P) = \bigcap_{\mathbf{u} \in \mathcal{S}^{d-1}} H_{\alpha, \mathbf{u}}(P)$  and the corresponding order- $\alpha$  contour is the boundary  $\partial R_\alpha(P)$  of  $R_\alpha(P)$ .

Substituting the quantiles  $q_\alpha(P_{\mathbf{u}})$  for the expectiles  $e_\alpha(P_{\mathbf{u}})$  in Definition 1 would provide the quantile hyperplanes/halfspaces from Paindaveine and Šiman (2011), which, through the construction in Definition 2, would provide the halfspace Tukey (1975) depth regions (see Theorem 2 in Kong and Mizera, 2012). The proposed regions can therefore be regarded as expectile analogs of the Tukey quantile ones. In line with this, the quantile regions in the univariate case  $d = 1$  are the interquantile intervals  $[q_\alpha(P), q_{1-\alpha}(P)]$ , whereas the regions from Definition 2 for  $d = 1$  reduce to the interexpectile intervals  $[e_\alpha(P), e_{1-\alpha}(P)]$ .

Since expectiles are monotone non-decreasing functions of their order  $\alpha$ , the regions  $R_\alpha(P)$  are nested (the larger  $\alpha$ , the smaller the region). The proposed regions enjoy many nice properties compared to their competitors resulting from point-valued expectiles, as we show on the basis of Theorem 2 below. To state the result, we define the *C-support* of  $P$  as  $C(P) := \{\mathbf{z} \in \mathbb{R}^d : \mathbb{P}[\mathbf{u}'\mathbf{Z} \leq \mathbf{u}'\mathbf{z}] > 0 \text{ for any } \mathbf{u} \in \mathcal{S}^{d-1}\}$ , where the random  $d$ -vector  $\mathbf{Z}$  has distribution  $P$ . Clearly,  $C(P)$  can be thought of as the convex hull of  $P$ 's support.

**Theorem 2.** Fix  $P \in \mathcal{P}_d$ . Then, for any  $\alpha \in (0, \frac{1}{2}]$ , the region  $R_\alpha(P)$  is a convex and compact subset of  $C(P)$ . Moreover,  $R_\alpha(P_{\mathbf{A}, \mathbf{b}}) = \mathbf{A}R_\alpha(P) + \mathbf{b}$  for any invertible  $d \times d$  matrix  $\mathbf{A}$  and  $d$ -vector  $\mathbf{b}$ .

No competing expectile regions combine these properties. For instance, the expectile regions from Herrmann et al. (2018) may extend far beyond the convex hull of the support, just like the original M-quantile regions from Breckling and Chambers (1988); see below. This was actually the motivation for the alternative proposals in Breckling et al. (2001) and Kokic et al. (2002). The regions introduced in these two papers, however, may fail to be convex, which is unnatural. More generally, none of the competing expectile regions are affine-equivariant. This may result in quite pathological behaviors. In the quantile case, for instance, Theorem 2.2 from Girard and Stupler (2017) implies that, if  $P$  is elliptically symmetric with density  $f$ , then, for small  $\alpha$ , the geometric quantile contours from Chaudhuri (1996) are “orthogonal” to the principal component structure of  $P$ , in the sense that these contours are furthest (resp., closest) to the symmetry center of  $P$  in the last (resp., first) principal direction. Our empirical results below show that geometric expectile regions suffer the same pathological behaviour. In contrast, the affine-equivariance result in Theorem 2 ensures that the shape of our expectile regions will match the principal component structure of  $P$ .

We illustrate this on the “cigar-shaped” data example from Breckling et al. (2001) and Kokic et al. (2002), for which  $P = P_n$  is the empirical probability measure associated with  $n = 200$  bivariate observations whose  $x$ -values form a uniform grid in  $[-1, 1]$  and whose  $y$ -values are randomly drawn from the normal distribution with mean 0 and variance .01. Figure 1 draws, for several orders  $\alpha$ , the various expectile contours mentioned in the previous paragraph (our contours  $\partial R_\alpha(P_n)$  were computed by replacing the intersection in Definition 2 by an intersection over 500 equispaced directions  $\mathbf{u}$  in  $\mathcal{S}^1$ ; all competing

contours require a similar discretization). For the sake of comparison, we also show the corresponding quantile contours. Clearly, the geometric expectiles from Herrmann et al. (2018) (as well as their quantile counterparts from Chaudhuri, 1996) extend much beyond the convex hull of the data points. Extreme geometric expectiles also show the aforementioned pathological behavior relative to the principal component structure of  $P$ . Finally, the outer expectile (and quantile) regions from Breckling et al. (2001) and Kokic et al. (2002) are non-convex in most cases. In line with Theorem 2, our expectile regions and contours do not exhibit these deficiencies.

## 4 Halfspace expectile depth

The expectile regions  $R_\alpha(P)$  are *centrality regions*, in the sense that they group locations  $\mathbf{z}$  according to their centrality with respect to  $P$ . This leads to the following expectile-flavoured concept of *statistical depth*.

**Definition 3.** Fix  $P \in \mathcal{P}_d$ . Then, the halfspace expectile depth (HED) of  $\mathbf{z}$  with respect to  $P$  is  $HED(\mathbf{z}, P) = \sup\{\alpha \in (0, \frac{1}{2}] : \mathbf{z} \in R_\alpha(P)\}$  (where we define  $\sup \emptyset$  as zero).

In this section, we investigate the properties of this new depth and we show its advantages over its quantile counterpart, namely the Tukey (1975) halfspace depth.

### 4.1 Basic properties and links with other depths

For any depth, the corresponding depth regions, that collect locations with depth larger than or equal to  $\alpha$ , are of particular interest and reveal interesting features of the underlying distribution (Section 5.2 will illustrate this in a regression framework). The following result

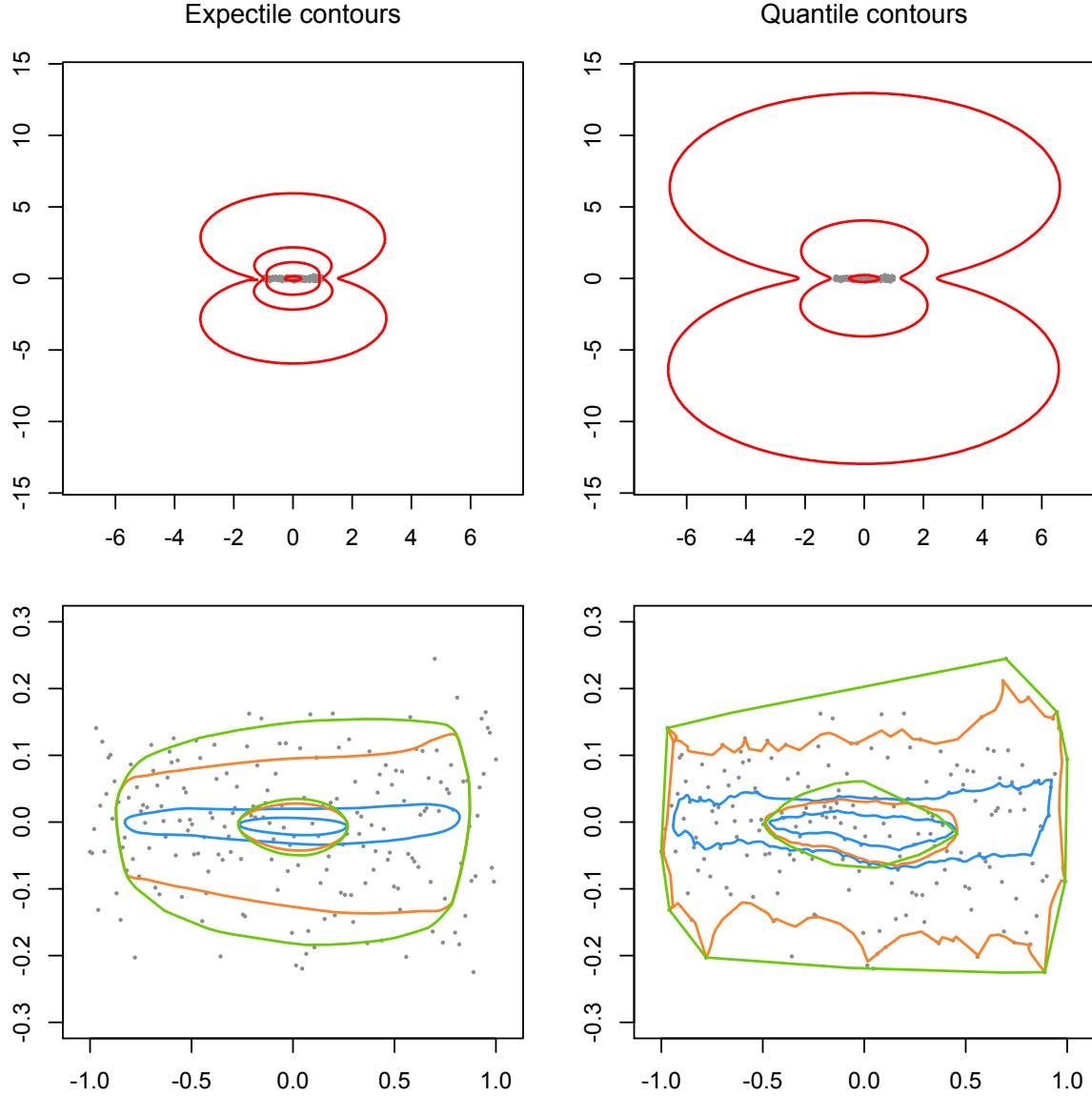


Figure 1: (Top:) the geometric expectile contours from [Herrmann et al. \(2018\)](#) (left) and geometric quantile contours from [Chaudhuri \(1996\)](#) (right), for the cigar-shaped data described in Section 3 and for  $\alpha = .00001, .0005, .005$ , and  $.25$  (for the smallest  $\alpha$ , the quantile contour is outside the plot). (Bottom left:) the expectile contours from [Breckling et al. \(2001\)](#) (blue), the  $(\delta = 10)$ -version of the [Kokic et al. \(2002\)](#) expectile contours (orange), and the proposed expectile contours (green), for the same data and for  $\alpha = 1/n = .005$  and  $.25$ ; we use the same value of  $\delta$  as in [Kokic et al. \(2002\)](#). (Bottom right:) the quantile versions of the contours in the bottom left panel. In each panel, the ( $n = 200$ ) data points are shown in grey.

shows that halfspace expectile depth regions coincide with the centrality regions introduced in the previous section.

**Theorem 3.** *Fix  $P \in \mathcal{P}_d$ . Then, for any  $\alpha \in (0, \frac{1}{2}]$ , the halfspace expectile depth region  $\{\mathbf{z} \in \mathbb{R}^d : HED(\mathbf{z}, P) \geq \alpha\}$  coincides with  $R_\alpha(P)$ .*

This result has several interesting consequences. First, it confirms that the halfspace expectile depth can be regarded as the expectile version of the halfspace depth (which of course justifies the terminology): substituting quantiles for expectiles in Definition 3, that is, substituting the quantiles  $q_\alpha(P_{\mathbf{u}})$  for the expectiles  $e_\alpha(P_{\mathbf{u}})$  in the halfspaces  $H_{\alpha,u}(P)$  leading to the regions  $R_\alpha(P)$  there, would indeed provide the halfspace depth, since we showed in Section 3 that the quantile version of the centrality regions  $R_\alpha(P)$  are the halfspace depth regions. Another consequence of Theorem 3 is that the HED is, like the halfspace depth, quasi-concave (this follows from the convexity of the regions  $R_\alpha(P)$ ). A further property the HED shares with the halfspace depth is that it is a *statistical depth function* in the sense of Zuo and Serfling (2000). We indeed have the following result.

**Theorem 4.** *For any  $P \in \mathcal{P}_d$ ,  $HED(\mathbf{z}, P)$  satisfies the following properties:* (i) (affine invariance:) *for any invertible  $d \times d$  matrix  $\mathbf{A}$  and  $d$ -vector  $\mathbf{b}$ ,  $HED(\mathbf{Az} + \mathbf{b}, P_{\mathbf{A},\mathbf{b}}) = HED(\mathbf{z}, P)$ , where  $P_{\mathbf{A},\mathbf{b}}$  was defined in Theorem 1;* (ii) (maximality at the center:) *if  $P$  is centrally symmetric about  $\boldsymbol{\theta}$  (i.e.,  $P[\boldsymbol{\theta} + B] = P[\boldsymbol{\theta} - B]$  for any  $d$ -Borel set  $B$ ), then  $HED(\boldsymbol{\theta}, P) \geq HED(\mathbf{z}, P)$  for any  $d$ -vector  $\mathbf{z}$ ;* (iii) (monotonicity along rays:) *if  $\boldsymbol{\theta}$  has maximum HED with respect to  $P$ , then, for any  $\mathbf{u} \in \mathcal{S}^{d-1}$ ,  $r \mapsto HED(\boldsymbol{\theta} + r\mathbf{u}, P)$  is monotone non-increasing in  $r (\geq 0)$ ;* (iv) (vanishing at infinity:) *as  $\|\mathbf{z}\| \rightarrow \infty$ ,  $HED(\mathbf{z}, P) \rightarrow 0$ .*

Other expectile depths could be defined from competing expectile regions by using the construction in Definition 3, but all of them would fail to meet one of the most classical requirements for depth, namely affine invariance (Theorem 4(i)). While it is not related

to expectiles, it is natural to also consider here the zonoid depth from Koshevoy and Mosler (1997), since it is maximized at the mean vector, hence is usually regarded as a depth of an  $L_2$ -nature. To get some insight on how the HED compares with the halfspace and zonoid depths, we consider the univariate example for which  $P$  is uniform over the interval  $\mathcal{I} = [0, 1]$  and the one for which it is uniform over the pair  $\{0, 1\}$ , which leads to

$$HED(z, P) = \frac{\min(z^2, (1-z)^2)}{z^2 + (1-z)^2} \mathbb{I}[z \in \mathcal{I}] \quad \text{and} \quad HED(z, P) = \min(z, 1-z) \mathbb{I}[z \in \mathcal{I}], \quad (4)$$

respectively; we refer to Koshevoy and Mosler (1997) for the corresponding expressions of the halfspace and zonoid depths. Figure 2, that plots these three depths for both  $P$ , suggests that the HED is smoother than the halfspace and zonoid depths and that it avoids ties in the support of the distribution (unlike the halfspace depth). This will be supported by our results in Section 4.2 below. Also, the zonoid depth regions for the uniform distribution over  $[0, 1]$  are the interquantile intervals  $[q_\alpha(P), q_{1-\alpha}(P)] = [\frac{\alpha}{2}, 1 - \frac{\alpha}{2}]$ , which is not so natural for a depth that has an  $L_2$  flavour (recall that the HED regions, irrespective of  $P$ , are rather the,  $L_2$  in nature, interexpectile intervals  $[e_\alpha(P), e_{1-\alpha}(P)]$ ). Multivariate examples will be considered in the supplementary materials (Section S.2).

The discussion below Theorem 3 shows that the HED bears strong links with the halfspace depth. As we now explain, however, the HED also shows key distinctive properties that should be most appealing to practitioners.

## 4.2 Distinctive properties

For halfspace depth, which is an  $L_1$ -concept, the deepest location is not always unique and the usual unique representative of the set of deepest locations, namely the Tukey median, is a multivariate extension of the univariate median. Also, the depth of the Tukey median may depend on  $P$ , which is inconvenient in practice (see below). The,  $L_2$  in nature, HED

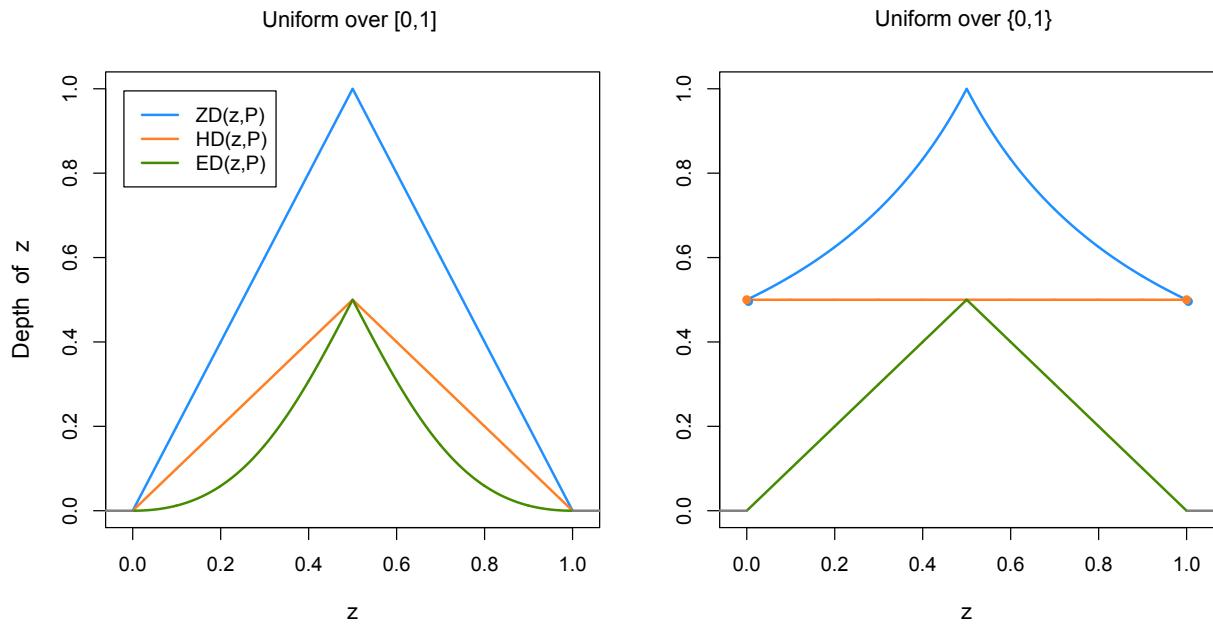


Figure 2: Plots, as functions of  $z$ , of the zonoid depth  $ZD(z, P)$  (blue), of the halfspace depth  $HD(z, P)$  (orange) and of the halfspace expectile depth  $HED(z, P)$  (green), when  $P$  is the uniform distribution over the interval  $[0, 1]$  (left) and the uniform distribution over the pair  $\{0, 1\}$  (right). For both  $P$ , all depth functions take value zero outside  $[0, 1]$ .

is much different in these respects.

**Theorem 5.** *For any  $P \in \mathcal{P}_d$ , the expectile depth  $HED(\mathbf{z}, P)$  is uniquely maximized at  $\mathbf{z} = \boldsymbol{\mu}(P) := E[\mathbf{Z}]$  (where  $\mathbf{Z}$  is a random  $d$ -vector with distribution  $P$ ) and the corresponding maximum depth is  $HED(\boldsymbol{\mu}(P), P) = 1/2$ .*

HED regions therefore always provide nested regions around the mean vector  $\boldsymbol{\mu}(P)$ , which should be appealing to practitioners. Also, the fact that the corresponding maximal depth is always  $1/2$  will allow practitioners to better interpret what it means that another location would have HED equal to, say,  $1/4$ : this would be, irrespective of the distribution, half as deep as the deepest location. In comparison, there is no way to evaluate the “relative” depth of a location with halfspace depth  $1/4$  without evaluating the depth of the Tukey median, which, at least in moderate to high dimensions, is notoriously difficult. Moreover, since the maximal HED is  $1/2$  for any  $P$ , a natural affine-invariant test for  $\mathcal{H}_0 : \boldsymbol{\mu}(P) = \boldsymbol{\mu}_0$ , where  $\boldsymbol{\mu}_0 \in \mathbb{R}^d$  is fixed, rejects  $\mathcal{H}_0$  for large values of  $T_n := (1/2) - HED(\boldsymbol{\mu}_0, P_n)$ , where  $P_n$  is the empirical probability measure associated with the sample  $\mathbf{Z}_1, \dots, \mathbf{Z}_n$  at hand. Due to the relation between the HED and the mean vector, this can be regarded as a nonparametric version of the Hotelling test.

We turn to another distinctive aspect of HED. As any statistical depth function, halfspace depth decreases monotonically when one moves away from a deepest location along any ray; see Theorem 4(iii). This monotonicity may fail to be strict, though (in the sample case, for instance, the halfspace depth is piecewise constant, hence is not strictly decreasing). In contrast, as Figure 2 already hinted, HED always offers a strict decrease (until, of course, the depth value zero is reached, if it is). We have the following result.

**Theorem 6.** *Fix  $P \in \mathcal{P}_d$  and  $\mathbf{u} \in \mathcal{S}^{d-1}$ . Then, letting  $r_{\mathbf{u}}(P) = \sup\{r > 0 : \boldsymbol{\mu}(P) + r\mathbf{u} \in C(P)\} \in (0, +\infty]$ , the function  $r \mapsto HED(\boldsymbol{\mu}(P) + r\mathbf{u}, P)$  is monotone strictly decreasing*

in  $[0, r_{\mathbf{u}}(P)]$  and  $HED(\boldsymbol{\mu}(P) + r\mathbf{u}, P) = 0$  for  $r \geq r_{\mathbf{u}}(P)$ .

A direct corollary is that if  $\mathbf{Z}_1, \dots, \mathbf{Z}_n$  are randomly sampled from a distribution  $P$  admitting a density, then the sample depths  $HED(\mathbf{Z}_i, P_n)$  will be pairwise different with probability one. This is precious in applications where ties in the depth values are to be avoided, such as, e.g., in supervised classification. For instance, the *max-depth classifiers* from Ghosh and Chaudhuri (2005) (see also Li et al., 2012) classify  $\mathbf{z}$  as arising from  $P_1$  rather than  $P_2$  if  $\mathbf{z}$  is deeper with respect to  $P_1$  than it is with respect to  $P_2$ , but obviously ties will lead to an unpleasant randomization.

The halfspace depth may also fail to be a continuous function of  $\mathbf{z}$ ; in particular, continuity will not hold in the sample case, due to the piecewise constant nature of the halfspace depth. In contrast, the HED is smooth even in the sample case, which should be appealing to practitioners who find it unpleasant that most depth values are not achieved at any  $\mathbf{z}$  or that a small change in  $\mathbf{z}$  has a strong impact on the corresponding depth.

**Theorem 7.** Fix  $P \in \mathcal{P}_d$ . Then, (i)  $\mathbf{z} \mapsto HED(\mathbf{z}, P)$  is uniformly continuous over  $\mathbb{R}^d$ ; (ii) for  $d = 1$ ,  $z \mapsto HED(z, P)$  is left- and right-differentiable over  $\mathbb{R}$ ; (iii) for  $d \geq 2$ , if  $P$  is smooth in a neighbourhood  $\mathcal{N}$  of  $\mathbf{z}_0$  (meaning that for any  $\mathbf{z} \in \mathcal{N}$ , any hyperplane containing  $\mathbf{z}$  has  $P$ -probability zero), then  $\mathbf{z} \mapsto HED(\mathbf{z}, P)$  admits directional derivatives at  $\mathbf{z}_0$  in all directions.

Figure 2 illustrates left- and right-differentiability of the HED, but it also reveals that, even inside the support of  $P$ , plain differentiability may fail (it does at  $z = 1/2$ ). The figure also shows that both the halfspace and zonoid depths may fail to be even continuous, hence are much less smooth than HED. Far from being a technical detail, smoothness of the HED will be an important asset for computational purposes. We discuss this in the next section, where we also present a last distinctive, actually most unexpected, property of the HED.

### 4.3 An alternative expression and computational aspects

Turning to computational aspects, the sample expectile regions  $R_\alpha(P_n)$  can be computed – or more precisely arbitrarily well approximated – by replacing the intersection in Definition 2 with an intersection over finitely many directions  $\mathbf{u}_\ell$ ,  $\ell = 1, \dots, L$ , with  $L$  large (a regular grid of directions can be used or, somewhat in the spirit of Cuesta-Albertos and Nieto-Reyes (2008), random directions can be generated). Many applications, however, do not require computing depth regions but rather the depth of a given location  $\mathbf{z}$  only. An important example is supervised classification through the max-depth approach; see Ghosh and Chaudhuri (2005) or Li et al. (2012). Obviously, the expression  $HED(\mathbf{z}, P) = \sup\{\alpha \in (0, \frac{1}{2}] : \mathbf{z} \in R_\alpha(P)\}$  allows one to compute the HED of  $\mathbf{z}$  from the corresponding regions, but it is extremely costly to have to compute many depth regions to evaluate the depth of a single  $\mathbf{z}$  only. This is an important motivation for the following result, which provides another expression for the HED.

**Theorem 8.** *For any  $\mathbf{z} \in \mathbb{R}^d$  and  $P \in \mathcal{P}_d$ ,*

$$HED(\mathbf{z}, P) = \min_{\mathbf{u} \in \mathcal{S}^{d-1}} G_{\mathbf{z}}^e(\mathbf{u}), \quad \text{with } G_{\mathbf{z}}^e(\mathbf{u}) := \frac{E[\mathbf{u}'(\mathbf{Z} - \mathbf{z}) | \mathbb{I}[\mathbf{u}'\mathbf{Z} \leq \mathbf{u}'\mathbf{z}]]}{E[\|\mathbf{u}'(\mathbf{Z} - \mathbf{z})\|]}, \quad (5)$$

where  $\mathbf{Z}$  has distribution  $P$ .

This result allows us to interpret the depth  $HED(\mathbf{z}, P)$  as the most severe *expectile* outlyingness of  $\mathbf{u}'\mathbf{z}$  with respect to the distribution of  $\mathbf{u}'\mathbf{Z}$ , in the same way

$$HD(\mathbf{z}, P) = \inf_{\mathbf{u} \in \mathcal{S}^{d-1}} G_{\mathbf{z}}^q(\mathbf{u}), \quad \text{with } G_{\mathbf{z}}^q(\mathbf{u}) := \mathbb{P}[\mathbf{u}'\mathbf{Z} \leq \mathbf{u}'\mathbf{z}], \quad (6)$$

defines the halfspace depth  $HD(\mathbf{z}, P)$  as the most severe *quantile* outlyingness of  $\mathbf{u}'\mathbf{z}$  with respect to the distribution of  $\mathbf{u}'\mathbf{Z}$ ; see Section 2 for the interpretation of these (expectile and quantile) scalar outlyingness measures in terms of  $G$  functions. This provides an intuitive

interpretation of  $HED(\mathbf{z}, P)$  in terms of depth and it also gives another insight on why the HED can be seen as the expectile variant of the halfspace depth.

Incidentally, Theorem 5 also states that a minimal direction  $\mathbf{u}$  always exists in (5), which directly follows from the continuity of  $G_{\mathbf{z}}^e(\cdot)$  (see Lemma S.2) and the compactness of  $\mathcal{S}^{d-1}$ . In contrast, a minimal direction does not always exist in (6). For instance, if  $\mathbf{z} = (1, 0)' \in \mathbb{R}^2$  and  $P = \frac{1}{2}P_1 + \frac{1}{2}P_2$ , where  $P_1$  is the bivariate standard normal distribution and  $P_2$  is the Dirac distribution at  $(1, 1)'$ , then no minimal direction exist in (6) (whereas  $\mathbf{u}_0 = (-1, 0)'$  is a minimal direction in (5)). This is another aspect for which the HED is easier to deal with than the halfspace depth.

Coming back to computational aspects, we now comment on how hard it is to evaluate  $HED(\mathbf{z}, P)$  and  $HD(\mathbf{z}, P)$  on the basis of (5) and (6), respectively. A key point is that (6) does not allow one to compute halfspace depth through standard algorithms such as Newton-Raphson-type methods. The reason is twofold: first, such iterative methods might converge to one of the many *local* minima of  $G_{\mathbf{z}}^q(\cdot)$ ; see the right panel of Figure 3. Second,  $G_{\mathbf{z}}^q(\cdot)$  is a piecewise constant, hence non-smooth, function in the sample case. In contrast, a further, most unexpected, distinctive property of the HED opens the door to a Newton-Raphson computation of this depth. More specifically, the following result ensures that the function  $G_{\mathbf{z}}^e(\cdot)$  that is to be minimized to compute the HED is not only smooth (see the proof of Theorem 7) but also never provides local-but-not-global minimizers.

**Theorem 9.** *Fix  $P \in \mathcal{P}_d$  and  $\mathbf{z} \in \mathbb{R}^d$  such that  $HED(\mathbf{z}, P) > 0$ . Assume that  $P[\Pi \setminus \{\mathbf{z}\}] = 0$  for any hyperplane  $\Pi$  containing  $\mathbf{z}$ . Let  $\mathbf{u}_0$  be an arbitrary minimizer of  $G_{\mathbf{z}}^e(\cdot)$  on  $\mathcal{S}^{d-1}$ . Let  $\mathbf{u}_t$ ,  $t \in [0, \pi]$ , be a geodesic path on  $\mathcal{S}^{d-1}$  from  $\mathbf{u}_0$  to  $\mathbf{u}_\pi = -\mathbf{u}_0$ . Then, there exist  $t_a, t_b$  with  $0 \leq t_a \leq t_b \leq \pi$  such that  $t \mapsto G_{\mathbf{z}}^e(\mathbf{u}_t)$  is constant over  $[0, t_a]$ , admits a strictly positive derivative at any  $t \in (t_a, t_b)$  (hence is strictly increasing over  $[t_a, t_b]$ ), and is constant over  $[t_b, \pi]$ .*

In Figure 3, we consider the case for which  $\mathbf{z} = (\frac{3}{4}, \frac{3}{4})'$  and  $P(\in \mathcal{P}_2)$  has independent  $\text{Exp}(1)$  marginals, and we plot the resulting functions  $\phi \mapsto G_{\mathbf{z}}^q((\cos \phi, \sin \phi)')$  and  $\phi \mapsto G_{\mathbf{z}}^e((\cos \phi, \sin \phi)')$  over  $[0, 2\pi]$ . Clearly, this illustrates the unexpected monotonicity property in Theorem 9 but also shows that the result may fail for halfspace depth. Jointly with the fact that, in the sample case,  $t \mapsto G_{\mathbf{z}}^e(\mathbf{u}_t)$  will be left- and right-differentiable, Theorem 9 opens the door to fast computation of HED, also in high dimensions, through Newton-Raphson-type methods. An algorithm evaluating HED in this way will be developed in a later work focusing on computational aspects.

## 5 Real data examples

In this section, we investigate how responsive the proposed expectile regions  $R_{\alpha}^{(n)}$  are to financial risk exposure linked with unforeseen catastrophic events, such as the 2007–2008 financial crisis or the COVID-19 pandemic. Like their benchmark quantile analogs, these centrality regions trim joint returns with extreme losses as well as those with extreme profits. In terms of riskiness, the required capital reserve against multivariate risk should cover any loss associated with joint returns inside  $R_{\alpha}^{(n)}$ , with a suitable security level  $\alpha$ . We will consider both unconditional regions (Section 5.1) and conditional ones (Section 5.2).

### 5.1 The non-regression setting

We first consider returns of several investment banks before and during the 2007–2008 financial crisis. We start with a “pre-crisis” dataset collecting daily returns for Morgan Stanley ( $Z_1$ ) and Lehman Brothers ( $Z_2$ ) from May 3rd, 1994, to June 29th, 2007 ( $n = 3,315$  trading days). The top left panel of Figure 4 shows the benchmark quantile contour of extreme order  $\alpha = .0003 (= 1/n)$  and the expectile contours of order  $\alpha \in \{.000005, .00005, .0001\}$ .

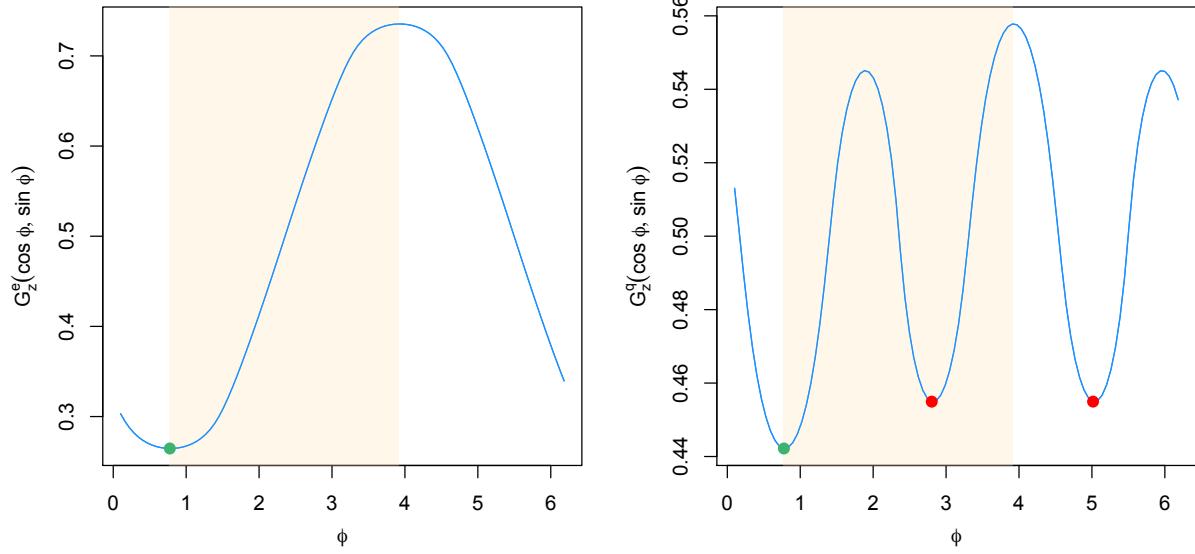


Figure 3: Plots of the HED objective function  $\phi \mapsto G_{\mathbf{z}}^e((\cos \phi, \sin \phi)')$  in (5) (left) and HD objective function  $\phi \mapsto G_{\mathbf{z}}^q((\cos \phi, \sin \phi)')$  in (6) (right), with  $\mathbf{z} = \left(\frac{3}{4}, \frac{3}{4}\right)'$  and  $P$  the probability measure over  $\mathbb{R}^2$  whose marginals are independent exponentials with mean one. Global minimizers providing the respective depths are marked in green, whereas local-but-not-global minimizers preventing the use of Newton-Raphson-type methods are marked in red. The shaded area shows the range considered in Theorem 9.

The most pessimistic of these trimming regions is the quantile one, which coincides with the convex hull of the data and expects the worst before crisis. However, the maxim “*expect the worst, and you won’t be disappointed*” fails during the crisis: the top right panel of Figure 4, that provides these centrality regions (with the same orders  $\alpha$ ) computed from returns ranging from May 3rd, 1994, to Lehman Brothers’ bankruptcy on September 15th, 2008, clearly reveals that expectile contours are much more alert to Lehman Brothers’ huge losses than the a priori more pessimistic quantile contour. We repeated the same exercise for Lehman Brothers ( $Z_1$ ) and Bear Stearns ( $Z_2$ ), with the same pre-crisis period and with a full time period ending on Bear Stearns’ collapse on May 30th, 2008. The results, that are displayed in the bottom panels of Figure 4, show that expectile contours here dominate the most pessimistic quantile contour not only in terms of responsiveness to Bear Stearns’ catastrophic losses but also in terms of reactivity to the few substantial profits.

## 5.2 The regression setting

To complement the “static” analysis conducted in Section 5.1, it is natural to try and estimate regression expectile regions, conditional on time. Denoting as  $P_x$  the conditional distribution of the  $d$ -variate response vector  $\mathbf{Y}$  given time  $X = x$ , our interest then lies in the conditional expectile regions  $R_{\alpha,x} = R_\alpha(P_x)$ ,  $\alpha \in (0, \frac{1}{2}]$ , that are obtained by intersecting the expectile halfspaces  $H_{\alpha,\mathbf{u},x} = H_{\alpha,\mathbf{u}}(P_x)$  over  $\mathbf{u} \in \mathcal{S}^{d-1}$  (obviously, all previous theorems extend to the conditional setting when substituting  $P_x$  for  $P$ ). If a sample  $(X_1, \mathbf{Y}_1), \dots, (X_n, \mathbf{Y}_n)$  is available, then one may estimate  $R_{\alpha,x}$  by the region  $R_{\alpha,x}^{(n)}$  obtained by intersecting over  $\mathbf{u} \in \mathcal{S}^{d-1}$  the sample halfspaces  $H_{\alpha,\mathbf{u},x}^{(n)} = \{\mathbf{y} \in \mathbb{R}^d : \mathbf{u}'\mathbf{y} \geq e_{\alpha,\mathbf{u},x}^{(n)}\}$ , where  $e_{\alpha,\mathbf{u},x}^{(n)}$  is the estimate of  $e_\alpha(P_{\mathbf{u}'\mathbf{Y}|[X=x]})$  resulting from a single-output (linear or nonparametric) regression using the responses  $\mathbf{u}'\mathbf{Y}_1, \dots, \mathbf{u}'\mathbf{Y}_n$  and covariates  $X_1, \dots, X_n$ . This conditional approach trivially extends to arbitrary  $p$ -variate covariate vectors  $\mathbf{X}$ , and,

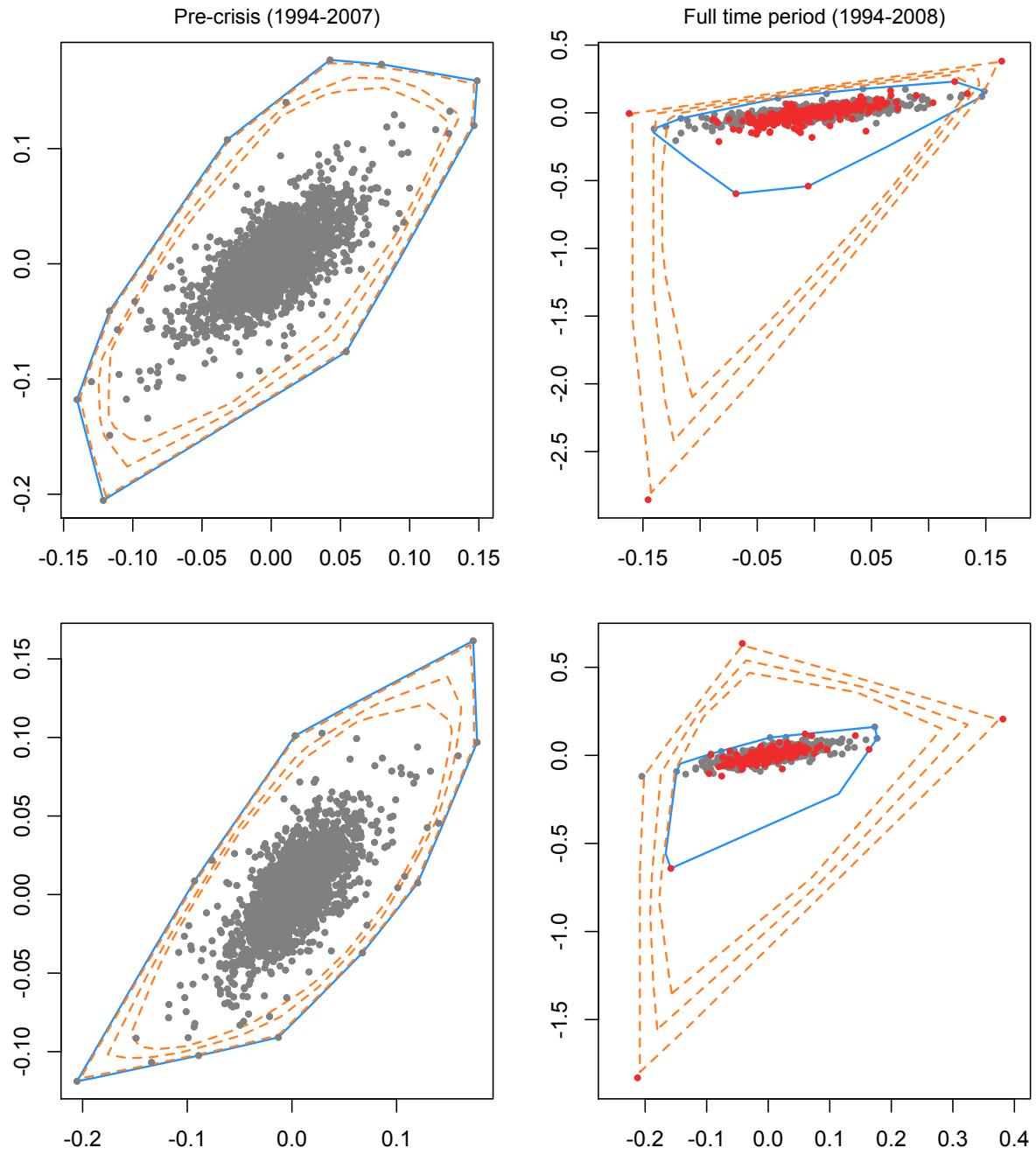


Figure 4: (Top:) Joint daily returns on Morgan Stanley ( $Z_1$ ) and Lehman Brothers ( $Z_2$ ), for the pre-crisis period (left) and the full time period (right); returns before (resp., during) the crisis are shown in grey (resp., red). The quantile contour associated with  $\alpha = .0003$  is plotted in solid blue and the expectile contours for  $\alpha \in \{.000005, .00005, .0001\}$  in dashed orange. (Bottom:) the same results for Lehman Brothers ( $Z_1$ ) and Bear Stearns ( $Z_2$ ).

of course, multiple-output *quantile* regression can be performed in the same way.

We illustrate this by considering the joint decline of Morgan Stanley and Goldman Sachs during the financial crisis. Our aim is to describe the unexpected joint extremal variation in these banks' equities ( $Y_1, Y_2$ , respectively) conditional on trading days ( $X$ ) from July 2, 2007, to September 19, 2008 (both investment banks converted themselves to bank holding companies on September 21st, 2008). The resulting marginal returns  $Y_{i1}$  and  $Y_{i2}$ ,  $i = 1, \dots, n = 309$ , are plotted in the top panel of Figure 5. The figure also provides the (tail) regression expectile contours  $R_{\alpha,x}^{(n)}$  with order  $\alpha = .0005$ , for several dates  $x$  selected as the 50%, 55%, ..., 95% empirical quantiles of  $X$ . For the sake of comparison, the corresponding conditional quantile contours, at order  $\alpha = .0008$ , are also provided. Both  $\alpha$ -values were chosen so that the earliest expectile and quantile contours (those associated with February 11, 2008) are roughly of the same size. To make the comparison as fair as possible, expectile and quantile contours were obtained by performing the same type of single-output (nonparametric) regression, namely those based on the same boosting principle using additive approaches from Sobotka and Kneib (2012) and Fenske et al. (2011) (we thus used the functions `expectreg.boost` and `quant.boost` from the R package *expectreg*).

The dynamics of the joint returns' evolution, which can hardly be visualized from the marginal time series displayed in the top panel of Figure 5, are nicely described by the resulting regression expectile and quantile contours. Both types of contours are qualitatively similar, indicating particularly that the joint returns were sensitive to the evolution of the systemic crisis: the contours indeed tend to spread out with time due to the high volatility and severity of infrequent returns, which is particularly seen for the last four contours (note that these seem to be more sensitive to severe losses than to high profits). Clearly, however, expectile contours much better flag these interesting features than quantile contours, and they produce more insightful fits of conditional location and spread of relevant

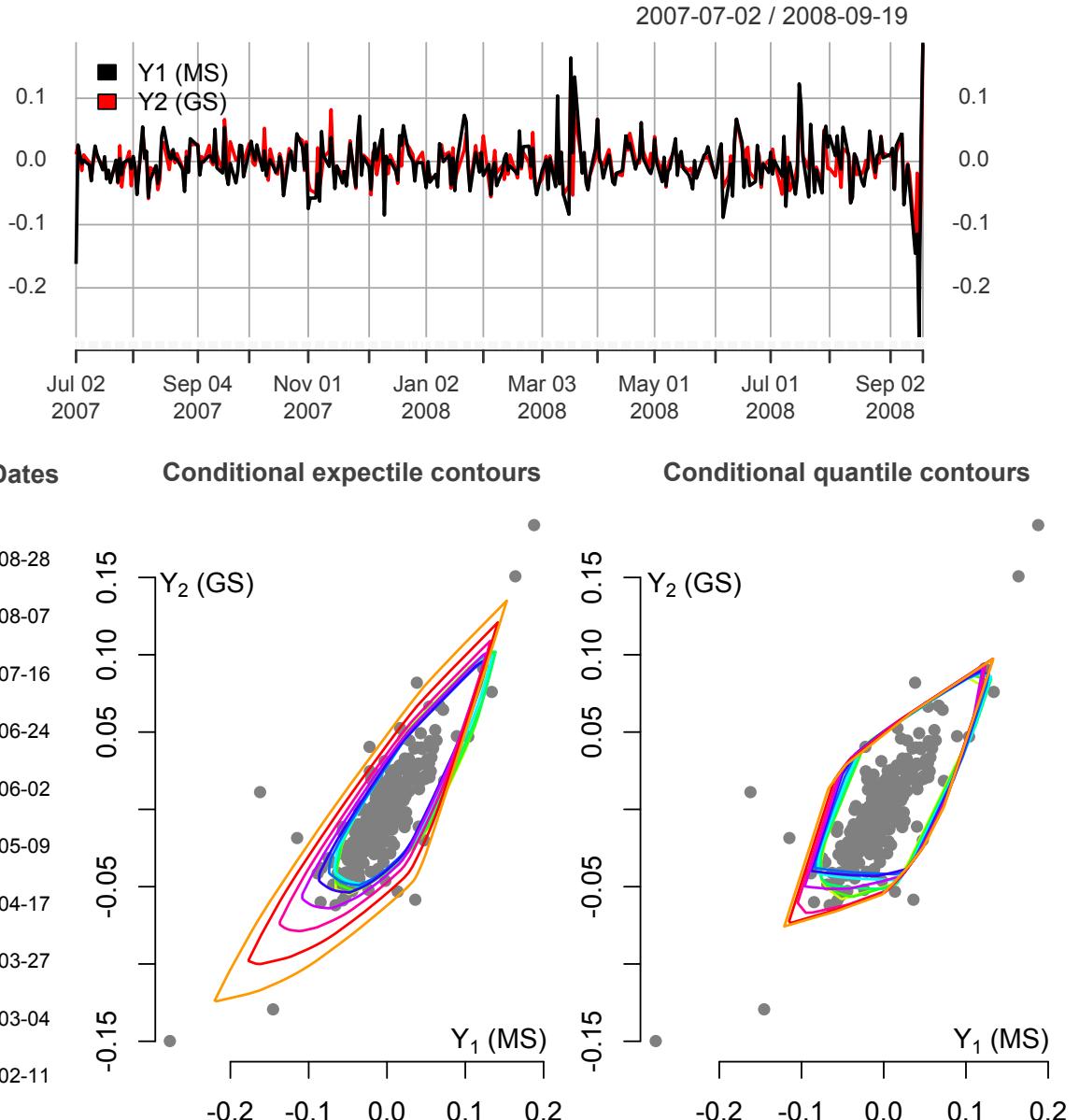


Figure 5: (Top:) daily returns of Morgan Stanley ( $Y_1$ ) and Goldman Sachs ( $Y_2$ ) from July 2, 2007, to September 19, 2008. (Bottom left:) the corresponding bivariate returns (gray), along with the resulting nonparametric regression expectile contours of order  $\alpha = .0005$ ,  
<sup>25</sup> conditional on trading days ( $X$ ), at 10 dates  $x$  selected as the 50% (light green), 55% (darker green), ..., 95% (orange) empirical quantiles of  $X$ . (Bottom right:) the corresponding results for regression quantiles of order  $\alpha = .0008$ ; see Section 5.2 for details.

data points. Multivariate expectiles may accordingly serve as an efficient instrument for detecting systemically risky firms.

As a final example, we analyze stock market index S&P500 ( $Y_1$ ) and Brent oil prices ( $Y_2$ ), conditional on trading days ( $X$ ), during the global spread of COVID-19 from December 31, 2019, to May 18, 2020 ( $n = 94$  trading days). Figure 6 provides marginal and bivariate returns, along with nonparametric regression expectile and quantile contours, at the extreme order  $\alpha = .0001$  in both cases, for dates  $x$  selected as the 5%, 10%, ..., 95% empirical quantiles of  $X$ . The results show that regression expectile contours are much smaller than their quantile analogs in the early stage of the pandemic, which corresponds to the dark blue contours, but that these contours are more sensitive to the day-on-day growth of the pandemic. Expectile contours capture more accurately both the location and spread of the data, and they better display the turbulent events experienced by financial and oil markets due to the twin shocks of the coronavirus pandemic and the Saudi-Russia oil price conflict: the initial expansion of these contours in blue (from January 7 to March 3) reflects the increasing fears over the spread of coronavirus. The following two contours in red (March, 9 and 16) successfully translate the panic seen in both financial and oil markets due (i) to the oil price war that started on March 8 and (ii) to the pandemic that went worldwide pandemic on March 11. Later (from March 23 to April 21), the expectile contours (in green) become less extreme in the  $Y_1$ -direction as the stock market index recovers somewhat progressively, while they continue expanding in the  $Y_2$ -direction as the oil prices remains down until reaching the historic crash on April 21. The decline of the final contours in all directions from dark to light gold marks the joint rebound of both oil prices and stock market index. Most clearly, these subtle joint dynamics much better show in expectile contours than in quantile ones.

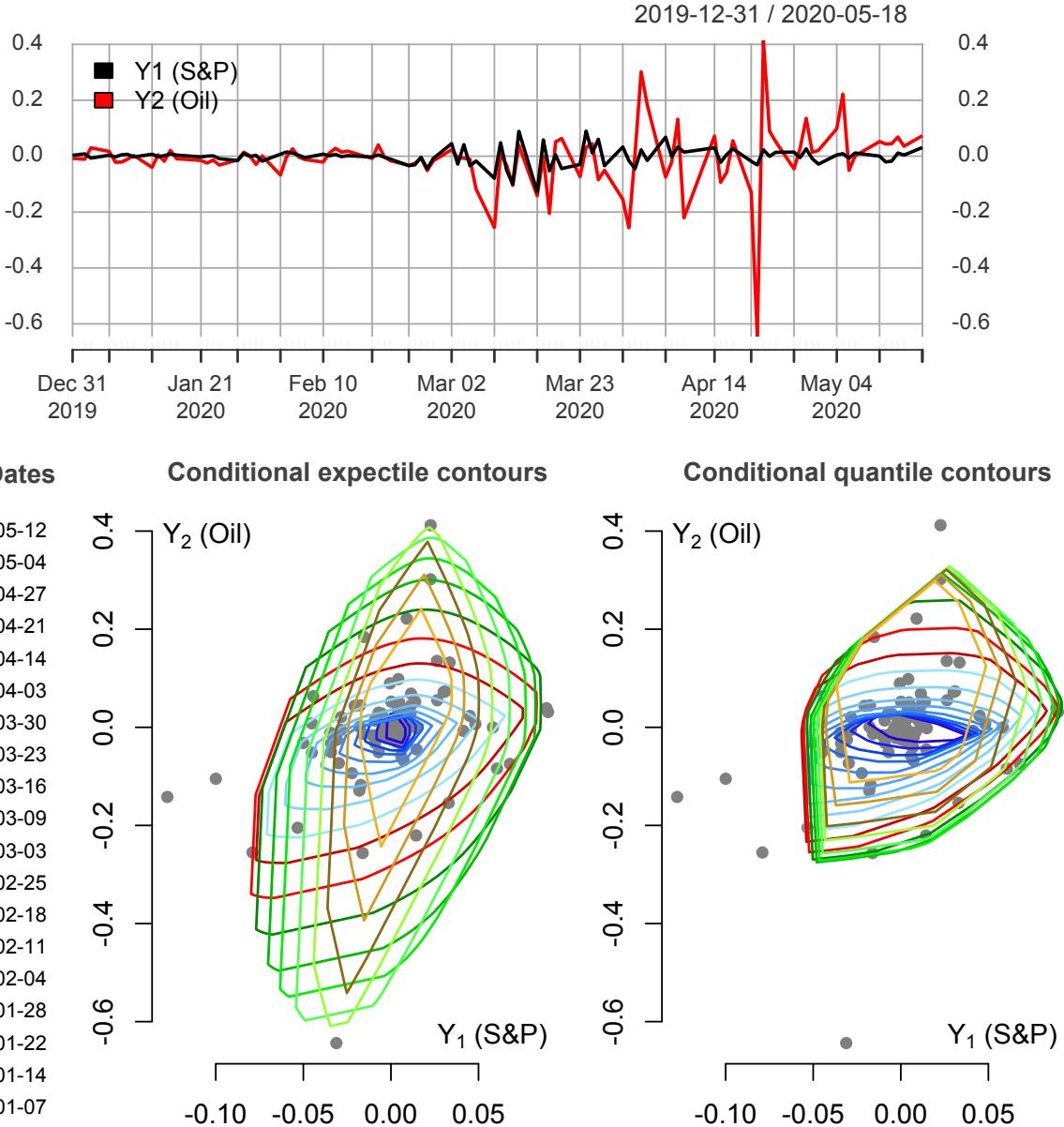


Figure 6: (Top:) daily S&P500 returns ( $Y_1$ ) and oil returns ( $Y_2$ ) from December 31, 2019, to May 18, 2020. (Bottom left:) the corresponding bivariate returns (gray), along with the resulting nonparametric regression expectile contours of order  $\alpha = .0001$ , conditional on 27 trading days ( $X$ ), at 19 dates  $x$  selected as the 5% (dark blue), 10% (lighter blue),  $\dots$ , 95% (gold) empirical quantiles of  $X$ . (Bottom right:) the corresponding results for regression quantiles, still of order  $\alpha = .0001$ ; see Section 5.2 for details.

## Acknowledgements

The authors acknowledge funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d’Avenir) program, grant ANR-17-EURE-0010. Support from the ANR under the grant ANR-19-CE40-0013-01 (ExtremReg project) is also gratefully acknowledged. Davy Paindaveine’s research is also supported by a research fellowship from the Francqui Foundation and by the Program of Concerted Research Actions (ARC) of the Université libre de Bruxelles.

**Supplement:** We provide the following further contributions in the supplementary materials: for the sake of completeness, we describe there some of the main competing multivariate expectile concepts (Section S.1). We compute expectile depth and expectile depth regions in several multivariate examples (Section S.2). We state asymptotic results for the proposed expectile depth (Section S.3). We illustrate on simulated data the proposed multiple-output expectile regression methods and show that these dominate the corresponding quantile-based methods in terms of crossings (Section S.4). We discuss the relation between multivariate expectiles and risk measures, and we show that our expectiles satisfy the coherency axioms of multivariate risk measures (Section S.5). Finally, we prove all results of the paper (Section S.6).

## References

- Breckling, J. and Chambers, R. (1988), “M-Quantiles,” *Biometrika*, 75, 761–771.
- Breckling, J., Kokic, P., and Lübke, O. (2001), “A note on multivariate M-quantiles,” *Statist. Probab. Lett.*, 55, 39–44.

- Carlier, G., Chernozhukov, V., and Galichon, A. (2016), “Vector quantile regression: An optimal transport approach,” *Ann. Statist.*, 44, 1165–1192.
- (2017), “Vector quantile regression beyond the specified case,” *J. Multivariate Anal.*, 161, 96–102.
- Chakraborty, B. (2003), “On multivariate quantile regression,” *J. Statist. Plann. Inference*, 110, 109–132.
- Chaudhuri, P. (1996), “On a geometric notion of quantiles for multivariate data,” *J. Amer. Statist. Assoc.*, 91, 862–872.
- Chavas, J.-P. (2018), “On multivariate quantile regression analysis,” *Stat. Methods Appl.*, 27, 365–384.
- Cheng, Y. and De Gooijer, J. (2007), “On the  $u$ th geometric conditional quantile,” *J. Statist. Plann. Inference*, 137, 1914–1930.
- Cousin, A. and Di Bernardino, E. (2013), “On multivariate extensions of value-at-risk,” *J. Multivariate Anal.*, 119, 32–46.
- (2014), “On multivariate extensions of conditional-tail-expectation,” *Insurance Math. Econom.*, 55, 272–282.
- Cuesta-Albertos, J. and Nieto-Reyes, A. (2008), “The random Tukey depth,” *Comput. Statist. Data Anal.*, 52, 4979–4988.
- Daouia, A., Girard, S., and Stupler, G. (2018), “Estimation of tail risk based on extreme expectiles,” *J. Roy. Statist. Soc. Ser. B*, 80, 263–292.
- De Rossi, G. and Harvey, H. (2009), “Quantiles, expectiles and splines,” *J. Econometrics*, 152, 179–185.

- Eilers, P. (2013), “Discussion: The beauty of expectiles,” *Stat. Model.*, 13, 317–322.
- Embrechts, P. and Hofert, M. (2014), “Statistics and quantitative risk management for banking and insurance,” *Annu. Rev. Stat. Appl.*, 1, 493–514.
- Fenske, N., Kneib, T., and Hothorn, T. (2011), “Identifying risk factors for severe childhood malnutrition by boosting additive quantile regression,” *J. Amer. Statist. Assoc.*, 106, 494–510.
- Ghosh, A. K. and Chaudhuri, P. (2005), “On maximum depth and related classifiers,” *Scand. J. Statist.*, 32, 327–350.
- Girard, S. and Stupler, G. (2017), “Intriguing properties of extreme geometric quantiles,” *REVSTAT*, 15, 107–139.
- Hallin, M., Lu, Z., Paindaveine, D., and Šiman, M. (2015), “Local bilinear multiple-output quantile/depth regression,” *Bernoulli*, 21, 1435–1466.
- Hallin, M., Paindaveine, D., and Šiman, M. (2010), “Multivariate quantiles and multiple-output regression quantiles: From  $L_1$  optimization to halfspace depth (with discussion),” *Ann. Statist.*, 38, 635–669.
- Herrmann, K., Hofert, M., and Mailhot, M. (2018), “Multivariate geometric expectiles,” *Scand. Actuar. J.*, 2018, 629–659.
- Jones, M. (1994), “Expectiles and M-quantiles are quantiles,” *Statist. Probab. Lett.*, 20, 149–153.
- Koenker, R. and Bassett, G. (1978), “Regression quantiles,” *Econometrica*, 46, 33–50.
- Kokic, P., Breckling, J., and Lübke, O. (2002), *A new definition of multivariate M-quantiles.*, Basel, Switzerland: Birkhäuser, chap. Statistical data analysis based on the

L1-norm and related methods, pp. 15–24.

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

Kong, L. and Mizera, I. (2012), “Quantile tomography: using quantiles with multivariate data,” *Statist. Sinica*, 22, 1589–1610.

Koshevoy, G. and Mosler, K. (1997), “Zonoid trimming for multivariate distributions,” *Ann. Statist.*, 25, 1998–2017.

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

Li, J., Cuesta-Albertos, J., and Liu, R. Y. (2012), “DD-Classifier: Nonparametric classification procedures based on DD-plots.” *J. Amer. Statist. Assoc.*, 107, 737–753.

Maume-Deschamps, V., Rullière, D., and Said, K. (2017a), “Asymptotic multivariate expectiles,” *ArXiv preprint arXiv:1704.07152v2*.

— (2017b), “Multivariate extensions of expectiles risk measures,” *Depend. Model.*, 5, 20–44.

McKeague, I., López-Pintado, S., Hallin, M., and Šiman, M. (2011), “Analyzing growth trajectories,” *J. Dev. Orig. Health Dis.*, 2, 322–329.

Newey, W. and Powell, J. (1987), “Asymmetric least squares estimation and testing,” *Econometrica*, 55, 819–847.

Paindaveine, D. and Šiman, M. (2011), “On directional multiple-output quantile regression,” *J. Multivariate Anal.*, 102, 193–212.

Schnabel, S. and Eilers, P. (2009), “Optimal expectile smoothing,” *Comput. Stat. Data Anal.*, 53, 4168–4177.

- Schulze Waltrup, L., Sobotka, F., Kneib, T., and Kauermann, G. (2015), “Expectile and quantile regression - David and Goliath?” *Stat. Model.*, 15, 433–456.
- Serfling, R. J. (2002), “Quantile functions for multivariate analysis: approaches and applications,” *Stat. Neerl.*, 56, 214–232.
- Sobotka, F. and Kneib, T. (2012), “Geoadditive expectile regression,” *Comput. Stat. Data Anal.*, 56, 755–767.
- Tukey, J. W. (1975), “Mathematics and the picturing of data,” in *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974)*, Vol. 2, Canad. Math. Congress, Montreal, Que., pp. 523–531.
- Waldmann, E. and Kneib, T. (2015), “Bayesian bivariate quantile regression,” *Stat. Model.*, 15, 326–344.
- Wei, Y. (2008), “An approach to multivariate covariate-dependent quantile contours with application to bivariate conditional growth charts,” *J. Amer. Statist. Assoc.*, 103, 397–409.
- Yao, Q. and Tong, H. (1996), “Asymmetric least squares regression and estimation: a nonparametric approach,” *J. Nonparametr. Stat.*, 6, 273–292.
- Zuo, Y. and Serfling, R. (2000), “General notions of statistical depth function,” *Ann. Statist.*, 28, 461–482.