Convergence of Multivariate Quantile Surfaces

Adil Ahidar-Coutrix, Philippe Berthet

Convergence of Multivariate Quantile Surfaces

ADIL AHIDAR-COUTRIX* PHILIPPE BERTHET†
Institut de Mathématiques de Toulouse
Université Paul Sabatier

December 6, 2016

Abstract

We define the quantile set of order $\alpha \in [1/2, 1)$ associated to a law P on $\mathbb{R}^d$ to be the collection of its directional quantiles seen from an observer $O \in \mathbb{R}^d$. Under minimal assumptions these star-shaped sets are closed surfaces, continuous in (O, α) and the collection of empirical quantile surfaces is uniformly consistent. Under mild assumptions – no density or symmetry is required for P – our uniform central limit theorem reveals the correlations between quantile points and a non asymptotic Gaussian approximation provides joint confident enlarged quantile surfaces. Our main result is a dimension free rate $n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}$ of Bahadur-Kiefer embedding by the empirical process indexed by half-spaces. These limit theorems sharply generalize the univariate quantile convergences and fully characterize the joint behavior of Tukey half-spaces.

1 Introduction

1.1 Short presentation

Let $\{X_n\}$ be a sequence of independent random vectors in $\mathbb{R}^d$ defined on a probability space $(\Omega, \mathcal{T}, \mathbb{P})$ and having the same law $P = \mathbb{P}^X$. Many procedures in multivariate data analysis have been proposed to picture out the structure of the data cloud $X_1, ..., X_n$ and distinguish between inner points, outer points and outliers. In particular it is worth mentioning generalized quantiles ([12],[19],[28],[30]), data depth ([16],[24],[23],[34],[35]), level sets ([27]), Tukey contours ([10],[21],[26],[33]), modal set estimation ([3],[25],[27],[28]), k-means

*adil.ahidar@math.univ-toulouse.fr
†philippe.berthet@math.univ-toulouse.fr

([7],[8]), trimming ([22],[26], [29]), quantile regression ([15]) among many others. The underlying generic problem is to infer about the mass localization of P in $\mathbb{R}^d$ – modal regions, support, main mass directions. Since probabilities and locations come into play together, the need of multivariate quantiles arises naturally. Now, the univariate α-th quantile can be defined in many ways, hence as many multivariate generalizations can be proposed in terms of points, vectors or sets satisfying some equation involving α.

The inference paradigm we promote below uses what we call quantile surfaces. They are defined in a purely nonparametric way, always exist and satisfy sharp convergence properties without too restrictive hypotheses. In this paper we focus on quantile surfaces built from half-spaces probabilities, so that our results can be applied to statistical procedures based on the popular Tukey half-spaces. Flexible extensions are studied in companion works, with applications to goodness of fit tests, depth vector fields and Lorens-Gini and Wasserstein type distances.

The paper is organized as follows. In Section 1 we discuss motivation and compare our approach to the main existing ones. In Section 2 we recall the limit theorems for univariate quantiles we intend to generalize. Then we provide notation, definitions and basic properties of the deterministic and empirical quantile surfaces, with a few illustrations and comments. Our results are stated in Section 3. Section 4 is devoted to proving continuity, uniform consistency, uniform weak convergence, strong approximation and a dimension free Bahadur-Kiefer representation of quantile surfaces.

1.2 Basic principles

It is important to point out that we depart from the following classical ideas, which have been extensively exploited.

It seems commonly admitted that localizing mass requires first a well defined mass center $M = M(P)$. On $\mathbb{R}$ the median corresponds to a robust central location M from where nested inter-quantiles intervals can grow up. In $\mathbb{R}^d$ it is then tempting to characterize some median point M, typically through a global minimization of some centrality expectation function. Seen from M the support of an unimodal P can be divided into central, inner, outer and extreme regions in a nested way. Such a contour description can be achieved by two main basic principles.

The depth principle consists in associating a real value to each point $O \in \mathbb{R}^d$, with a maximum at some mass center M. The latter typically depends on a notion of central or angular symmetry and depth contours stand as level sets of some depth function depending on P and M.

The quantile principle consists in associating a set of points to a value $\alpha \in (0,1)$. Typical quantile sets are selected among a small entropy collection of sets by means of argmax estimation, and centering sets at M helps making them nested like contours.

Outer spatial quantile sets or less deep contours are used to characterize outliers and build trimmed areas before processing, for sake of robustness. Inner spatial quantile sets or deeper contours are used to depict central regions of the support of P. In this spirit the depth axioms are formalized in [34]. Other approaches provide a similar center-outward ordering of points. Note that centered quantile sets have a probability α whereas depth contours may or not rely on α-th quantiles of some associated real valued random variable. Even when α is not a probability, contours require a central median point to cross directions. This is the case in [19] where the inverse of a multivalued function is used to represent directional quantiles.

1.3 Motivation

The limitations of the framework of quantile sets and depth contours motivate our notion of arbitrarily anchored quantile surfaces.

Firstly, focusing on a unique mass center $M \in \mathbb{R}^d$ could be misleading and excludes interesting cases like mixtures or low dimensional supports. We would like to depict mass localization beyond the center-outward case, with no need of any objective center $M = M(P)$. We thus suggest to learn about P by moving a subjective viewpoint $O \in \mathbb{R}^d$ – like turning around a geometrical structure to see all faces rather than observing it from a central point inside. If P is M-symmetric then all expected properties hold at $O = M$ and we recover radial quantiles.

Secondly, few limit theorems are available besides consistency compared to the variety of proposed methods. We would like to generalize the sharpest limit theorems on univariate quantiles. Using directional projections seen from $O \in \mathbb{R}^d$ allows to go back to $\mathbb{R}$ and our main contribution is to control them jointly.

Thirdly, known results hold under restrictive assumptions on P. In particular, P often has density and contiguous support or is regular with respect to the indexing sets or a depth function. We would like to impose no stronger assumptions than for univariate quantiles. Moreover in higher dimension the statistical dependency of the coordinates of X could make P very concentrated around low dimensional manifolds or geometrical structures, and such a sparsity means no density. Thus a special effort is made to relax the density and support requirement.

Sometimes theoretical methods have unrealistic computational aspects. Consider for instance plug-in procedures such as computing level sets after a d-dimensional density estimation. The quantile surfaces we introduce are quickly computed by orthogonal projections and confident bands follow from our Gaussian approximation by tractable Monte-Carlo simulations.

Lastly, in our opinion a non reductive notion of $\alpha-$th quantile set in $\mathbb{R}^d$ should be $(d-1)$-dimensional and informative depth should be d-dimensional. This is what quantile surfaces and their depth vector fields are.

1.4 A new principle

Imagine an observer located in $O \in \mathbb{R}^d$ looking at the sample $X_1, ..., X_n$ in all directions $u \in \mathbb{S}_{d-1}$ where $\mathbb{S}_{d-1}$ is the unit sphere of $\mathbb{R}^d$. Let him picture out the data cloud in $\mathbb{R}^d$ from O by drawing the collection of u-directional α-th quantile point $Q_n(O,u,\alpha) = O + Y_n(O,u,\alpha)u$ where $Y_n(O,u,\alpha)$ is the univariate α-th quantile of the projected sample $\langle X_i - O, u\rangle$ on the oriented line (O,u), and $\langle .,. \rangle$ is the inner product. We thus associate a star-shaped quantile set $Q_n(O,\alpha)$ to every $(O,\alpha) \in \mathbb{R}^d \times (1/2,1)$. This is a multivariate quantile principle with no mass center, no α-mass quantile set and no global contour.

Under minimal assumptions the sets $Q(O,\alpha)$ associated to P are nested surfaces starting at O then extending toward modal areas. For fixed O, increasing α indicates main mass directions and concentrations. For fixed α, the deepest is O the "smaller" is $Q(O,\alpha)$. This leads to new kinds of depth. For instance a depth vector can be assigned to each O by integrating along the surface $Q(O,\alpha)$. Vectors of the resulting depth field point to the main mass – not always central or even multi-modal – then rotate and grow longer as α increases. Appropriate limit theorems are derived elsewhere from the forthcoming results.

Informative quantile multivariate data analysis can be performed by moving O and changing the projection rule φ. This new paradigm is rich and can be stated as follows. Facing the fact that $\mathbb{R}^d$ is not naturally ordered, one should simply admit subjectivity and collect viewpoints. The statistical challenge is then to learn about P by comparing the surfaces $Q(O,\alpha)$ while changing (O,α) and φ.

Results don't depend on the observer O only in the orthogonal projection case, which is fully analyzed below. Our limit theorems are uniform in (O,α) and as sharp as for $d=1$, even when P has no density or low dimensional support. Essentially, we jointly control the quantile processes $(\sqrt{n}\,(Y_n(O,u,\alpha) - Y(O,u,\alpha)))$ associated to the projected samples $\langle X_i - O, u\rangle$ in each direction $u \in \mathbb{S}_{d-1}$. The main result is an optimal and surprisingly dimension free Bahadur-Kiefer approximation ([2],[17],[31]). The most useful result is a non asymptotic Brownian approximation.

The closest results we can compare with concern the Tukey contour ([10], [24],[33]). This central region is the intersection of half-spaces having probability α. The main difference is that we study the location of Tukey half-spaces themselves rather than their possibly empty intersection – if $\alpha < d/d+1$, see [10]–, in order to catch all the statistical information. In [26] a central limit theorem is stated for the empirical Tukey contour under strong regularity assumption on P and a mass center. We go further by proving results uniform in α together with rates, approximations and weaker assumptions.

2 From quantiles to quantile surfaces

2.1 Univariate quantiles

It is useful to recall the limiting behavior of the univariate quantile process since our goal is to obtain similar results jointly for a d-dimensional collection of real random samples, each being strongly dependent of the others, namely $Y_n = \langle X_n - O, u \rangle$ where $X_n \in \mathbb{R}^d$, $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$. Consider on $(\Omega, \mathcal{T}, \mathbb{P})$ a sequence $\{Y_n\}$ of independent copies of a real random variable Y. Write, for $y \in \mathbb{R}$ and $\alpha \in (0,1)$, $F_Y(y) = \mathbb{P}(Y \leqslant y)$, $F_Y^{-1}(\alpha) = \inf\{y \in \mathbb{R} : F_Y(y) \geq \alpha\}$ and δ_y the Dirac mass at y. Define the empirical measure $P_n = \sum_{i \leq n} \delta_{Y_i}/n$, the empirical distribution function $F_n = P_n((-\infty, y])$ and the empirical quantile function $F_n^{-1}(\alpha) = \inf\{y \in \mathbb{R} : F_n(y) \geq \alpha\}$, $\alpha \in (0,1)$.

Two problems make the estimation of F_Y^{-1} a not so easy task. First, $F_n^{-1}(\alpha_0)$ is not consistent if F_Y^{-1} is not continuous at α_0 . Second, if S_Y is unbounded then $\sup_{\alpha \in [0,1]} \left| F_n^{-1}(\alpha) - F_Y^{-1}(\alpha) \right| = +\infty$ so that tail quantiles of F cannot be estimated by using extreme values without extra hypotheses and appropriate truncation see [5, 6, 31]. We won't consider this situation here. Let $\Delta = [\alpha^-, \alpha^+]$ where $0 < \alpha^- \leq \alpha^+ < 1$.

Proposition 2.1 (Uniform consistency)**.** *If F_Y is continuous on $F_Y^{-1}(\Delta)$ then*

$$\lim_{n\to\infty} \sup_{\alpha \in \Delta} \left| F_n^{-1}(\alpha) - F_Y^{-1}(\alpha) \right| = 0 \quad a.s. \tag{2.1}$$

if, and only if, F_Y^{-1} is continuous on Δ. This remains true for $\Delta = (0,1)$ if $F_Y^{-1}((0,1))$ is bounded.

Proof. If F_Y^{-1} is continuous on Δ see Section 4.1 where the proof is not classical even for $d = 1$. Conversely, if F_Y^{-1} is not continuous at $\alpha_0 \in (0,1)$ we almost surely have $\limsup_{n\to\infty} \left| F_n^{-1}(\alpha_0) - F_Y^{-1}(\alpha_0) \right| > 0$. To see this, observe that $F_Y^{-1}(\alpha_0) = y_0 < y_1 = \lim_{\alpha \downarrow \alpha_0} F_Y^{-1}(\alpha)$ implies $\mathbb{P}(Y \in (y_0, y_1)) = 0$ thus, with probability one, we have $\inf_n \inf\{Y_i > y_0 : i \leqslant n\} \geqslant y_1$ and also $F_n(y_0) < \alpha_0$ infinitely often, since by the law of the iterated logarithm it holds

$$\liminf_{n\to\infty} \frac{\sqrt{n}(F_n(y_0)-\alpha_0)}{\sqrt{2\alpha_0(1-\alpha_0)\log\log n}} = -1 \quad a.s.$$

therefore $F_n^{-1}(\alpha_0) \geqslant y_1$ happens infinitely often, and the above lim sup is bounded from below by $y_1 - y_0 > 0$. □

In order to establish the weak convergence of quantiles a well behaved density is needed. Assume that Y has density $f_Y > 0$ on $F_Y^{-1}((0,1))$ and define the so-called density quantile function to be

$$h_Y = f_Y \circ F_Y^{-1}. \tag{2.2}$$

Note that h_Y is translation invariant since for all $a, b \in \mathbb{R}_*$ it holds $h_{aY+b} = h_Y / |a|$. Also, $1/h_Y$ is the quantile density function. Few hypotheses on h_Y are required when considering quantiles of order Δ instead of $(0,1)$, thus avoiding controlling tails.

Let $\mathcal{D}(\Delta)$ be the set of left continuous functions on Δ endowed either with the Skorokod topology and Borel sigma field or with the sup-norm topology and the sigma field generated by open balls. A sufficient condition for the Donsker type convergence is the following.

(H) *There exists an open set Δ_0 such that $\Delta \subset \Delta_0$ and f_Y is differentiable on $S_0 = F_Y^{-1}(\Delta_0)$ with* $\inf_{S_0} f_Y > 0$ and $\sup_{S_0} |f'_Y| < \infty$.

Proposition 2.2 (Uniform Central Limit Theorem)**.** *Under **(H)** the sequence of weighted quantile processes $\sqrt{n}\left(F_n^{-1} - F_Y^{-1}\right) h_Y$ indexed by Δ weakly converges on $\mathcal{D}(\Delta)$ to the Brownian Bridge B restricted to Δ.*

Proof. This is Theorem 3.2 when $d = 1$. The differentiability assumption **(H)** corresponds to **(A4)** in Section 3 and is weakened into **(A2)**. □

The convergence of finite dimensional marginals immediately follows, and helps understanding the covariance structure of our multivariate quantiles.

Corollary 2.1. *Fix $0 < \alpha_1 < ... < \alpha_k < 1$. If f_Y is continuous and away from zero on some neighborhood of $\{\alpha_1, ..., \alpha_k\}$ then*

$$\sqrt{n}\begin{pmatrix} F_n^{-1}(\alpha_1) - F_Y^{-1}(\alpha_1) \\ ... \\ F_n^{-1}(\alpha_k) - F_Y^{-1}(\alpha_k) \end{pmatrix} \underset{n\to\infty}{\overset{\mathcal{L}aw}{\longrightarrow}} \mathcal{N}(0_k, \Sigma), \quad \Sigma_{i,j} = \frac{\alpha_i \wedge \alpha_j - \alpha_i\alpha_j}{h_Y(\alpha_i)h_Y(\alpha_j)}. \tag{2.3}$$

Proof. The limiting process B is Gaussian, centered, with covariance function $cov(B(\alpha_1), B(\alpha_2)) = \alpha_1 \wedge \alpha_2 - \alpha_1\alpha_2$, $\alpha_i \in (0,1)$. Thus (2.3) holds under **(H)** with $\alpha^- < \alpha_1 < \alpha_k < \alpha^+$. However the assumption on f'_Y is useless when $\{\alpha_1, ..., \alpha_k\}$ are fixed, it serves in the proof of THEOREM 3.2 for $d = 1$ only to ensure uniform tightness on Δ. Likewize continuity of f_Y is only required locally. $\square$

A way to strengthen and prove PROPOSITION 2.2 is to make use of the Hungarian construction. Starting from [18, 5, 6] this strategy consists in using the quantile transform to control $\sqrt{n}\left(F_n^{-1} - F_Y^{-1}\right) h_Y$ by the easier to handle uniform quantile process uniformly on Δ. Then by KMT ([20]) and the representation of order statistics by partial sums of exponential random variables, the latter can in turn be approximated at rate $(\log n)/\sqrt{n}$ by a sequence of Brownian Bridges built jointly.

Proposition 2.3 (Gaussian Approximation)**.** *Assume that **(H)** holds. Then one can construct on the same probability space $(\Omega, \mathcal{T}, \mathbb{P})$ an i.i.d. sequence Y_n with law F_Y together with a sequence $\{B_n\}$ of standard Brownian Bridges in such a way that*

$$\limsup_{n\to\infty} \frac{\sqrt{n}}{\log n} \sup_{\alpha\in\Delta} \left| \sqrt{n}\left(F_n^{-1}(\alpha) - F_Y^{-1}(\alpha)\right) - \frac{B_n(\alpha)}{h_Y(\alpha)} \right| < \infty \quad a.s.$$

Proof. See [6]. Assumption **(H)** is weakened into **(A3)** at THEOREM 3.3. $\square$

This approach can not be generalized to our quantile surfaces since no quantile transform or partial sum representation hold in $\mathbb{R}^d$. Fortunately, a second strategy works on $\mathbb{R}$. It is based on the Bahadur-Kiefer approximation of the quantile process by the empirical process at rate

$$b_n = n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}.$$

Proposition 2.4 (Bahadur-Kiefer Approximation)**.** *Under **(H)** we have*

$$\limsup_{n\to\infty} \frac{1}{b_n} \sup_{\alpha\in\Delta} \left| \sqrt{n}\left(F_n^{-1}(\alpha) - F_Y^{-1}(\alpha)\right) + \sqrt{n}\left(\frac{F_n(F_Y^{-1}(\alpha)) - \alpha}{h_Y(\alpha)}\right) \right| < \infty \quad a.s.$$

Proof. See [6], [9], [11], [31]. This also follows from THEOREM 3.4 where **(H)** is weakened into **(A3)**. $\square$

This yields an approximation of $\sqrt{n}\left(F_n^{-1} - F_Y^{-1}\right) h_Y$ at this suboptimal order b_n by the KMT Brownian Bridges B'_n built jointly with the empirical process at sup-norm distance $(\log n)/\sqrt{n}$. This further means that the same process B'_n is simultaneously close to the empirical and quantile processes, which could help deriving joint limit laws in statistical applications.

We make use of the second strategy to extend the above results to $\mathbb{R}^d$. Thus the key result is a Bahadur-Kiefer type approximation of the quantile surfaces by the empirical process, and surprisingly b_n turns out to be dimension free. The ensuing Gaussian approximation rates are distribution free, but depends on the dimension through the strong approximation of [4].

2.2 Directional quantiles

In DEFINITION 2.2 below the directional quantile points are built from projections $\{\langle X_n, u\rangle : u \in \mathbb{S}_{d-1}\}$ and are related to each other through a common anchoring point $O \in \mathbb{R}^d$. The resulting quantile points no more depend on O if, and only if, $d = 1$. In this case the left and right directions are associated to the unit vectors $u = -1$ and $u = +1$ and, for $\alpha \in [1/2, 1]$, the left and right directed α-th quantile points are, respectively, $Q(-1, \alpha) = F_{-Y}^{-1}(\alpha)$ and $Q_\alpha(+1) = F_Y^{-1}(\alpha)$. We call $Q_\alpha = \{Q(-1, \alpha), Q(+1, \alpha)\}$ the α-th quantile set.

The usual univariate quantiles use only the right direction $+1$ and $\alpha \in [0, 1]$. They can be deduced from Q_α as follows. Since $Q(-1, \alpha)$ is the right limit of F_Y^{-1} at $1 - \alpha$ it holds $Q(-1, \alpha) \geq F_Y^{-1}(1 - \alpha)$ with equality if and only if F_Y is strictly increasing just after $F_Y^{-1}(1 - \alpha)$. Let $Q^-(-1, 1 - \alpha)$ denote the left continuous version of the increasing function $\alpha \to Q(-1, 1 - \alpha)$ on $[0, 1/2]$. In particular, $Q^-(-1, 1/2) = \inf\{y : F_Y(y) \geq 1/2\}$ and $Q^-(-1, 1) = \inf\{y : F_Y(y) > 0\}$. Also write $Q^+(+1, 1/2) = \sup\{y : F_Y(y) \leq 1/2\}$ the right limit of $Q(+1, \alpha)$ at $\alpha = 1/2$. Then we have

$$F_Y^{-1}(\alpha) = \mathbb{1}_{\alpha<1/2} Q^-(-1, 1 - \alpha) + \mathbb{1}_{\alpha>1/2} Q(+1, \alpha), \quad \alpha \in (0, 1) \setminus \{1/2\}$$

and $Q_{1/2} = [Q^-(-1, 1/2), Q^+(+1, 1/2)]$ is the median interval of Y. Let $Q_n(-1, \alpha)$ be the empirical α-th quantile of $-Y_1, ..., -Y_n$ and $Q_n(+1, \alpha) = F_n^{-1}(\alpha)$. Write $h(-1, \alpha) = h_{-Y}(\alpha)$ and $h(+1, \alpha) = h_Y(\alpha)$.

In the univariate case all subjective viewpoints are the same since O plays no role and THEOREM 3.2 reduces exactly to the following.

Corollary 2.2. *Assume that **(H)** holds. The sequence of real random processes* $\sqrt{n}\,(Q_n(u, \alpha) - Q(u, \alpha))$ *indexed by* $(u, \alpha) \in \{-1, 1\} \times \Delta$ *weakly converges to a centered Gaussian process* G_P *indexed by* $(u, \alpha) \in \{-1, 1\} \times \Delta$ *having covariance given by*

$$cov(G_P(u_1, \alpha_1), G_P(u_2, \alpha_2)) = \frac{\alpha_1 \wedge \alpha_2 - \alpha_1\alpha_2}{h(u_1, \alpha_1)h(u_2, \alpha_2)}. \tag{2.4}$$

Proof. Take $d = 1$ in THEOREM 3.2. This is also a simple consequence of PROPOSITION 2 since by hypothesis F_Y^{-1} is strictly increasing on Δ and thus $Q(-1, \alpha) = F_Y^{-1}(1 - \alpha)$. The limiting process is then defined by $G_P(+1, \alpha) = B(\alpha)$ and $G_P(-1, \alpha) = B(1 - \alpha)$ so that (2.3) yields (2.4). □

Here is our flexible general definition of multivariate quantile surfaces.

Definition 2.1. *(Generalized quantile sets). Let $O \in \mathbb{R}^d$, $u_0 \in \mathbb{S}_{d-1}$, 0_d be the origin and φ be a u_0-symmetric continuous function from $\mathbb{R}^d$ to $\mathbb{R}$ satisfying*

$$\varphi^{-1}((-\infty, y_1]) = A_{y_1} \subset A_{y_2}, \quad y_1 \le y_2,$$
$$\lambda_d(\varphi^{-1}(\{y\})) = 0, \quad y \in \mathbb{R}.$$

For any $u \in S_{d-1}$ write r_u any rotation of $\mathbb{R}^d$ having center 0_d and angle $u_0 \hookrightarrow u$ and t_O the translation directed by O. For $\alpha \in [1/2, 1)$ define

$$\begin{aligned} Y(O,u,\alpha) &= \inf\{y : \mathbb{P}(t_O \circ r_u(A_y)) \ge \alpha\} \\ Q(O,\alpha) &= \{O + Y_\alpha(O,u)u : u \in \mathbb{S}_{d-1}\} \end{aligned}$$

to be the u-directional (φ, u_0)-shaped α-th quantile range and set seen from O.

Hence each α-th quantile point $O + Y_\alpha(O,u)u$ corresponds to a set having probability α, symmetric with respect to the line (O,u). Put together this points form a surface $Q(O,\alpha)$ under appropriate conditions. It is easily seen that DEFINITION 2.1 reduces to DEFINITION 2.2 in the special case $\varphi(x) = \langle x, u_0\rangle$, $A_y = \varphi^{-1}((-\infty, y]) = H(0_d, u_0, y)$. This orthogonal projection case is our main focus.

2.3 Multivariate quantile surfaces

Let $\mathcal{H}$ denote the family of all half-spaces and $\mathcal{H}_\alpha$ the sub-family of half-spaces H having probability $P(H) = \alpha > 0$. Let

$$H(O,u,y) = \left\{x \in \mathbb{R}^d : \langle x - O, u\rangle \le y\right\} \in \mathcal{H} \tag{2.5}$$

be the half-space standing at distance $y \in \mathbb{R}$ from O in direction $u \in \mathbb{S}_{d-1}$. Given $\alpha \in [1/2, 1)$ and $u \in \mathbb{S}_{d-1}$ let

$$Y(O,u,\alpha) = \inf\{y : P(H(O,u,y)) \ge \alpha\}$$

be the u-directional α-th quantile range from O and

$$H(u,\alpha) = H(O,u,Y(O,u,\alpha))$$

be the u-directional α-th quantile half-space, that does not depend on O. Conversely, for $y \in \mathbb{R}$, $P(H(O,u,y))$ is the u-directional p-value at y. It is noteworthy that $P(H(O,u,y)) = F_{\langle X-O,u\rangle}(y)$ and thus

$$Y(O,u,\alpha) = F^{-1}_{\langle X-O,u\rangle}(\alpha) = F^{-1}_{\langle X,u\rangle}(\alpha) - \langle O,u\rangle \tag{2.6}$$

is the α-th quantile of the real random variable $\langle X - O, u\rangle$.

Definition 2.2 (Multivariate quantile set)**.** *For $\alpha \in [1/2, 1)$, $O \in \mathbb{R}^d$ and $u \in \mathbb{S}_{d-1}$ define the u-directional α-th quantile point seen from O to be*

$$Q(O, u, \alpha) = O + Y(O, u, \alpha)u \tag{2.7}$$

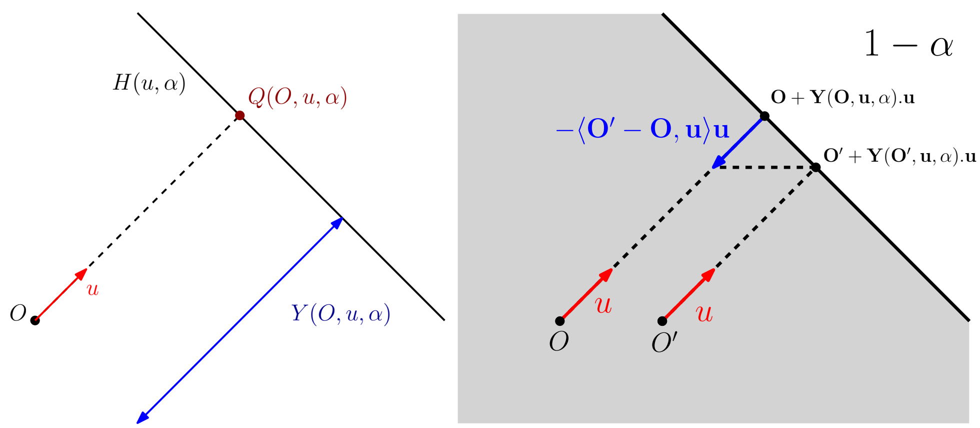
and the α-th quantile set seen from O to be the star-shaped collection of points

$$Q(O, \alpha) = \{Q(O, u, \alpha) : u \in \mathbb{S}_{d-1}\}. \tag{2.8}$$

Since

$$Q(O', u, \alpha) = Q(O, u, \alpha) + O' - O - \langle O' - O, u \rangle u \tag{2.9}$$

it is easy to get all quantile sets $Q(O, \alpha)$ from any of them. However, from a statistical point of view, looking at several $Q(O, \alpha)$ simultaneously by moving O, is a good way to learn about P.



We restrict ourselves to laws P for which the α-th quantile sets $Q(O, \alpha)$ are surfaces, but we do not require that P is absolutely continuous.

Remark 2.1. *The boundary of the intersection $\mathcal{T}_\alpha$ of all $H(u, \alpha)$ is the so-called "Tukey contour". If $\mathcal{T}_\alpha$ is not empty then it is a compact convex set and $u \to Y(O, u, \alpha)$ is its support function. Hence it is continuous, subadditive and, in general, not differentiable. However, $\mathcal{T}_\alpha$ is likely to be empty if P is multimodal and α is small enough.*

A median surface simply corresponds to $\alpha = 1/2$ and has no special feature except maybe at central points where it is more self intersecting than for $\alpha > 1/2$ or at outliers.

Basic assumptions. Let 0_d be the origin of $\mathbb{R}^d$ and $\Delta = [\alpha^-, \alpha^+] \subset [1/2, 1)$. Assume that the hyperplanes

$$\partial H(u, \alpha) = \left\{ x \in \mathbb{R}^d : \langle x, u \rangle = Y(0_d, u, \alpha) \right\}$$

satisfy

$$(\mathbf{A_0^-}) \quad P(\partial H(u, \alpha)) = 0, \quad u \in \mathbb{S}_{d-1}, \quad \alpha \in \Delta.$$

Under $(\mathbf{A_0^-})$, for $\alpha \in \Delta$ we have $P(H(u, \alpha)) = \alpha$ and $\mathcal{H}_\alpha = \{H(u, \alpha) : u \in \mathbb{S}_{d-1}\}$. This excludes laws P partly supported by one or more hyperplanes, for instance laws P with discrete component. Assume moreover that hyperbands

$$H(O, u, y, z) = H(O, u, z) \setminus H(O, u, y) \quad y < z \tag{2.10}$$

satisfy for all $u \in \mathbb{S}_{d-1}$

$$(\mathbf{A_0^+}) \quad P(H(O, u, y, z)) > 0, \quad Y(O, u, \alpha^-) \leq y < z \leq Y(O, u, \alpha^+).$$

Remark 2.2. *Let $A \triangle B = (A \backslash B) \cup (B \backslash A)$. Under $(\boldsymbol{A_0^-})$ and $(\boldsymbol{A_0^+})$ it holds*

$$\lim_{\beta \to \alpha} P(H(u, \alpha) \triangle H(u, \beta)) = 0, \quad u \in \mathbb{S}_{d-1}, \quad \alpha \in [1/2, 1)\,.$$

These two assumptions are sufficient to make the natural non-parametric estimator of $Q(O, u, \alpha)$ consistent uniformly in (O, u, α). Define the set of admissible distances by

$$\mathcal{Y}_\Delta(O, u) = \{y : P(H(O, u, y)) \in \Delta\} = F^{-1}_{\langle X - O, u \rangle}(\Delta). \tag{2.11}$$

Since the probability measure P is tight, there exists $r^+ > 0$ such that $P(B(O, r^+)) > \alpha^+$ and thus $H(u, \alpha) \cap B(O, r^+) \neq \emptyset$ for $\alpha \in \Delta$, hence

$$\sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |Y(O, u, \alpha)| = \sup_{u \in \mathbb{S}_{d-1}} \sup_{y \in \mathcal{Y}_\Delta(O, u)} |y| < +\infty. \tag{2.12}$$

Theorem 2.1. *Under $(\boldsymbol{A_0^-})$, assumption $(\boldsymbol{A_0^+})$ is equivalent to the fact that $(u, \alpha) \mapsto Q(O, u, \alpha)$ is continuous on $\mathbb{S}_{d-1} \times \Delta$ for any $O \in \mathbb{R}^d$.*

By Theorem 2.1, the set $Q(O, \alpha)$ from (2.8) is the image of the compact set $\mathbb{S}_{d-1}$ through a continuous application, it is a surface we call the α-th quantile surface seen from O.

Corollary 2.3. *Under $(\boldsymbol{A_0^-})$ and $(\boldsymbol{A_0^+})$, the set $Q(O, \alpha)$ is a closed surface, for all $O \in \mathbb{R}^d$ and $\alpha \in \Delta$.*

Let define the set of admissible bands of width $\varepsilon > 0$ allowed by Δ to be

(2.13)
$$\mathcal{B}_\varepsilon = \left\{ H(O,u,y,y+\varepsilon) : \quad O \in \mathbb{R}^d, u \in \mathbb{S}_{d-1}, y, y+\varepsilon \in \mathcal{Y}_\Delta(O,u) \right\}.$$

Note that $\mathcal{B}_\varepsilon$ depends on Δ through $\mathcal{Y}_\Delta(O,u)$. It is useful to rewrite $(\mathbf{A_0^-})$ and $(\mathbf{A_0^+})$, in terms of the function

$$\Psi(\varepsilon) = \inf_{B \in \mathcal{B}_\varepsilon} P(B), \quad \varepsilon > 0 \tag{2.14}$$

Proposition 2.5. *Under* $(\boldsymbol{A_0^-})$ *and* $(\boldsymbol{A_0^+})$ *the following two conditions hold true*

$(\boldsymbol{A^-_{0,\Psi}})$ $\lim_{\varepsilon \to 0} \Psi(\varepsilon) = 0$

$(\boldsymbol{A^+_{0,\Psi}})$ $\Psi(\varepsilon) > 0,\ 0 < \varepsilon < \varepsilon^+ = \sup\{\varepsilon > 0, \mathcal{B}_\varepsilon \neq \emptyset\}$.

Proposition 2.6. *The function* Ψ *is right-continuous on* $(0, \varepsilon^+)$. *Moreover, under* $(\boldsymbol{A_0^-})$ *and* $(\boldsymbol{A_0^+})$ *the function* Ψ *is continuous on* $[0, \varepsilon^+)$.

By Proposition 2.6 under $(\mathbf{A_0^-})$ and $(\mathbf{A_0^+})$ Ψ is *càdlàg* and strictly increasing with

$$\Psi \circ \Psi^{-1}(\alpha) = \alpha \quad \text{and} \quad \Psi^{-1} \circ \Psi(\alpha) \geq \alpha, \quad \alpha \in \Delta. \tag{2.15}$$

2.4 Empirical quantile surfaces

We intend to estimate $Q(O,\alpha)$ jointly in $\alpha \in \Delta \subset [1/2, 1)$ and $O \in \mathbb{R}^d$ by applying the definition of quantile surfaces from section 2.3 to the empirical measure $P_n = \frac{1}{n} \sum_{i \leq n} \delta_{X_i}$, where δ_x is the Dirac mass at $x \in \mathbb{R}^d$. For $u \in \mathbb{S}_{d-1}$ let

$$Y_n(O,u,\alpha) = \inf \left\{ y : P_n(H(O,u,y)) \geq \alpha \right\}.$$

Define the u-directional α-th empirical quantile point seen from O to be

$$Q_n(O,u,\alpha) = O + Y_n(O,u,\alpha)u$$

and associate to this point the α-th empirical quantile half-space

$$H_n(u,\alpha) = H(O,u,Y_n(O,u,\alpha)).$$

Let the α-th empirical quantile set seen from O be

$$Q_n(O,\alpha) = \left\{ Q_n(O,u,\alpha) : u \in \mathbb{S}_{d-1} \right\}.$$

The quantile half-spaces indexed by points $Q_n(O,u,\alpha)$ are collected into

$$\mathcal{H}_{n,\alpha} = \left\{ H_n(u,\alpha) : u \in \mathbb{S}_{d-1} \right\}$$

For O, O' in $\mathbb{R}^d$ we have $Y_n(O', u, \alpha) = Y_n(O, u, \alpha) - \langle O' - O, u \rangle$ and combining this with (2.9) we can highlight the following important property of the directional quantiles process

$$Y_n(O', u, \alpha) - Y(O', u, \alpha) = Y_n(O, u, \alpha) - Y(O, u, \alpha) \tag{2.16}$$

which means that $Y_n - Y$ is independent of O.

Proposition 2.7. *Under* $(\mathbf{A_0^-})$*, for all* $n > d$*,*

$$\mathcal{H}_{n,\alpha} \subset \left\{ H : H \textit{ half-space, } P_n(H) \in \left[\alpha, \alpha + \frac{d}{n} \right] \right\}.$$

2.5 Illustrations and comments

We picture out several examples in dimension 2. On Fig 1 we show shapes of quantile surfaces obtained for symmetric laws, here the symmetry point is $O = (0, 0)$ and thus $Q(O, \alpha)$ is a circle. The function $\alpha \to Y(O, (1, 0), \alpha)$ corresponds to the univariate quantile function of the radial law. Moving O at $O_2 = (-3, 0)$, $O_3 = (-5, 0)$, $O_4 = (-7, 0)$ gives examples of typical shapes when the observer is away from the central point. This typical shape has one inner and one outer loops intersecting at O, each corresponding to connex subsets of directions in $\mathbb{S}_{d-1}$.

Fig 2 shows that the previous typical shape is preserved even if P has no density but obeys $(\mathbf{A_0^-})$ and $(\mathbf{A_0^+})$, here a spiral support with uniform law. The empirical surface for $\alpha = 0.7$ is shown to be less smooth with $n = 1000$ points than the almost true one with $n = 10000$ points.

Next we consider a mixture of two Gaussian distributions $\mathcal{N}((-2, 0), I)$ and $\mathcal{N}((2, 0), 3I)$ with weights $1/4$ and $3/4$ respectively, where I is the identity matrix. In Fig 3 the surfaces are contours since the observer is inside the central area, here we take $\alpha = 0.6$, 0.7, 0.8 and 0.9. In Fig 4 α is fixed and O is moving outside the data. Note that any of the surfaces can be deduced from the other by (2.9) so drawing several O is very fast and facilitates a visual human interpretation.

In Fig 5 and 6, P is a similar gaussian mixture but the two modes are more separated compared to the standard deviation. The Tukey contours are sometimes empty, however the quantile surfaces always exist and are shown from an observer standing between the two modes. In Fig 5 increasing α results in resorbing the left part of the initial contours to create an inside loop at the right hand side, associated to the left oriented directions for which the mass has to be catched behind the observer – here $\alpha = 0.6$, 0.7, 0.8 and 0.9. In Fig 6, moving O for a fixed $\alpha = 0.7$ is a simple computation and drawing all surfaces helps understanding where the modal areas are located – for alpha large enough the main modal area is easily revealed in between the surfaces. In cases where the data cloud is so big that no study can be performed visually such a data summary can be useful.

On FIG 8 we show in red color, the median surface seen from $O = (0,0)$ which is almost a point since the spiral uniform law is "almost" symmetric. By zooming toward the median surface we can see on FIG 9 that it is indeed a very oscillating surface around O with a very small volume. Obviously if P is symmetric about M then the median surface seen from $O = M$ is reduced to the point O itself and the median surface seen from another point is a sphere (a circle here) passing through the symmetry point M. Such a central median point can then be localized by intersecting median surfaces. If P is not symmetric the median surface has not necessarily a small volume somewhere. For instance at FIG 11 the point where the median surface is almost of minimum volume for the second gaussian mixture is at $O = (4.1, 0)$. The associated median surface shows three loops – one cutting mass from the right and two from the upper left or lower left respectively. Moreover the median surfaces seen from $O_1 = (-15, 6)$, $O_2 = (-5, 6)$, $O_3 = (-5, -6)$ and $O_4 = (15, 6)$, intersects around the median surface of FIG 10 but they are not circles.

It is noteworthy that under $(\mathbf{A_0^-})$ and $(\mathbf{A_0^+})$, every median surface is a "double" surface, in the sense that every point of $Q(O, 1/2)$ corresponds at the same time to $Q(O, u, 1/2)$ and the point $Q(O, -u, 1/2)$.

In FIG 13 we show a case where at the special point $O = (0,0)$ even for α large more than two loops appear inside the quantile surface. Here P is a mixture of several laws having disjoint supports separated by lines containing O. Moving slightly O at FIG 14 provides again the typical shapes and the transition merging the two inner loops into one is smooth as O moves. Then sending O far away confirms the usual shape seen from outer points, see FIG 15.

As a conclusion we promote the technique of moving alpha and O to analyze data from the mass localization viewpoint. Since all is under the control of sharp limit theorems we can also think about using deterministic and random projections on low dimensional spaces minimizing quantile surfaces, as for linear data analysis. It is possible to build many kind of tests based on quantile surfaces, and also depth vector fields summarizing for each O the distance and average direction to move in order to recover α mass.

3 Results

3.1 Uniform Strong Consistency

The following result reduces exactly to PROPOSITION 2.1, when $d = 1$.

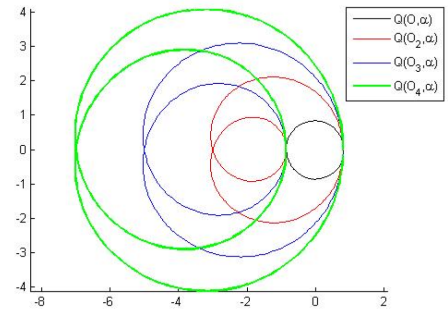


Figure 1: Central symmetric law, $\alpha = 0.8$, moving O.

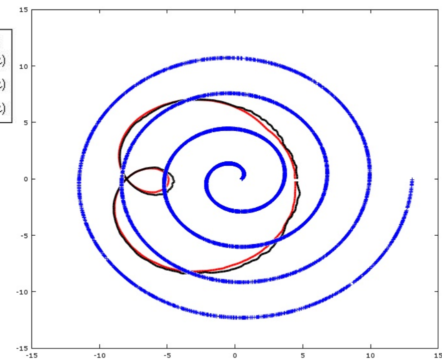


Figure 2: Law with Lebesgue zero support.

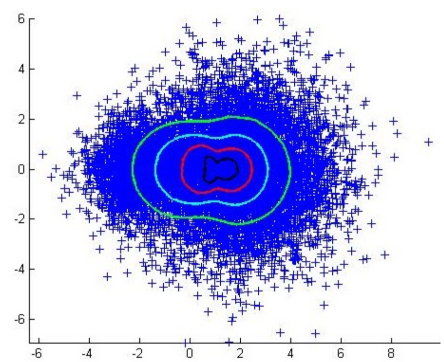


Figure 3: O inward, moving α.

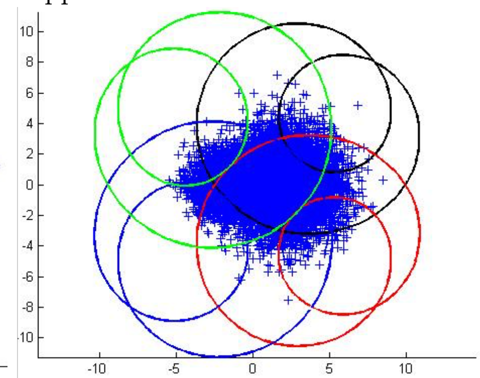


Figure 4: α fixed, $\alpha = 0.7$, moving O around.

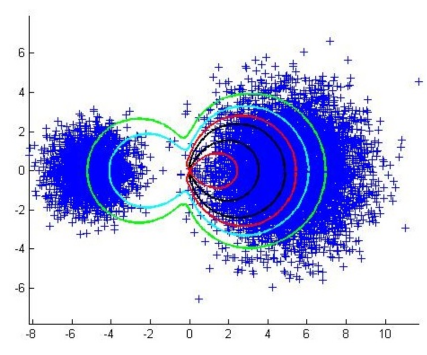


Figure 5: O fixed between two modes, moving α.

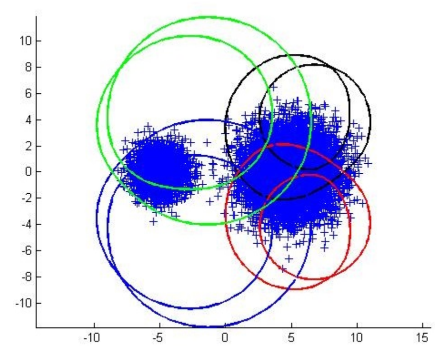


Figure 6: α fixed, $\alpha = 0.7$, moving O around.

Figure 7: Examples of quantiles surfaces in dimension 2

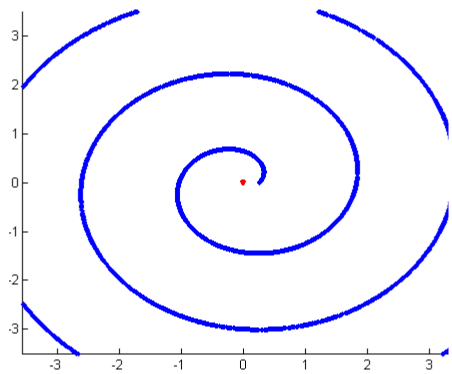


Figure 8: Median surface for a Lebsgue zero supported measure

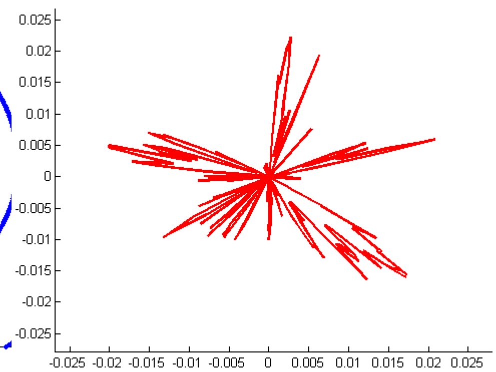


Figure 9: Zoom on the median of Fig 8

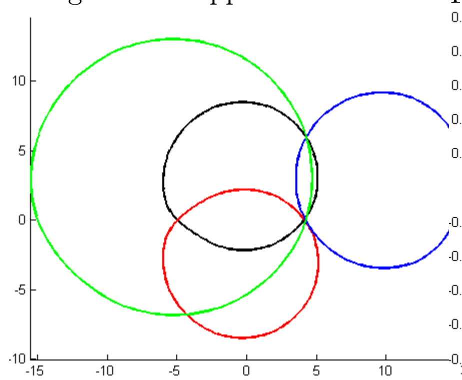


Figure 10: Median surfaces for an asymmetric law while moving O

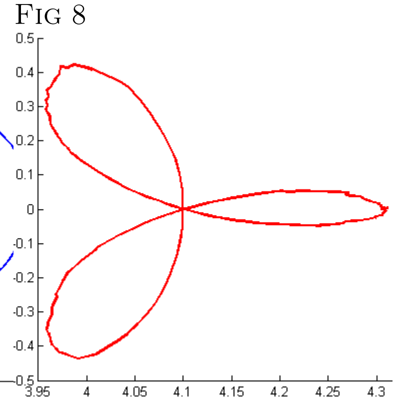


Figure 11: Median surface for an asymmetric law for $O = (4.1, 0)$

Figure 12: Examples of median surfaces.

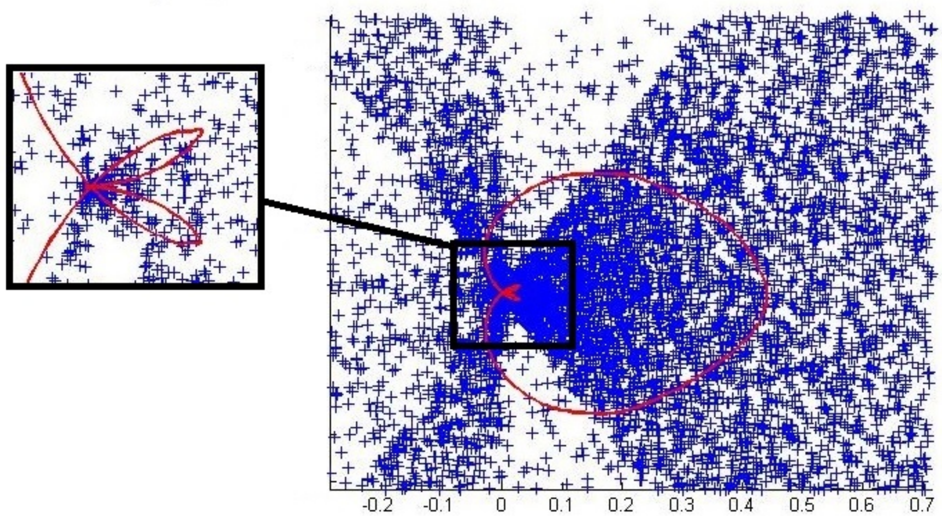


Figure 13: Two inner loops at $O = (0, 0)$, $\alpha = 0, 7$.

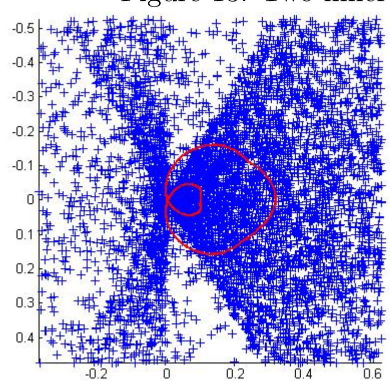


Figure 14: O moved outward, $\alpha = 0, 7$.

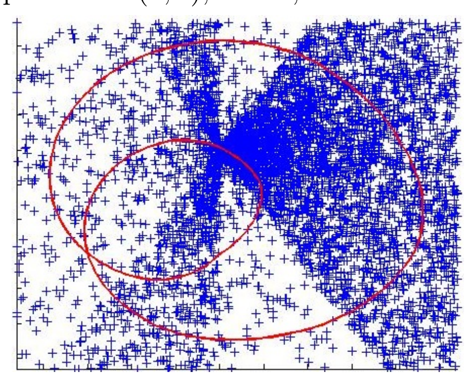


Figure 15: One inner loop at $O = (0, 0)$, $\alpha = 0, 8$.

Figure 16: An example of law with multi loops for some O.

Theorem 3.1 (Uniform Strong Consistency)**.** *Under the assumption* $(\boldsymbol{A_0^-})$*,* $(\boldsymbol{A_0^+})$ *is equivalent to*

$$\lim_{n\to\infty} \|Y_n(O,u,\alpha) - Y(O,u,\alpha)\|_{\mathbb{R}^d\times S_{d-1}\times\Delta} = 0 \quad a.s.$$

Hence, we have

$$\lim_{n\to\infty} \sup_{O\in\mathbb{R}^d} \sup_{\alpha\in\Delta} d_H(Q_{n,\alpha}(O), Q_\alpha(O)) = 0 \quad a.s.$$

where d_H *denotes the Hausdorff distance.*

To go beyond this consistency result, we require the existence of a directional density quantile as in (2.2). For $(u,\alpha) \in \mathbb{S}_{d-1}\times\Delta$ define

$$h(u,\alpha) = h_u(\alpha) = f_{\langle X,u\rangle} \circ F^{-1}_{\langle X,u\rangle}(\alpha).$$

$(\mathbf{A_1})$ For all $u \in \mathbb{S}_{d-1}$, the random variable $\langle X,u\rangle$ has a continuous density $f_{\langle X,u\rangle} > 0$ on $F^{-1}_{\langle X,u\rangle}(\Delta)$, moreover, for some m and M

$$0 < m \le \inf_{\alpha\in\Delta} \inf_{u\in\mathbb{S}_{d-1}} h_u(\alpha) \le \sup_{\alpha\in\Delta} \sup_{u\in\mathbb{S}_{d-1}} h_u(\alpha) \le M < +\infty$$

Remark that $(\mathbf{A_1})$ does not imply that P has a density on $\mathbb{R}^d$. However $(\mathbf{A_1})$ implies $(\mathbf{A_0^-})$ and $(\mathbf{A_0^+})$ with

$$m(z-y) \le P(H(O,u,y,z)) \le M(z-y), \quad y<z, \;\; u\in\mathbb{S}_{d-1}, \;\; y,z \in \mathcal{Y}_\Delta(O,u)$$

In particular, under $(\mathbf{A_1})$ P has no discrete component in its Lebesgue–Nikodym decomposition, and likewise none of the marginal laws of P, has a discrete component since none of their linear combinations has.

3.2 Uniform Weak Convergence

In order to state the central limit theorem, we first define the limiting Gaussian process $\mathbb{G}_P$. Let $\mathbb{B}_P$ be the P-Brownian bridge indexed by half-spaces, that is the zero mean Gaussian process on $\mathcal{H}$ having covariance $cov(\mathbb{B}_P(H), \mathbb{B}_P(H')) = P(H\cap H') - P(H)P(H')$, for $(H,H') \in \mathcal{H}\times\mathcal{H}$. Under $(\mathbf{A_1})$ the random function

$$\mathbb{G}_P(u,\alpha) := \frac{\mathbb{B}_P(H(u,\alpha))}{h(u,\alpha)}, \text{ for } (u,\alpha)\in\mathbb{S}_{d-1}\times\Delta \tag{3.1}$$

is a bounded centered Gaussian process indexed by the compact parameter set $\mathbb{S}_{d-1}\times\Delta$ with covariance function given by

$$cov(\mathbb{G}_P(u_1,\alpha_1), \mathbb{G}_P(u_2,\alpha_2)) = \frac{P(H(u_1,\alpha_1)\cap H(u_2,\alpha_2)) - \alpha_1\alpha_2}{h(u_1,\alpha_1)h(u_2,\alpha_2)}. \tag{3.2}$$

We also set $\overrightarrow{\mathbb{G}}_P(u,\alpha) := \mathbb{G}_P(u,\alpha)\cdot u,$ for $(u,\alpha)\in\mathbb{S}_{d-1}\times\Delta$. To state the regularity condition $\mathbf{(A_2)}$ ensuring the weak convergence, we need to introduce for all $0<\gamma<\gamma_0$, the quantity

(3.3)
$$\rho(\gamma) = \sup_{|\varepsilon'|<\gamma}\sup_{u\in\mathbb{S}_{d-1}}\sup_{\alpha\in\Delta}\left|F_{\langle X,u\rangle}(Y(O,u,\alpha)+\varepsilon')-\alpha-h(u,\alpha)\varepsilon'\right|$$

that controls the expansion of $F_{\langle X,u\rangle}$ in the γ-neighborhood of $Y(O,u,\alpha)$.

$\mathbf{(A_2)}$ We assume that

$$\lim_{\gamma\to 0}\frac{\sqrt{\log\log(1/\gamma)}}{\gamma}\rho(\gamma)=0.$$

Theorem 3.2 (Uniform Central Limit Theorem)**.** *Under* $\boldsymbol{(A_1)}$ *and* $\boldsymbol{(A_2)}$ *the process* $\sqrt{n}(Y_n-Y)$ *weakly converges to* $\mathbb{G}_P$ *on the set of bounded real functions on* $\mathbb{S}_{d-1}\times\Delta$*, endowed with the supremum norm. Likewise, the process* $\sqrt{n}(Q_n-Q)$ *indexed by* $\mathbb{S}_{d-1}\times\Delta$ *weakly converges to* $\overrightarrow{\mathbb{G}}_P$ *on the set of bounded* $\mathbb{R}^d$ *valued functions on* $\mathbb{S}_{d-1}\times\Delta$*, endowed with the supremum norm.*

Theorem 3.2 is a weak convergence statement involving jointly all quantile surfaces for $\alpha\in\Delta$. In particular, we have the following CLT for finite set of points on any of these surfaces.

Corollary 3.1. *Let* $(O_1,u_1,\alpha_1),...,(O_k,u_k,\alpha_k)$ *dans* $\mathbb{R}^d\times\mathbb{S}_{d-1}\times\Delta$*. Under* $\boldsymbol{(A_1)}$ *and* $\boldsymbol{(A_2)}$*, we have*

$$\sqrt{n}\begin{pmatrix} Y_n(O_1,u_1,\alpha_1)-Y(O_1,u_1,\alpha_1)\\ \dots\\ Y_n(O_k,u_k,\alpha_k)-Y(O_k,u_k,\alpha_k)\end{pmatrix}\underset{n\to\infty}{\overset{\mathcal{L}aw}{\longrightarrow}}\mathcal{N}(0_k,\Sigma)$$

with Σ *the covariance matrix defined by*

$$\begin{aligned}\Sigma_{i,j} &= \frac{P(H(O_i,u_i,Y(O_i,u_i,\alpha_i))\cap H(O_j,u_j,Y(O_j,u_j,\alpha_j)))-\alpha_i\alpha_j}{f_{\langle X-O_i,u_i\rangle}\circ F^{-1}_{\langle X-O_i,u_i\rangle}(\alpha_i).f_{\langle X-O_j,u_j\rangle}\circ F^{-1}_{\langle X-O_j,u_j\rangle}(\alpha_j)}\\ &= \frac{P(H(u_i,\alpha_i)\cap H(u_j,\alpha_j))-\alpha_i\alpha_j}{h(u_i,\alpha_i)h(u_j,\alpha_j)}.\end{aligned}$$

Note that Theorem 3.2 and Corollary 3.1 are exact generalizations of Proposition 2.2 and Corollary 2.1, respectively.

3.3 The main result

To ensure the Bahadur-Kiefer type representation, we need the following stronger condition.

$\mathbf{(A_3)}$ We suppose that

$$\lim_{\gamma\to 0}\frac{\rho(\gamma)\sqrt{\log\log(1/\gamma)}}{\gamma^{3/2}\sqrt{\log(1/\gamma)}}=0.$$

This condition can be replaced by one of the following conditions, which are more restrictive but easier to check.

$\mathbf{(A_3')}$ There exists $r > 1/2$ and $C^* > 0$ such that for all $0 < \gamma < \gamma_0$

$$\rho(\gamma) \leq C^* \gamma^{1+r}.$$

$\mathbf{(A_4)}$ The function h is differentiable on $\mathbb{S}_{d-1} \times \Delta$ with uniformly bounded derivatives.

Under $\mathbf{(A_4)}$, the assumption $\mathbf{(A_3')}$ holds true with $r = 1$. Moreover we have $\mathbf{(A_4)} \Rightarrow \mathbf{(A_3')} \Rightarrow \mathbf{(A_3)} \Rightarrow \mathbf{(A_2)}$. Let $\Lambda_n = \sqrt{n}(P_n - P)$ be the empirical process indexed by $\mathcal{H}$ and define

$$\mathbb{E}_n(u,\alpha) = \Lambda_n(H(u,\alpha)) = \sqrt{n}(P_n(H(u,\alpha)) - \alpha)$$

Theorem 3.3 (Bahadur-Kiefer type representation). *Under $\boldsymbol{(A_1)}$ and $\boldsymbol{(A_2)}$ we have*

$$\lim_{n\to\infty} \left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{S_{d-1}\times\Delta} = 0 \quad a.s. \tag{3.4}$$

and for any $\theta > 0$ there exists $c_\theta(m,M,d) > 0$ and $n_\theta(m,M,d) > 0$ such that we have, for all $n > n_\theta$,

$$\mathbb{P}\left(\left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{R}^d\times S_{d-1}\times\Delta} \geq c_\theta a_n \right) \leq \frac{1}{n^\theta}, \tag{3.5}$$

where

$$a_n = \sqrt{n}\rho\left(\sqrt{\log\log n/n}\right) \vee \frac{(\log n)^{1/2}(\log\log n)^{1/4}}{n^{1/4}}.$$

If moreover $\boldsymbol{(A_3)}$ holds then

$$\left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{S_{d-1}\times\Delta} = O_{a.s.}\left(\frac{(\log n)^{1/2}(\log\log n)^{1/4}}{n^{1/4}} \right). \tag{3.6}$$

Note that THEOREM 3.3 contains PROPOSITION 2.4 for $d = 1$. It is a good surprise that the order of the rate of convergence in (3.6) is dimension free. Note that c_θ can be computed explicitly and depends on the dimension d and P. By THEOREM 3.3 the multivariate empirical quantile surfaces inherit the properties of the empirical process.

3.4 Non asymptotic strong approximation

The following Gaussian approximation is useful to construct explicit confident bands around empirical quantile surfaces by using Monte-Carlo methods. As a matter of fact, using (3.7) and (3.8) it remains to plug-in any estimator of h in the covariance (3.2) in order to compute joint confident intervals along a very large set of points from $Q_n(O, \alpha)$. Even for fixed n the probability of such confident band has an explicit upper bound.

Theorem 3.4 (Uniform Strong Approximation with rate). *Under* ***(A₁)*** *and* ***(A₂)*** *one can construct on the same probability space* $(\Omega, \mathcal{T}, \mathbb{P})$ *an i.i.d. sequence* X_n *with distribution* P *and a sequence* $\mathbb{G}_n$ *of versions of* $\mathbb{G}_P$ *in such a way that for* $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$, $\alpha \in \Delta$

$$Y_n(O, u, \alpha) = Y(O, u, \alpha) + \frac{\mathbb{G}_n(u, \alpha)}{\sqrt{n}} + \frac{\mathbb{Z}_n(u, \alpha)}{\sqrt{n}} \tag{3.7}$$

where $\mathbb{Z}_n = \sqrt{n}(Y_n - Y) - \mathbb{G}_n$ *is such that*

$$\lim_{n \to \infty} \|\mathbb{Z}_n\|_{S_{d-1} \times \Delta} = 0 \quad a.s. \tag{3.8}$$

If P *moreover satisfies* ***(A₃)*** *then* $\mathbb{G}_n$ *can be constructed such that for* $v_d = 1/(2 + 10d)$ *and* $w_d = (4 + 10d)/(4 + 20d)$, *there exists* $n_\theta(m, M, d) > 0$ *such that we have, for all* $n > n_\theta$,

$$\mathbb{P}\left(\|\mathbb{Z}_n\|_{S_{d-1} \times \Delta} \geq c_\theta \frac{(\log n)^{w_d}}{n^{v_d}} \right) \leq \frac{1}{n^\theta}. \tag{3.9}$$

3.5 Law of the iterated logarithm

Recall that Ψ is given in (2.14).

Theorem 3.5 (Law of the Iterated Logarithm). *Under* $(\boldsymbol{A_0^-})$ *and* $(\boldsymbol{A_0^+})$

$$\limsup_{n \to \infty} \frac{\|Y_n - Y\|_{\mathbb{S}_{d-1} \times \Delta}}{\Psi^{-1}\left(\sqrt{(\log \log n)/n}\right)} < \infty \quad a.s.$$

Remark 3.1. *If instead of* $(\boldsymbol{A_0^-})$ *and* $(\boldsymbol{A_0^+})$ *we assume the stronger* ***(A₁)***, *then the law of iterated logarithm can be rewritten in the following more classical form*

$$\limsup_{n \to \infty} \frac{\|Y_n - Y\|_{\mathbb{S}_{d-1} \times \Delta}}{\sqrt{(\log \log n)/n}} < \infty \quad a.s.$$

In the particular case of a central symmetric distribution, we obtain exactly same result as for the quantile process on $\mathbb{R}$.

4 Proofs

4.1 Proof of Theorem 2.1

The proof of Theorem 2.1 relies on the technical Lemma 5.1. Its proof is postponed to the appendix.
Necessary Condition. First, under $(\mathbf{A_0^-})$ and $(\mathbf{A_0^+})$ $Q(O,\alpha)$ is a bounded set since (2.12) holds, then for all $O \in \mathbb{R}^d$ there exists $r > 0$ such that for all $u \in \mathbb{S}^d$, $\alpha \in \Delta$ we have $O+Y(O,u,\alpha)u = Q(O,u,\alpha) \in B(O,r)$. Now, we show that $(u,\alpha) \mapsto Q(O,u,\alpha)$ is continuous. If $(u,\alpha) \mapsto Q(O,u,\alpha)$ is not continuous, then there exists a sequence $(u_n)_{n\geq 1}$ in $\mathbb{S}_{d-1}$ and $(\alpha_n)_{n\geq 1}$ in Δ with $u_n \to u$ and $\alpha_n \to \alpha$ such that

$$\lim_{n\to\infty} Q(O,u_n,\alpha_n) \neq Q(O,u,\alpha).$$

Since $Q(O,u_n,\alpha_n)$ is bounded there exists a subsequence $(u_{n_j})_{j\geq 1}$ such that $u_{n_j} \to u$ and $(\alpha_{n_j})_{j\geq 1}$ such that $\alpha_{n_j} \to \alpha$ with moreover

$$\lim_{j\to\infty} Q(O,u_{n_j},\alpha_{n_j}) = Q_\infty = O + y_\infty u \in \mathbb{R}^d$$

where $y_\infty = \lim_{j\to\infty} Y(O,u_{n_j},\alpha_{n_j}) < +\infty$ and

$$y_\infty \neq y = Y(O,u,\alpha)$$

so that $Q_\infty \neq Q(O,u,\alpha)$. Suppose that $y < y_\infty$ and choose y' such that $y < y' < y_\infty$. By Lemma 5.1, there exists an increasing subsequence $(n_{j(k)})_{k\geq 1}$ with $n_{j(k)} \to +\infty$ and a decreasing sequence of sets H_k such that

$$\bigcap_{k\geq 1} H_k = \emptyset, \quad H(O,u,y') \setminus H(u_{n_{j(k)}},\alpha_{n_{j(k)}}) \subset H_k \subset H(O,u,y')$$

and it follows that $(H(O,u,y') \setminus H_k)_{k\geq 1}$ is increasing with

$$\lim_{k\to\infty} \uparrow (H(O,u,y') \setminus H_k) = \bigcup_{k\geq 1} (H(O,u,y') \setminus H_k) = H(O,u,y') \setminus \bigcap_{k\geq 1} H_k$$

hence

$$\bigcup_{k\geq 1} (H(O,u,y') \setminus H_k) = H(O,u,y').$$

By the lower continuity property of P and $(\mathbf{A_0^-})$, we get

$$\begin{aligned}
\alpha &\leq P(H(O,u,y')) = \lim_{k\to\infty} \uparrow P(H(O,u,y') \setminus H_k)) \\
&\leq \lim_{k\to\infty} P(H(O,u,y') \cap H(u_{n_{j(k)}},\alpha_{n_{j(k)}})) \\
&\leq \lim_{k\to\infty} P(H(u_{n_{j(k)}},\alpha_{n_{j(k)}})) = \alpha
\end{aligned}$$

and consequently,

$$P(H(O,u,y,y')) = P((H(O,u,y') \setminus H(u,\alpha)) = 0$$

then $P(H(O,u,y,y')) = 0$ which contradicts $\mathbf{(A_0^+)}$. The case $y_\infty < Y(O,u,\alpha)$ is analogous.

Sufficient Condition. We prove that if $(u,\alpha) \to Y(O,u,\alpha)$ is bounded and continuous on $\mathbb{S}_{d-1} \times \Delta$ then $\mathbf{(A_0^+)}$ holds true. To do so, we show that

$$\neg\mathbf{(A_0^+)} \Rightarrow Y(O,\cdot,\cdot) \text{ is not continuous on } \mathbb{S}_{d-1} \times \Delta$$

where $\neg\mathbf{(A_0^+)}$ is the converse property of $\mathbf{(A_0^+)}$. Suppose that $\neg\mathbf{(A_0^+)}$ holds true and $(u,\alpha) \mapsto Q(O,u,\alpha)$ is bounded and continuous on $\mathbb{S}_{d-1} \times \Delta$. By $\neg\mathbf{(A_0^+)}$, there exists $\varepsilon_0 > 0$ such that $B_0 \in \mathcal{B}_{\varepsilon_0}$ with $P(B_0) = 0$. Pick $u \in \mathbb{S}_{d-1}$ such that $B_0 = H(O,u,y,y+\varepsilon_0)$ and $\alpha_0 = P(H(O,u,y+\varepsilon_0))$. We have $\alpha_0 = P(H(O,u,y+\varepsilon_0)) = P(H(O,u,y) \cup B_0) = P(H(O,u,y))$ then $Y(O,u,\alpha_0) \le y$. Let $(\alpha_k^+)_{k\in\mathbb{N}}$ be a strictly decreasing sequence with $\alpha_k^+ \downarrow \alpha_0$. Under $\mathbf{(A_0^-)}$ we have $H(O,u,y+\varepsilon_0) \subsetneq H(u,\alpha_k^+)$ hence $Y(O,u,\alpha_k^+) \ge y+\varepsilon_0$. By continuity of $Y(O,\cdot,\cdot)$ we get

$$\lim_{k\to\infty} Y(O,u,\alpha_k^+) = Y(O,u,\alpha_0) \ge y+\varepsilon_0$$

and consequently $y+\varepsilon_0 \le Y(O,u,\alpha_0) \le y$ which contradicts $\varepsilon_0 > 0$.

4.2 Proof of Proposition 2.5

The assumption $\mathbf{(A_0^-)}$ implies $\mathbf{(A_{0,\Psi}^-)}$, since for $y = Y(O,u,\alpha)$ with $u \in \mathbb{S}_{d-1}$ and $\alpha \in \Delta$ we have by the continuity property of P

$$\lim_{\varepsilon\to 0} \Psi(\varepsilon) \le \lim_{\varepsilon\to 0} P(H(O,u,y,y+\varepsilon)) = P(\partial H(u,\alpha)) = 0.$$

It remains to show that under $\mathbf{(A_0^-)}$ the assumption $\mathbf{(A_0^+)}$ implies $\mathbf{(A_{0,\Psi}^+)}$. Suppose that $\Psi(\varepsilon_0) = 0$ for some $\varepsilon_0 > 0$. There exists u_k in $\mathbb{S}_{d-1}$ and $y_k, y_k+\varepsilon_0 \in \mathcal{Y}_\Delta(O,u_k)$ such that

$$\lim_{k\to\infty} P(H(O,u_k,y_k,y_k+\varepsilon_0)) = 0. \tag{4.1}$$

Since $\mathcal{Y}_\Delta(O) = \bigcup_{u\in\mathbb{S}_{d-1}} \mathcal{Y}_\Delta(O,u)$ is compact, we can extract a subsequence $(u'_k, y'_k) \in \mathbb{S}_{d-1} \times \mathcal{Y}_\Delta(O,u'_k)$ having limit (u_0,y_0). We have

$$\mathcal{Y}_\Delta(O,u'_k) = [Y(O,u'_k,\alpha^-),\ Y(O,u'_k,\alpha^+)]$$

so by continuity of $u \to Y(O,u,\alpha)$ we get that $(u_0,y_0) \in \mathbb{S}_{d-1} \times \mathcal{Y}_\Delta(O,u_0)$, i.e. for $(u_0, y_0+\varepsilon_0)$. Set

$$B'_k = H(O,u'_k,y'_k,y'_k+\varepsilon_0).$$

By (4.1) we have $\lim_{k\to\infty} P(B'_k) = 0$. We now show that

$$\mathbb{1}_{B_0} \geqslant \lim_{k\to\infty} \mathbb{1}_{B'_k} \geqslant \mathbb{1}_{B_0\setminus\partial B_0} = \mathbb{1}_{B_0} - \mathbb{1}_{\partial B_0}$$

where $B_0 = H(O, u_0, y_0, y_0 + \varepsilon_0)$. First, if $x \notin B_0$ then there exists a $\delta-$neighborhood V_δ of (u_0, y_0) in $\mathbb{S}_{d-1}\times\mathbb{R}$ such that $x \notin \bigcup_{(u,y)\in V_\delta} H(O, u, y, y+\varepsilon_0)$ thus for every k big enough, $x \notin B'_k$. If $x \in \partial B_0$ we always have $\mathbb{1}_{B'_k}(x) \geqslant \mathbb{1}_{B_0\setminus\partial B_0}(x) = 0$. Finally, if $x \in B_0\setminus\partial B_0$ there exists a $\delta-$neighborhood V_δ of (u_0, y_0) in $\mathbb{S}_{d-1} \times \mathbb{R}$ such that $x \in \bigcap_{(u,y)\in V_\delta} H(O, u, y, y + \varepsilon_0)$ so for all k big enough, $x \in B'_k$. Consequently,

$$P(B_0) \geqslant \lim_{k\to\infty} P(B'_k) \geqslant P(B_0) - P(\partial B_0)$$

but by $(\mathbf{A_0^-})$ and $(\mathbf{A_0^+})$ we know that $P(B_0) > 0$ and $P(\partial B_0) = 0$. This implies that $\lim_{k\to\infty} P(B'_k) = P(B_0) > 0$, which is contradictory.

4.3 Proof of Proposition 2.6

The monotonous function Ψ has a right limit at any $\varepsilon_0 \geqslant 0$ and a left limit at any $\varepsilon_0 > 0$. Let $\varepsilon_k \downarrow \varepsilon_0 > 0$. For every $\theta > 0$ there exists $B_{\theta,0} \in \mathcal{B}_{\varepsilon_0}$ such that $B_{\theta,0} = H(O, u_\theta, y_\theta, y_\theta + \varepsilon_0)$ satisfies

$$(1 + \theta)\Psi(\varepsilon_0) > P(B_{\theta,0}) \geqslant \Psi(\varepsilon_0).$$

Consider a decreasing sequence of sets $B_{\theta,k} = H(O, u_\theta, y_\theta, y_\theta+\varepsilon_k)$ with limit $\bigcap_k B_{\theta,k} = B_{\theta,0}$ so that $P(B_{\theta,k}) \downarrow P(B_{\theta,0})$. There exists $k_\theta > 0$ such that for every $k \geqslant k_\theta$

$$(1 + \theta)\Psi(\varepsilon_0) > P(B_{\theta,k}) \geqslant P(B_{\theta,0}) \geqslant \Psi(\varepsilon_0).$$

Since Ψ is increasing we have $P(B_{\theta,k}) \geqslant \Psi(\varepsilon_k) \geqslant \Psi(\varepsilon_0)$. As $\Psi(\varepsilon_k)$ converges to a right limit $\Psi(\varepsilon_0^+)$ at ε_0, we have for every $\theta > 0$,

$$(1 + \theta)\Psi(\varepsilon_0) > \lim_{k\to\infty} \Psi(\varepsilon_k) = \Psi(\varepsilon_0^+) \geqslant \Psi(\varepsilon_0).$$

In other words $\lim_{k\to\infty} \Psi(\varepsilon_k) = \Psi(\varepsilon_0^+) = \Psi(\varepsilon_0)$. Likewise, if $\varepsilon_k \uparrow \varepsilon_0 > 0$ then to every $\theta > 0$ we associate a sequence $B_{\theta,k} \in \mathcal{B}_{\varepsilon_k}$ such that

$$(1 + \theta)\Psi(\varepsilon_k) > P(B_{\theta,k}) > \Psi(\varepsilon_k)$$

and by compacity in u and y we can extract a stabilized sequence,

$$B_{\theta,k_n} = H(O, u_{k_n}, y_{k_n}, y_{k_n} + \varepsilon_{k_n})$$

with $u_{k_n} \to u_\theta$, $y_{k_n} \to y_\theta$. We set $B_{\theta,0} = H(O, u_\theta, y_\theta, y_\theta + \varepsilon_0)$. Under $(\mathbf{A_0^-})$ and $(\mathbf{A_0^+})$ by Proposition 2.1 we have $P(B_{\theta,k_n}) \to P(B_{\theta,0})$ hence for all $k \geqslant k_\theta$ it follows that

$$\lim_{n\to\infty} P(B_{\theta,k_n}) \to \Psi(\varepsilon_0^-) \geqslant (1 - \theta)P(B_{\theta,0}) \geqslant (1 - \theta)\Psi(\varepsilon_0)$$

$$\Psi(\varepsilon_0) \geqslant \Psi(\varepsilon_0^-) \geqslant (1 - \theta)\Psi(\varepsilon_0)$$

for every $\theta > 0$. Therefore, $\Psi(\varepsilon_0^-) = \Psi(\varepsilon_0)$.

4.4 Proof of Proposition 2.7

Under $(\mathbf{A_0^-})$ we want to show that we almost surely have for all O, u, α and $n > d$ that

$$\alpha \leqslant P_n(H(O,u,Y_n(O,u,\alpha))) = P_n(H_n(u,\alpha))) \leqslant \alpha + \frac{d}{n}.$$

By definition of $Y_n(O,u,\alpha)$ we have $P_n(H(O,u,Y_n(O,u,\alpha))) \geqslant \alpha$. Fix $n > d$. Under $(\mathbf{A_0^-})$ the probability that $X_1, ..., X_{d+1}$ stand on the same hype-plan is null. As a matter of fact, by denoting $\partial H(x_1, ..., x_d)$ the unique hyper-plan containing d distinct points $x_1, ..., x_d$ we have

$$\begin{aligned}
&\mathbb{P}\left(X_{d+1} \in \partial H(X_1, ..., X_d)\right) \\
&= \int_{x_1 \in \mathbb{R}^d} ... \int_{x_d \in \mathbb{R}^d} \mathbb{P}\left(X_{d+1} \in \partial H(x_1, ..., x_d) \ \mid \ X_1 = x_1, ..., X_d = x_d\right) dP(x_1)...dP(x_d) \\
&= \int_{x_1 \in \mathbb{R}^d} ... \int_{x_d \in \mathbb{R}^d} \mathbb{P}\left(X_{d+1} \in \partial H(x_1, ..., x_d)\right) dP(x_1)...dP(x_d) = 0
\end{aligned}$$

since $\mathbb{P}\left(X_{d+1} \in \partial H\right) = 0$ for all hyper-plan ∂H, by $(\mathbf{A_0^-})$. It follows that

$$\begin{aligned}
&\mathbb{P}\left(\left\{X_{i_{d+1}} \in \partial H(X_{i_1}, ..., X_{i_d}), \text{ for distinct } i_1, ..., i_{d+1}\right\}\right) \\
&\leqslant \sum_{0 \leqslant i_1, ..., i_{d+1} \leqslant n} \mathbb{P}\left(X_{i_{d+1}} \in \partial H(X_{i_1}, ..., X_{i_d})\right) = 0
\end{aligned}$$

Therefore, almost surely, no hyper-plan contains more than d sample points,

$$\mathbb{P}\left(\sup_{H \in \mathcal{H}} P_n(\partial H) \geqslant \frac{d+1}{n}\right) = 0.$$

By denoting $\operatorname{int}(H(O,u,y)) = \left\{x \in \mathbb{R}^d : \langle x - O, u\rangle < y\right\}$ we have, with probability one, for all u, α

$$\begin{aligned}
P_n(H_n(u,\alpha)) &= P_n(\operatorname{int}(H_n(u,\alpha))) + P_n(\partial H_n(u,\alpha)) \\
&\leqslant P_n(\operatorname{int}(H_n(u,\alpha))) + \frac{d}{n}.
\end{aligned}$$

We also have $P_n(\operatorname{int}(H_n(u,\alpha))) \leqslant \alpha$ because if $P_n(\operatorname{int}(H_n(u,\alpha))) > \alpha$ then there is at least $\lceil n\alpha \rceil$ points $X_i \in \operatorname{int}(H_n(u,\alpha))$ hence we have $\langle X_i, u\rangle < Y_n(O,u,\alpha)$ and by denoting

$$\mathring{Y}_n(O,u,\alpha) = \max_{X_i \in \operatorname{int}(H_n(u,\alpha))} (\langle X_i, u\rangle) < Y_n(O,u,\alpha)$$

it follows that

$$P_n(H(O,u,\mathring{Y}_n(O,u,\alpha))) \geqslant \frac{\lceil n\alpha \rceil}{n} \geqslant \alpha$$

which contradicts the definition

$$Y_n(O,u,\alpha) = \inf\left\{y \in \mathbb{R} : P_n(H(O,u,y)) \geqslant \alpha\right\} \leqslant \mathring{Y}_n(O,u,\alpha).$$

4.5 Proof of uniform consistency with rate

Proof of Theorem 3.1. Under $(\mathbf{A_0^-})$ and $(\mathbf{A^+_{0,\Psi}})$ suppose that there exists $\delta > 0$ and an increasing random sequence $n_k \to \infty$ such that

$$|Y_{n_k}(O, u_{n_k}, \alpha_{n_k}) - Y(O, u_{n_k}, \alpha_{n_k})| > \delta.$$

Let $(u_{n'_k}, \alpha_{n'_k})$ be a subsequence on $\mathbb{S}_{d-1} \times [\alpha^-, \alpha^+] \subset \mathbb{S}_{d-1} \times (1/2, 1)$ with $u_{n'_k} \neq u_0$ and $\alpha_{n'_k} \neq \alpha_0$ and $u_{n'_k} \to u_0$ and $\alpha_{n'_k} \to \alpha_0$. It is possible to extract from (n'_k) an increasing sequence (m_k) with $m_k \to \infty$ such that either

$$Y(O, u_{m_k}, \alpha_{m_k}) - Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta \tag{4.2}$$

or $Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) - Y(O, u_{m_k}, \alpha_{m_k}) \geq \delta$. We assume (4.2) and we set

$$A_k = H_{m_k}(O, u_{m_k}, \alpha_{m_k}), \quad C_k = H(O, u_{m_k}, \alpha_{m_k}), \quad B_k = C_k \setminus A_k.$$

Since $\mathcal{H}$ is a VC-class we have

$$\lim_{n\to\infty} \sup_{H\in\mathcal{H}} |P_n(H) - P(H)| = 0, \quad a.s.$$

Under $(\mathbf{A_0^-})$, Proposition 2.7 implies

$$\begin{aligned} \sup_{H\in\mathcal{H}} |P_{m_k}(H) - P(H)| &\geq P_{m_k}(A_k) - P(A_k) \\ &\geq \alpha_{m_k} - \frac{d}{m_k} - (\alpha_{m_k} - P(B_k)) \\ &\geq -\frac{d}{m_k} + \Psi(\delta) \end{aligned}$$

so that we have,

$$\Psi(\delta) \leq \sup_{H\in\mathcal{H}} |P_{m_k}(H) - P(H)| + \frac{d}{m_k}. \tag{4.3}$$

Therefore there exists $\delta > 0$ such that $\Psi(\delta) = 0$ which contradict $(\mathbf{A^+_{0,\Psi}})$. In the alternative case of (4.2), a similar arguments holds.

Proof of Theorem 3.5 Under $(\mathbf{A_0^-})$ and $(\mathbf{A^+_{0,\Psi}})$ suppose that there exists a random increasing sequence $n_k \to \infty$ such that

$$|Y_{n_k}(O, u_{n_k}, \alpha_{n_k}) - Y(O, u_{n_k}, \alpha_{n_k})| > \delta_{n_k} = \Psi^{-1}\left(\sqrt{\frac{\log\log n_k}{n_k}}\right).$$

Let $(u_{n'_k}, \alpha_{n'_k})$ be a sequence of $\mathbb{S}_{d-1} \times [\alpha^-, \alpha^+] \subset \mathbb{S}_{d-1} \times (0, 1)$ with $u_{n'_k} \neq u_0$ and $\alpha_{n'_k} \neq \alpha_0$ and $u_{n'_k} \to u_0$ and $\alpha_{n'_k} \to \alpha_0$. There exists an increasing sequence $(m_k)_{k\geq 1}$ such that $m_k \to \infty$ and $Y(O, u_{m_k}, \alpha_{m_k}) - Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta_{m_k}$. We set

$$A_k = H_{m_k}(O, u_{m_k}, \alpha_{m_k}), \quad C_k = H(O, u_{m_k}, \alpha_{m_k}), \quad B_k = C_k \setminus A_k.$$

Since $\mathcal{H}$ is a VC-class, by the law of the iterated logarithm of Alexander [1] we know that

$$\limsup_{n\to\infty} \frac{\|P_n - P\|_{\mathcal{H}}}{\sqrt{(\log\log n)/n}} \leq \frac{\sqrt{2}}{2} \quad a.s.$$

since $4/5 > \sqrt{2}/2$, there exists $k(\omega) > 0$ such that for all $k \geq k(\omega)$

$$\frac{4}{5}\sqrt{\frac{\log\log m_k}{m_k}} \geq \sup_{H\in\mathcal{H}} |P_{m_k}(H) - P(H)|.$$

Under $(\mathbf{A_0^-})$ by PROPOSITIONS 2.6 and 2.7 we have

$$\begin{aligned}
\sup_{H\in\mathcal{H}} |P_{m_k}(H) - P(H)| &\geq P_{m_k}(A_k) - P(A_k) \\
&\geq \alpha_{m_k} - \frac{d}{m_k} - (\alpha_{m_k} - P(B_k)) \\
&\geq -\frac{d}{m_k} + \Psi(\delta_n) \\
&\geq \sqrt{\frac{\log\log m_k}{m_k}} - \frac{d}{m_k}
\end{aligned}$$

hence

$$\frac{4}{5}\sqrt{\frac{\log\log m_k}{m_k}} \geq \sup_{H\in\mathcal{H}} |P_{m_k}(H) - P(H)| \geq \sqrt{\frac{\log\log m_k}{m_k}} - \frac{d}{m_k}. \tag{4.4}$$

This implies that $1 \geq \frac{1}{5d}\sqrt{m_k \log\log m_k}$ which is absurd, so we have

$$\limsup_{n\to\infty} \frac{\|Y_n - Y\|_{\mathbb{S}_{d-1}\times\Delta}}{\Psi^{-1}\left(\sqrt{(\log\log n)/n}\right)} < \infty \quad a.s.$$

the case when $Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) - Y(O, u_{m_k}, \alpha_{m_k}) \geq \delta_{m_k}$ is identical.

4.6 Preliminary to the proofs of the main theorem

Let us write the empirical process indexed in different ways, as follows. For $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$, $y \in \mathbb{R}$, $\alpha \in \Delta$ and $H \in \mathcal{H}$,

$$\alpha_n(O, u, y) = \sqrt{n}\left(P_n(H(O, u, y)) - P(H(O, u, y))\right),$$

$$\mathbb{E}_n(u,\alpha) = \sqrt{n}\left(P_n(H(u,\alpha)) - P(H(u,\alpha))\right),$$
$$\Lambda_n(H) = \sqrt{n}(P_n(H) - P(H)),$$

and the quantile process

$$\mathbb{D}_n(u,\alpha) = \sqrt{n}\left(Y_n(O,u,\alpha) - Y(O,u,\alpha)\right).$$

Thus, for $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ and $\alpha \in \Delta$, we have

$$\alpha_n(O,u,Y(O,u,\alpha)) = \mathbb{E}_n(u,\alpha) = \Lambda_n(H(u,\alpha)) \tag{4.5}$$

and the increments

$$\begin{aligned} \Lambda_n(H(O,u,y,y+\varepsilon)) &= \sqrt{n}\left(P_n(H(O,u,y,y+\varepsilon)) - P(H(O,u,y,y+\varepsilon))\right) \\ &= \Lambda_n(H(O,u,y+\varepsilon)) - \Lambda_n(H(O,u,y)). \end{aligned}$$

For $n \geq 3$, $C > 1$, denote $\varepsilon_n = C\sqrt{\dfrac{\log\log n}{n}}$ and

$$\mathcal{B}_n = \bigcup_{0<\varepsilon<\varepsilon_n} \mathcal{B}_\varepsilon, \quad \mathcal{F}_n = \{\mathbb{1}_B : \ B \in \mathcal{B}_n\}.$$

The next proposition is crucial for the upcoming proofs. It's about the sharp control of the modulus of continuity of the empirical process Λ_n for the bands of width smaller than ε_n.

Proposition 4.1. *Under* ***(A₁)****, for all $\zeta > 1$ there exists $C_0, C_1 > 0$, then for all $n \geq 3$ we have*

$$\mathbb{P}\left\{\|\Lambda_n\|_{\mathcal{B}_n} \geq C_0 \frac{(\log n)^{1/2}(\log\log n)^{1/4}}{n^{1/4}}\right\} \leq \frac{C_1}{n^\zeta}.$$

Proof. Let $n \geq 3$. By Remark 5.1 from the appendix, the class $\mathcal{F}_n$ satisfies **(F.i)** and **(F.ii)**, thus by applying the Talagrand inequality [32] there exists $A_0, A_1 > 0$ such that

$$\begin{aligned} \mathbb{P}\left\{\|\Lambda_n\|_{\mathcal{B}_n} \geq A_0\left(\mathbb{E}\left(\sup_{B\in\mathcal{B}_n}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n \tau_i \mathbb{1}_{X_i\in B}\right|\right) + t_n\right)\right\} \\ \leq 2\exp\left(-\frac{A_1 t_n^2}{\sigma_n^2}\right) + 2\exp\left(-A_1 t_n\sqrt{n}\right) \end{aligned}$$

with

$$\sigma_n^2 = \sup_{B\in\mathcal{B}_n} Var(\mathbb{1}_{X\in B}), \quad t_n = \sqrt{\frac{CM\zeta}{A_1}}\frac{(\log n)^{1/2}(\log\log n)^{1/4}}{n^{1/4}}. \tag{4.6}$$

By $\mathbf{(A_1)}$ we have

$$Var(\mathbb{1}_{X\in B}) = P(B)\left(1 - P(B)\right) \leq M\varepsilon_n(1 - m\varepsilon_n) \leq M\varepsilon_n.$$

Thus,

$$\exp\left(-\frac{A_1 t_n^2}{\sigma_n^2}\right) \leq \exp\left(-\frac{A_1 t_n^2}{M\varepsilon_n}\right) = \frac{1}{n^\zeta}.$$

Moreover, for $n \geq 3$ we have

$$\exp\left(-A_1 t_n \sqrt{n}\right) = \exp\left(-\sqrt{MCA_1\zeta}(\log n)^{1/2}(n\log\log n)^{1/4}\right) \leq \frac{C_1'}{n^\zeta}$$

By Remark 5.2 the class $\mathcal{F}_n$ obeys the conditions of Theorem 5.2 thus there exists $A_2 > 0$ such that for all $n \geq n_0$ we have

$$\begin{aligned} \mathbb{E}\left(\sup_{B\in\mathcal{B}_n}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n \tau_i \mathbb{1}_{X_i\in B}\right|\right) &\leq A_2\sqrt{vM\varepsilon_n \log\left(1 \vee 1/\sqrt{M\varepsilon_n}\right)} \\ &\leq C_0' \frac{(\log n)^{1/2}(\log\log n)^{1/4}}{n^{1/4}} \end{aligned}$$

hence the result is proved. □

4.7 Proof of the main theorem

Preliminary step For $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ and $\alpha \in \Delta$, $\gamma > 0$

$$y_\alpha = Y(O,u,\alpha) \quad \text{and} \quad v_\gamma(y_\alpha) = \{y \in \mathbb{R},\ |y - y_\alpha| < \gamma\}.$$

We know that $\lim_{n\to\infty} \|Y_n(O,\, u,\, \alpha) - y_\alpha\|_{\mathbb{S}_{d-1}\times\Delta} = 0$ *a.s.* then there exists $\gamma_0 > 0$ such that for all $0 < \gamma < \gamma_0$ we have for $n \geq n(\omega, \gamma)$

$$\begin{aligned} & Y_n(O,\, u,\, \alpha) \\ &= \inf_{y\in v_\gamma(y_\alpha)} \{P_n(H(O,u,y)) \geq \alpha\} \\ &= \inf_{y\in v_\gamma(y_\alpha)} \{P_n(H(O,u,y)) - P(H(O,u,y)) \geq \alpha - P(H(O,u,y))\} \\ &= \inf_{y\in v_\gamma(y_\alpha)} \left\{P_n(H(O,u,y)) - P(H(O,u,y)) \geq F_{\langle X-O,u\rangle}(y_\alpha) - F_{\langle X-O,u\rangle}(y)\right\} \\ &= \inf_{y\in v_\gamma(y_\alpha)} \left\{\alpha_n(O,u,y) \geq \sqrt{n}\left(F_{\langle X-O,u\rangle}(y_\alpha) - F_{\langle X-O,u\rangle}(y)\right)\right\} \end{aligned}$$

with $\alpha_n(O,u,y) = \sqrt{n}\left(P_n(H(O,u,y)) - P(H(O,u,y))\right)$. Under $(\mathbf{A_1})$ and $(\mathbf{A_2})$, $y \mapsto F_{\langle X-O,u\rangle}(y)$ is continuous and differentiable on $\mathbb{R}$ thus by Taylor expansion to the first order in the neighborhood of y_α, we have for all $y \in v_\gamma(y_\alpha)$

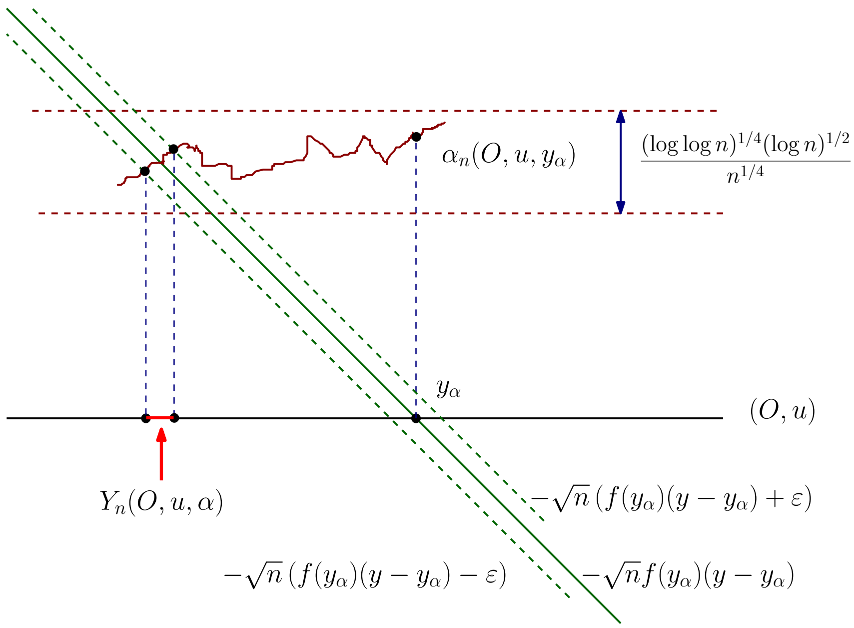
$$F_{\langle X-O,u\rangle}(y_\alpha) - F_{\langle X-O,u\rangle}(y) = f_{\langle X-O,u\rangle}(y_\alpha)(y_\alpha - y) + \varepsilon_\gamma(u,\alpha,y_\alpha - y)$$

with

$$\lim_{\gamma\to 0}\sup_{u\in\mathbb{S}_{d-1}}\sup_{\alpha\in\Delta}|\varepsilon_\gamma(u,\alpha,y_\alpha - y)| = 0.$$

From now on, we study the following,

$$\begin{aligned}&Y_n(O,\, u,\, \alpha)\\&= \inf_{y\in v_\gamma(y_\alpha)} \left\{\alpha_n(O,u,y) \geq \sqrt{n}\left(f_{\langle X-O,u\rangle}(y_\alpha)(y_\alpha - y) + \varepsilon_\gamma(u,\alpha,y_\alpha - y)\right)\right\}.\end{aligned}$$



Step I Under $(\mathbf{A_1})$ and $(\mathbf{A_2})$, we show that

$$\lim_{n\to\infty} \left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1}\times\Delta} = 0 \quad a.s. \tag{4.7}$$

By LEMMA 5.2 there exists $C_\Delta > 0$ and $n(\omega) > 0$ such that for all $n \geq n(\omega)$, we have for all $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ and $\alpha \in \Delta$, $Y_n(O,u,\alpha) \in v_{\gamma_n}(y_\alpha)$ where

$$v_{\gamma_n}(y_\alpha) = [y_\alpha - \gamma_n, y_\alpha + \gamma_n]\,,\ \gamma_n = C_\Delta\sqrt{\frac{\log\log n}{n}}.$$

For all $y \in v_{\gamma_n}(y_\alpha)$, denote

$$z_n(O,u,y,y_\alpha) = \alpha_n(O,u,y_\alpha) - \alpha_n(O,u,y) = \Lambda_n(H(O,u,y_\alpha,y))$$

the increments of the empirical process Λ_n on the bands of width less than γ_n. By PROPOSITION 4.1,

(4.8)

$$\sup_{u\in\mathbb{S}^d}\sup_{\alpha\in\Delta}\sup_{y\in v_{\gamma_n}(y_\alpha)} |z_n(O,u,y,y_\alpha)| = O_{a.s.}\left(\frac{(\log n)^{1/2}(\log\log n)^{1/4}}{n^{1/4}}\right)$$

hence

$$\lim_{n\to\infty} \sup_{u\in\mathbb{S}^d} \sup_{\alpha\in\Delta} \sup_{y\in v_{\gamma_n}(y_\alpha)} |z_n(O,u,y,y_\alpha)| = 0 \quad a.s. \tag{4.9}$$

and for all $n \geq n(\omega)$, we get

$$\begin{aligned}
& Y_n(O,\, u,\, \alpha) \\
&= \inf_{y\in v_{\gamma_n}(y_\alpha)} \left\{\alpha_n(O,u,y) \geq \sqrt{n}\left(f_{\langle X-O,u\rangle}(y_\alpha)(y_\alpha - y) + \varepsilon_{\gamma_n}(u,\alpha,y_\alpha - y)\right)\right\} \\
&= \inf_{y\in v_{\gamma_n}(y_\alpha)} \left\{ y \geq y_\alpha - \frac{\alpha_n(O,u,y)}{n^{1/2} f_{\langle X-O,u\rangle}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(u,\alpha,y_\alpha - y)}{f_{\langle X-O,u\rangle}(y_\alpha)} \right\} \\
&= \inf_{y\in v_{\gamma_n}(y_\alpha)} \left\{ y \geq y_\alpha - \frac{\alpha_n(O,u,y_\alpha)}{n^{1/2} f_{\langle X-O,u\rangle}(y_\alpha)} - \frac{z_n(O,u,y,y_\alpha)}{n^{1/2} f_{\langle X-O,u\rangle}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(u,\alpha,y_\alpha - y)}{f_{\langle X-O,u\rangle}(y_\alpha)} \right\}.
\end{aligned}$$

Since $|Y_n(O,u,\alpha)| < \infty$, we have on the one hand,

$$\begin{aligned}
Y_n(O,u,\alpha) &\geq y_\alpha - \frac{\alpha_n(O,u,y_\alpha)}{n^{1/2} f_{\langle X-O,u\rangle}(y_\alpha)} \\
&\quad + \inf_{u\in\mathbb{S}^d} \inf_{\alpha\in\Delta} \inf_{y\in v_{\gamma_n}(y_\alpha)} \left(- \frac{z_n(O,u,y,y_\alpha)}{n^{1/2} f_{\langle X-O,u\rangle}(y_\alpha)} + \frac{\varepsilon_{\gamma}(u,\alpha,y_\alpha - y)}{f_{\langle X-O,u\rangle}(y_\alpha)} \right) \\
&\geq y_\alpha - \frac{\alpha_n(O,u,y_\alpha)}{n^{1/2} f_{\langle X-O,u\rangle}(y_\alpha)} - \frac{\Theta_{1,n}}{n^{1/2}} - \Theta_{2,n}
\end{aligned}$$

where

$$\Theta_{1,n} = \sup_{u\in\mathbb{S}^d} \sup_{\alpha\in\Delta} \sup_{y\in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O,u,y,y_\alpha)}{f_{\langle X-O,u\rangle}(y_\alpha)} \right| \tag{4.10}$$

$$\Theta_{2,n} = \sup_{u\in\mathbb{S}^d} \sup_{\alpha\in\Delta} \sup_{y\in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u,\alpha,y_\alpha - y)}{f_{\langle X-O,u\rangle}(y_\alpha)} \right|. \tag{4.11}$$

Likewise, we have

$$Y_n(O,u,\alpha) \leq y_\alpha - \frac{\alpha_n(O,u,y_\alpha)}{n^{1/2} f_{\langle X-O,u\rangle}(y_\alpha)} + \frac{\Theta_{1,n}}{n^{1/2}} + \Theta_{2,n}$$

thus

$$\left| n^{1/2}(Y_n(O,u,\alpha) - y_\alpha) + \frac{\alpha_n(O,u,y_\alpha)}{f_{\langle X-O,u\rangle}(y_\alpha)} \right| \leq \Theta_{1,n} + n^{1/2}\Theta_{2,n}. \tag{4.12}$$

By $(\mathbf{A_1})$, we have

$$\Theta_{1,n} \leq \frac{1}{m} \sup_{u\in\mathbb{S}^d} \sup_{\alpha\in\Delta} \sup_{y\in v_{\gamma_n}(y_\alpha)} |z_n(O,u,y,y_\alpha)|\,.$$

Hence by (4.9), we have $\lim_{n\to\infty}\Theta_{1,n}=0$ $a.s.$ Observe that ρ of (3.3) can be written as

$$\rho(\gamma)=\sup_{u\in\mathbb{S}^d}\sup_{\alpha\in\Delta}\sup_{y\in v_\gamma(y_\alpha)}\left|\varepsilon_\gamma(u,\alpha,y_\alpha-y)\right|.$$

Since $f_{\langle X-O,u\rangle}(y_\alpha)=h(u,\alpha)$, under $\mathbf{(A_1)}$ we obtain

$$n^{1/2}\Theta_{2,n}\le\frac{n^{1/2}\rho(\gamma_n)}{m}. \tag{4.13}$$

By $\mathbf{(A_2)}$, we have

$$\lim_{n\to\infty}\frac{\sqrt{\log\log(1/\gamma_n)}}{\gamma_n}\rho(\gamma_n)=\lim_{n\to\infty}n^{1/2}\rho(\gamma_n)=0$$

hence $\lim_{n\to\infty}n^{1/2}\Theta_{2,n}=0$ thus,

$$\lim_{n\to\infty}\left|n^{1/2}(Y_n(O,u,\alpha)-y_\alpha)+\frac{\alpha_n(O,u,y_\alpha)}{f_{\langle X-O,u\rangle}(y_\alpha)}\right|=0\quad a.s.$$

and with previous notation

$$\frac{\alpha_n(O,u,y_\alpha)}{f_{\langle X-O,u\rangle}}=\frac{\mathbb{E}_n(u,\alpha)}{h(u,\alpha)}$$

then (4.7) holds.

Step II Under $\mathbf{(A_1)}$ and $\mathbf{(A_2)}$ we show that we can construct on the same probability space $(\Omega,\mathcal{T},\mathbb{P})$ and i.i.d. sequence (X_n) of law P and a sequence $(\mathbb{G}_n)$ of versions of $\mathbb{G}_P$ such that

$$\lim_{n\to\infty}\|\mathbb{E}_n-\mathbb{G}_n\|_{\mathbb{S}_{d-1}\times\Delta}=0\quad a.s. \tag{4.14}$$

The set $\mathcal{H}$ is a class of Vapnik-Chervonenkis, thus it is a Donsker class,

$$(\Lambda_n(H))_{H\in\mathcal{H}}\overset{\mathcal{L}aw}{\underset{n\to\infty}{\longrightarrow}}(\mathbb{G}_P(H))_{H\in\mathcal{H}}$$

with $\mathbb{G}_P$ a Brownian bridge indexed by $\mathcal{H}$ of covariance

$$cov(\mathbb{G}_P(H),\mathbb{G}_P(H'))=P(H\cap H')-P(H)P(H'),\quad H,H'\in\mathcal{H}.$$

Then, by applying Theorem 5.3 to Λ_n, we can construct on the same probability space $(\Omega,\mathcal{T},\mathbb{P})$ and i.i.d. sequence (X_n) of law P and a sequence $(\mathbb{G}_n)$ of versions of $\mathbb{G}_P$ such that

$$\Lambda_n(H(O,u,y))=\mathbb{G}_n(H(O,u,y))+\xi_n(H(O,u,y)) \tag{4.15}$$

with

$$\lim_{n\to\infty} \|\xi_n(H)\|_{\mathcal{H}} = 0 \quad a.s. \tag{4.16}$$

and for all $\theta > 1$ there exists $K_1 > 0$

$$\mathbb{P}\left(\sup_{u\in\mathbb{S}_{d-1}} \sup_{\alpha\in\Delta} |\xi_n(u,\alpha)| \geq K_1 \frac{(\log n)^{w_d}}{n^{v_d}}\right) \leq \frac{1}{n^\theta} \tag{4.17}$$

with the notation $\xi_n(u,\alpha) = \xi_n(H(u,\alpha))$ and $v_d = 1/(2+10d)$, $w_d = (4+10d)/(4+20d)$. Consequently (4.14) holds.

Step III Under $(\mathbf{A_1})$ and $(\mathbf{A_2})$, we show that

$$\lim_{n\to\infty} \left\| \mathbb{D}_n + \frac{\mathbb{G}_n}{h} \right\|_{\mathbb{S}_{d-1}\times\Delta} = 0 \quad a.s.$$

By **Step I** we have

$$\lim_{n\to\infty} \left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1}\times\Delta} = 0 \quad a.s.$$

and by **Step II**

$$\lim_{n\to\infty} \|\mathbb{E}_n - \mathbb{G}_n\|_{\mathbb{S}_{d-1}\times\Delta} = 0 \quad a.s.$$

By $(\mathbf{A_1})$ the function h is bounded thus under $(\mathbf{A_1})$ and $(\mathbf{A_2})$, we have

$$\lim_{n\to\infty} \left\| \mathbb{D}_n + \frac{\mathbb{G}_n}{h} \right\|_{\mathbb{S}_{d-1}\times\Delta} = 0 \quad a.s.$$

which readily implies

$$\begin{aligned} \lim_{n\to\infty} d_{PL}(\sqrt{n}(Y_n - Y), -\frac{\mathbb{G}_n}{h}) &= \lim_{n\to\infty} d_{PL}(\sqrt{n}(Y_n - Y), -\frac{\mathbb{G}_P}{h}) \\ &= \lim_{n\to\infty} d_{PL}(\sqrt{n}(Y_n - Y), \frac{\mathbb{G}_P}{h}) = 0 \end{aligned}$$

where d_{PL} is Prokhorov-Levy distance. Therefore

$$\mathbb{D}_n = \sqrt{n}(Y_n - Y) \underset{n\to\infty}{\overset{\mathcal{L}aw}{\longrightarrow}} \tilde{\mathbb{G}} := \frac{\mathbb{G}_P}{h}$$

in the sense of the weak convergence on the space of bounded function on $\mathbb{S}_{d-1} \times \Delta$ endowed with the supremum norm. Note that

$$cov(\tilde{\mathbb{G}}(u,\alpha), \tilde{\mathbb{G}}(u',\alpha')) = \frac{P(H(u,\alpha) \cap H(u',\alpha')) - \alpha\alpha'}{h(u,\alpha)h(u',\alpha')}.$$

Step IV *(Rate in Bahadur-Kiefer representation).* We show that under $\mathbf{(A_1)}$, $\mathbf{(A_2)}$, $\mathbf{(A_3)}$, it holds

$$\left\|\mathbb{D}_n + \frac{\mathbb{E}_n}{h}\right\|_{\mathbb{S}_{d-1}\times\Delta} = O_{a.s.}\left(\frac{(\log\log n)^{1/4}(\log n)^{1/2}}{n^{1/4}}\right). \tag{4.18}$$

The class $\mathcal{H}$ is a Vapnik-Cervonenkis class of dimension $d+1$, so by the law of the iterated logarithm (Alexander 1984 [1]) we have

$$\limsup_{n\to\infty} \frac{\|\Lambda_n\|_{\mathcal{H}}}{\sqrt{2\log\log n}} \leq \frac{1}{2} \quad a.s.$$

then with probability 1, there exists $n(\omega) > 0$ such that for all $n \geq n(\omega)$, we have for all $u \in \mathbb{S}_{d-1}$

$$\alpha_n(O,u,y) := \Lambda_n(H(O,u,y)) \in \left[-\sqrt{\log\log n}, \sqrt{\log\log n}\right].$$

For all $n \geq n(\omega)$, recall (4.12)

$$\left| n^{1/2}(Y_n(O,u,\alpha) - y_\alpha) + \frac{\alpha_n(O,u,y_\alpha)}{f_{\langle X-O,u\rangle}(y_\alpha)}\right| \leq \Theta_{1,n} + n^{1/2}\Theta_{2,n} \tag{4.19}$$

and by (4.8) and $\mathbf{(A_1)}$ and $\mathbf{(A_2)}$ there exists $C' > 0$, such that for all $n \geq n(\omega)$

$$\Theta_{1,n} \leq \frac{C'}{m}\left(\frac{(\log n)^{1/2}(\log\log n)^{1/4}}{n^{1/4}}\right)$$

and by (4.13)

$$n^{1/2}\Theta_{2,n} \leq \frac{n^{1/2}\rho(\gamma_n)}{m}.$$

thus for $n \geq n(\omega)$,

$$\begin{aligned}
\left|\mathbb{D}_n(u,\alpha) + \frac{\mathbb{E}_n(u,\alpha)}{h(u,\alpha)}\right| &\leq \Theta_{1,n} + n^{1/2}\Theta_{2,n} \\
&\leq \frac{C'}{m}\left(\frac{(\log\log n)^{1/4}(\log n)^{1/2}}{n^{1/4}}\right) + \frac{n^{1/2}\rho(\gamma_n)}{m} \\
&\leq \frac{(\log\log n)^{1/4}(\log n)^{1/2}}{n^{1/4}m}\left(\frac{\rho(\gamma_n)(\log\log n)^{1/2}}{(\gamma_n)^{3/2}(\log n)^{1/2}} + C'\right) \\
&\leq \frac{t_n'}{m}\left(\frac{\rho(\gamma_n)(\log\log(1/\gamma_n))^{1/2}}{\gamma_n^{3/2}(\log(1/\gamma_n))^{1/2}}\left(\frac{\log(1/\gamma_n)(\log\log n)}{\log\log(1/\gamma_n)(\log n)}\right)^{1/2} + C'\right)
\end{aligned}$$

with $t_n' = n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}$. We have

$$\lim_{n\to\infty}\left(\frac{\log(1/\gamma_n)(\log\log n)}{\log\log(1/\gamma_n)(\log n)}\right)^{1/2} = \frac{1}{2}$$

and by $\mathbf{(A_3)}$

$$\lim_{n\to\infty} \frac{\rho(\gamma_n)(\log\log(1/\gamma_n))^{1/2}}{\gamma_n^{3/2}(\log(1/\gamma_n))^{1/2}} = \lim_{n\to\infty} \frac{\rho(\gamma_n)}{\gamma_n^{3/2}\sqrt{\log(1/\gamma_n)}} = 0.$$

Consequently, under $\mathbf{(A_1)}$, $\mathbf{(A_2)}$ and $\mathbf{(A_3)}$ we have proved (4.18).

Remark 4.1. *Under $\boldsymbol{(A_1)}$ and $\boldsymbol{(A_2)}$ we have*

$$\lim_{n\to\infty} n^{1/2}\rho(\gamma_n) = 0$$

without the additional assumptions $\boldsymbol{(A_3)}$ or $\boldsymbol{(A'_3)}$, we can only state that there exists $n(\omega) > 0$ such that for all $n \geq n(\omega)$

$$\left\|\mathbb{D}_n + \frac{\mathbb{E}_n}{h}\right\|_{\mathbb{S}_{d-1}\times\Delta} = O_{a.s.}\left(\left(n^{1/2}\rho(\gamma_n)\right) \vee \frac{(\log\log n)^{1/4}(\log n)^{1/2}}{n^{1/4}}\right).$$

Step V *(Rate of the Gaussian approximation).* We have shown that under $\mathbf{(A_1)}$ and $\mathbf{(A_2)}$, we can construct on the same probability space $(\Omega, \mathcal{T}, \mathbb{P})$ an i.i.d. sequence (X_n) of law P and $(\mathbb{G}_n)$ of versions of $\mathbb{G}$ such that for all $u \in \mathbb{S}_{d-1}$ and $\alpha \in \Delta$, we have

$$\mathbb{D}_n(u,\alpha) = -\frac{\mathbb{G}_n(u,\alpha)}{h(u,\alpha)} + \mathbb{Z}_n(u,\alpha)$$

$\lim_{n\to\infty} \|\mathbb{Z}_n\|_{\mathbb{S}_{d-1}\times\Delta} = 0$ *a.s.* We have

$$\left\|\mathbb{D}_n + \frac{\mathbb{G}_n}{h}\right\|_{\mathbb{S}_{d-1}\times\Delta} \leq \left\|\mathbb{D}_n + \frac{\mathbb{E}_n}{h}\right\|_{\mathbb{S}_{d-1}\times\Delta} + \frac{1}{m}\|\mathbb{E}_n - \mathbb{G}_n\|_{\mathbb{S}_{d-1}\times\Delta}.$$

Under $\mathbf{(A_3)}$, by (4.18), there exists $n(\omega) > 0$ and $C''_\Delta > 0$ such that for all $n \geq n(\omega)$, we have

$$\left\|\mathbb{D}_n + \frac{\mathbb{E}_n}{h}\right\|_{\mathbb{S}_{d-1}\times\Delta} \leq C''_\Delta \frac{(\log\log n)^{1/4}(\log n)^{1/2}}{n^{1/4}}$$

and by (4.17) and the Borel-Cantelli lemma, we have

$$\|\mathbb{E}_n - \mathbb{G}_n\|_{\mathbb{S}_{d-1}\times\Delta} = O_{a.s.}\left(\frac{(\log n)^{w_d}}{n^{v_d}}\right).$$

Conclusion Under $\mathbf{(A_1)}$ and $\mathbf{(A_2)}$ one can construct on the same probability space $(\Omega, \mathcal{T}, \mathbb{P})$ an i.i.d. sequence X_n with distribution P and a sequence $\mathbb{G}_n$ of versions of $\mathbb{G}_P$ in such a way that for $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$, $\alpha \in \Delta$

$$Y_n(O,u,\alpha) = Y(O,u,\alpha) + \frac{\mathbb{G}_n(u,\alpha)}{\sqrt{n}} + \frac{\mathbb{Z}_n(u,\alpha)}{\sqrt{n}}$$

where $\lim_{n\to\infty} \|\mathbb{Z}_n\|_{S_{d-1}\times\Delta} = 0$ *a.s.* If P moreover satisfies $(\mathbf{A_3})$ then $\mathbb{G}_n$ can be constructed such that for $v_d = 1/(2+10d)$ and $w_d = (4+10d)/(4+20d)$, there exists $n_\theta(m,M,d) > 0$ such that we have, for all $n > n_\theta$,

$$\mathbb{P}\left(\|\mathbb{Z}_n\|_{S_{d-1}\times\Delta} \geq c_\theta \frac{(\log n)^{w_d}}{n^{v_d}}\right) \leq \frac{1}{n^\theta}.$$

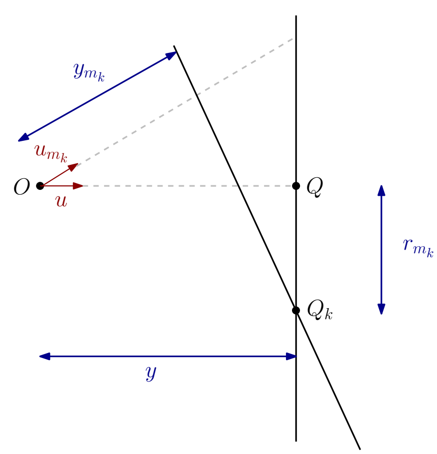
5 Appendix

5.1 Technical Lemmas

Lemma 5.1. *Let $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$, $y \in \mathbb{R}$ and $y_\infty \in \overline{\mathbb{R}}$ with $y \neq y_\infty$. For every sequence $(u_n)_{n\in\mathbb{N}}$ of $\mathbb{S}_{d-1}$ with $u_n \neq u$ and $u_n \to u$ and for every sequence $(y_n)_{n\in\mathbb{N}}$ of reals with $y_n \neq y_\infty$ and $y_n \to y_\infty$, there exists an increasing sequence of integers $(n_k)_{k\in\mathbb{N}}$ with $n_k \to \infty$ and a sequence of sets $(H_k)_{k\geq 1}$ such that*

$$H_{k+1} \subset H_k, \quad \bigcap\nolimits_{k\geqslant 1} H_k = \emptyset, \quad H(O,u_{n_k},y_{n_k})\backslash H(O,u,y) \subset H_k \subset H(O,u,y).$$

Proof. Let $u \in \mathbb{S}_{d-1}$, $y \in \mathbb{R}$ and $y_\infty \in \overline{\mathbb{R}}$ with $y < y_\infty$, and let $p_{H(O,u,y)}$ denote the orthogonal projection on $\partial H(O,u,y)$. We denote $Q = p_{H(O,u,y)}(O) = O + yu$. For (u_n) in $\mathbb{S}_{d-1}$ with $u_n \neq u$ and $\lim_{n\to\infty} u_n = u$ and (y_n) sequence of reals with $y_n \neq y_\infty$ and $\lim_{n\to\infty} y_n = y_\infty$, one can extract $((u_{m_k}, y_{m_k}))_{k\geq 1}$ in $\mathbb{S}_{d-1}\times\mathbb{R}$ such that $(\langle u_{m_k}, u\rangle)_{k\geq 1}$ is increasing with $\lim_{k\to\infty}\langle u_{m_k}, u\rangle = 1$.



We set $D_{m_k} = \partial H(O,u,y) \cap \partial H(O,u_{m_k},y_{m_k})$, which is not empty since $u_{m_k} \neq u$ and it is an hyper-plan of dimension $d-2$. Denote the distance between Q and D_{m_k} by $r_{m_k} = \inf_{Q_k\in D_{m_k}} \|Q_k - Q\|$.

Step I We show that

$$\lim_{k\to\infty} r_{m_k} = +\infty.$$

Fix $Q' = O + y'u$ with $y < y' < y_\infty$ an element from the line (O,u) and $A'_k = H(O, u_{m_k}, y'_{m_k})$ the half-space of normal u_{m_k} intersecting (O,u) exactly in Q', i.e.

$$A'_k \cap (O,u) = Q'.$$

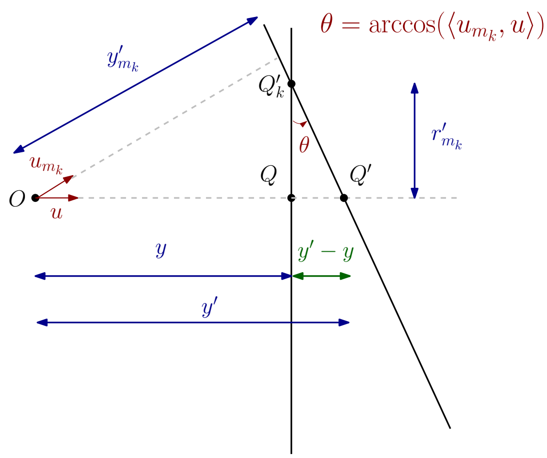
we can easily see that $y'_{m_k} = y'\langle u_{m_k}, u\rangle$, hence $(y_{m_k})_{k\geq 1}$ is increasing with $y'_{m_k} \to y'$. For k big enough, $y'_{m_k} < y_{m_k}$, thus $A'_k \subsetneq H(O, u_{m_k}, y_{m_k})$. From now on, denote

$$D'_{m_k} = \partial H(O,u,y) \cap \partial A'_k, \quad r'_{m_k} = \inf_{Q'_k \in D'_{m_k}} \|Q'_k - Q\| .$$

By observing that $A'_k \subsetneq H(O, u_{m_k}, y_{m_k})$, we have $r'_{m_k} < r_{m_k}$. Consequently, $r'_{m_k} = \dfrac{y'-y}{\tan(\arccos(\langle u_{m_k}, u\rangle))}$ and $r'_{m_k} \to \infty$, hence $r_{m_k} \to \infty$. Now, we can extract an increasing subsequence $(r_{n_k})_{k\geq 1}$ with

$$\lim_{k\to\infty} r_{n_k} = +\infty.$$

Step I figure



Step II We construct $(H_k)_{k\geq 1}$ of Lemma 5.1.
Let $k \geq 1$, define the set of directions $\mathbb{V}_k = \{v \in \mathbb{S}_{d-1} : \ \langle v,u\rangle = \langle u_{n_k}, u\rangle\}$

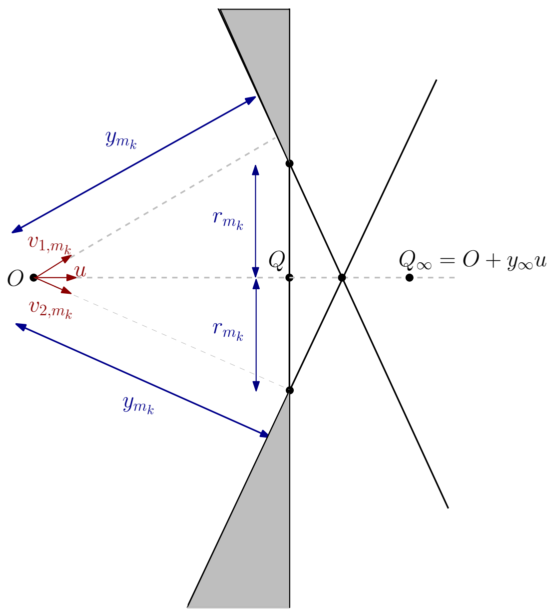
and the set of half-spaces $\mathbb{U}_k = \{H(O, v, y_{n_k}) : \quad v \in \mathbb{V}_k\}$ obtained by revolution of $H(O, u_{n_k}, y_{n_k})$ around (O, u). Finally, define

$$\mathbb{T}_k = \bigcap_{v \in \mathbb{V}_k} H(O, v, y_{n_k}) = \bigcap_{\hat{H} \in \mathbb{U}_k} \hat{H}$$

and

$$H_k = \bigcup_{v \in \mathbb{V}_k} H(O, u, y) \setminus H(O, v, y_{n_k}) = H(O, u, y) \setminus \mathbb{T}_k.$$

As $r_{n_k} \uparrow +\infty$ we have $H_{k+1} \subset H_k$. And since $u_{n_k} \neq u$ then for all $\hat{H} \in \mathbb{U}_k$ we have $H(O, u, y) \cap \hat{H} \neq \emptyset$ in particular, $H_k \neq \emptyset$. We have $H_k \subset H(O, u, y)$ and $u_{n_k} \in \mathbb{V}_k$ we get by definition of H_k, that $H(O, u, y) \setminus H(O, u_{n_k}, y_{n_k}) \subset H_k$.



In Grey $H_k = H(O, u, y) \setminus T_k$

We have $\lim_{k \to \infty} \langle u_{n_k}, u \rangle = 1$ then $V_\infty = \{u\}$ and $\lim_{k \to \infty} y_{n_k} = y$. Moreover $\bigcap_{k \geq 1} H_k = \emptyset$. The case where $y_\infty < y' < y$ is analogous. $\square$

Lemma 5.2. *Under **(A₁)**, almost surely there exists* $C_\Delta > 0$ *and* $n(\omega) > 3$ *such that for all* $n \geq n(\omega)$ *we have*

$$\|Y_n - Y\|_{\mathbb{S}_{d-1} \times \Delta} \leq C_\Delta \sqrt{\frac{\log \log n}{n}}.$$

Proof. Under $(\mathbf{A_1})$, we have

$$m\varepsilon \leq \Psi(\varepsilon) \leq M\varepsilon, \quad \varepsilon \geq 0.$$

By taking $\varepsilon = \Psi^{-1}\left(\sqrt{\log\log n/n}\right)$, we obtain by PROPOSITION 2.6 that for all $n > 3$

$$\Psi^{-1}\left(\sqrt{\frac{\log\log n}{n}}\right) \leq \frac{1}{m}\sqrt{\frac{\log\log n}{n}}. \tag{5.1}$$

and by THEOREM 3.5, we know that almost surely there exists $c_\Delta > 0$ and $n(\omega) > 3$ such that for all $n \geq n(\omega)$ we have

$$\|Y_n - Y\|_{\mathbb{S}_{d-1}\times\Delta} \leq c_\Delta \Psi^{-1}\left(\sqrt{\frac{\log\log n}{n}}\right)$$

and by (5.1) for $C_\Delta = c_\Delta/m$, we get

$$\|Y_n - Y\|_{\mathbb{S}_{d-1}\times\Delta} \leq C_\Delta\sqrt{\frac{\log\log n}{n}}.$$

□

5.2 Tools needed in the proof of main theorem

Let $\mathcal{F}$ be a class of measurable real valued functions of $\mathcal{X}$, suppose that

(F.i) for $S_* > 0$, for all $f \in \mathcal{F}$, $\sup_{x\in\mathcal{X}}|f(x)| \leq S_*/2$.

(F.ii) The class $\mathcal{F}$ is point-wide measurable, i.e. there exists a countable subclass $\mathcal{F}_\infty$ of $\mathcal{F}$ such that for every f there exists $(f_m)_{m\in\mathbb{N}} \subset \mathcal{F}_\infty^{\mathbb{N}}$ for which $\lim_{m\to\infty} f_m(x) = f(x)$ for all $x \in \mathcal{X}$.

the **(F.ii)** is set to avoid measurability problems and the use of outer integrals.

Theorem 5.1 (Talagrand Inequality [32])**.** *If $\mathcal{G}$ satisfies **(F.i)** and **(F.ii)** then for all $n \geq 1$ and $t > 0$ we have for finite constants $A_0 > 0$ and $A_1 > 0$*

$$\mathbb{P}\left\{\|\alpha_n\|_\mathcal{G} \geq A_0\left(\mathbb{E}\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \tau_i g(X_i)\right\|_\mathcal{G}\right) + t\right)\right\}$$
$$\leq 2\exp\left(-\frac{A_1 t^2}{\sigma_\mathcal{G}^2}\right) + 2\exp\left(-\frac{A_1 t\sqrt{n}}{S_*}\right)$$

*where $\sigma_\mathcal{G}^2 = \sup_{g\in\mathcal{G}} Var(g(X))$, and S_* from **(F.i)**.*

The constants A_0, A_1 are universals and do not depend in $\mathcal{G}$ and S_*.

Remark 5.1. *Let $n \geq 3$, $C > 1$, for $\varepsilon_n = C\sqrt{\dfrac{\log\log n}{n}}$, we set*

$$\mathcal{B}_n = \bigcup_{0<\varepsilon<\varepsilon_n} \mathcal{B}_\varepsilon, \quad \mathcal{F}_n = \{\mathbb{1}_B : \ B \in \mathcal{B}_n\}.$$

*$\mathcal{F}_n$ satisfies **(F.i)** and **(F.ii)**, as a matter of fact*

- *for all $g \in \mathcal{F}_n$ we have $\sup_{x\in\mathcal{X}}|g(x)| \leq 1 = 2/2$, thus, with notations of **(F.i)** we have $S_* = 2$.*
- *for all $\varepsilon > 0$, $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ and $y \in \mathcal{Y}_\Delta(O,u)$ such that $H(O,u,y,y+\varepsilon) \in \mathbb{B}_n$ there exists a sequence of rational numbers $\delta_k \to \varepsilon$, and a sequence of $u_k \to u$ of $\mathbb{Q}_{d-1} = \{v \in \mathbb{Q}^2 : \|v\|_2 = 1\}$ and a sequence of rational numbers $y_k \to y$ such that for all $x \in \mathbb{R}^d$ we have*

$$\lim_{k\to\infty} g_k(x) = \lim_{k\to\infty} \mathbb{1}_{H(O,u_k,y_k,y_k+\delta_k)}(x) = g(x) = \mathbb{1}_{H(O,u,y,y+\delta)}(x)$$

Theorem 5.2 (Moments inequality [13], [14])**.** *Let $\mathcal{G}$ satisfy **(F.i)** and **(F.ii)** with envelope G and be such that for some positive constants $\beta, v, c > 1$ and $\sigma \leq 1/(8c)$ the following conditions holds*

$$\mathbb{E}(G^2(X)) \leq \beta^2; \quad N_G(\varepsilon, \mathcal{G}) \leq c\varepsilon^{-v}, 0 < \varepsilon < 1; \quad \sup_{g\in\mathcal{G}} \mathbb{E}(g^2(X)) \leq \sigma^2;$$

and

$$\sup_{g\in\mathcal{G}} \sup_{x\in\mathcal{X}} |g(x)| \leq \frac{\sqrt{n\sigma^2/\ln(\beta \vee 1/\sigma)}}{2\sqrt{v+1}}.$$

Then we have for a universal constant A_2 not depending on β,

$$\mathbb{E}\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \tau_i g(X_i)\right\|_{\mathcal{G}}\right) \leq A_2\sqrt{v\sigma^2 \ln(\beta \vee 1/\sigma)}.$$

Remark 5.2. *Let $n \geq 3$, $g \in \mathcal{F}_n$ and $G = 1$ the envelope function of $\mathcal{F}_n$, we have*

$$\mathbb{E}(G(X)^2) = 1 \leq \beta = 2$$

*Under **(A₁)** for all $B \in \mathcal{B}_n$*

$$E(\mathbb{1}^2_{X\in B}) = P(B) \leq M\varepsilon_n =: \theta_n^2$$

since $\mathcal{B}_n$ is a VC class of dimension $2d+1$ there exists $c > 1$

$$N(\varepsilon, \mathcal{F}_n) \leq c\varepsilon^{-v}, 0 < \varepsilon < 1$$

with $v = 2((2d+1)-1) = 4d$. *Finally, there exists* $n_0 > 0$ *such that for all* $n > n_0$ *we have*

$$\frac{1}{\theta_n} = \frac{1}{\sqrt{M\varepsilon_n}} = \left(\frac{1}{MC}\sqrt{\frac{n}{\log\log n}}\right)^{1/2} = \frac{1}{\sqrt{MC}} \cdot \frac{n^{1/4}}{(\log\log n)^{1/4}} > 2.$$

consequently, $\ln(\beta \vee 1/\theta_n) = \log(2 \vee 1/\theta_n) = \log(1/\theta_n)$ *and*

$$n\theta_n^2 = CM\sqrt{n\log\log n}$$

so

$$\frac{n\theta_n^2}{\log(\beta \vee \frac{1}{\theta_n})} = \frac{CM\sqrt{n\log\log n}}{\log\left(\frac{1}{\sqrt{MC}} \cdot \frac{n^{1/4}}{(\log\log n)^{1/4}}\right)}$$

thus

$$\lim_{n\to\infty} \frac{n\theta_n^2}{\log(\beta \vee \frac{1}{\theta_n})} = +\infty.$$

Since $\sup_{g\in\mathcal{F}_n}\sup_{x\in\mathcal{X}}|g(x)| = 1$ *there exists* $n_1 > n_0 > 0$ *such that for* $n \geq n_1$ *we have*

$$\sup_{g\in\mathcal{F}_n}\sup_{x\in\mathcal{X}}|g(x)| \leq \frac{\sqrt{n\theta_n^2/\log(\beta \vee 1/\theta_n)}}{2\sqrt{v+1}}.$$

Theorem 5.3 (Berthet and Mason 2006 [4])**.** *Let* $\mathcal{G}$ *be a VC class of dimension* $VC(\mathcal{G})$ *satisfying* ***(F.i)*** *and* ***(F.ii)*** *with envelope* $G := S_*/2$. *For all* $\lambda > 1$ *there exists* $\rho(\lambda) > 1$ *such that for all* $n \geq 1$ *we can construct on the same probability space, the vectors* $X_1, \cdots, X_n$ *and a sequence* $(\mathbb{G}_n)$ *of versions of* $\mathbb{G}$ *such that*

$$\mathbb{P}\{\|\alpha_n - \mathbb{G}_n\|_{\mathcal{G}} > \rho(\lambda)n^{-v_1}(\log n)^{v_2}\} \leq n^{-\lambda}$$

with $v_1 = 1/(2+5v_0)$ *and* $v_2 = (4+5v_0)/(4+10v_0)$ *and* $v_0 = 2(VC(\mathcal{G})-1)$ *and where* $\mathbb{G}$ *is P-Brownian Bridge indexed by* $\mathcal{G}$.

References

[1] K. S. Alexander. Probability inequalities for empirical processes and a law of the iterated logarithm. The Annals of Probability, 12(4):1041–1067, 11 1984.

[2] R. R. Bahadur. A note on quantiles in large samples. The Annals of Mathematical Statistics, 37(3):577–580, 06 1966.

[3] P. Berthet and C. El-Nouty. Almost sure asymptotic behaviour of the shorth estimators. Mathematical Methods of Statistics, 14(4):379–403, 2005.

[4] P. Berthet and D. M. Mason. Revisiting two strong approximation results of dudley and philipp. High Dimensional Probability, 51:155–172, 2006.

[5] M. Csörgő and P. Révész. Strong approximations of the quantile process. The Annals of Statistics, 6(4):882–894, 07 1978.

[6] M. Csörgő and P. Révész. Strong approximations in probability and statistics. Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981.

[7] J. A. Cuesta and C. Matrán. The strong law of large numbers for k-means and best possible nets of Banach valued random variables. Probability Theory and Related Fields, 78(4):523–534, 1988.

[8] J. A. Cuesta-Albertos, A. Gordaliza, and C. Matrán. Trimmed k-means: an attempt to robustify quantizers. The Annals of Statistics, 25(2):553–576, 1997.

[9] P. Deheuvels and D. M. Mason. Bahadur-kiefer-type processes. The Annals of Probability, 18(2):669–697, 04 1990.

[10] D. L. Donoho and M. Gasko. Breakdown properties of location estimates based on halfspace depth and projected outlyingness. The Annals of Statistics, 20(4):1803–1827, 12 1992.

[11] J. H. J. Einmahl. A short and elementary proof of the main Bahadur-Kiefer theorem. The Annals of Probability, 24(1):526–531, 1996.

[12] J. H. J. Einmahl and D. M. Mason. Generalized quantile processes. The Annals of Statistics, 20(2):pp. 1062–1078, 1992.

[13] U. Einmahl and D. M. Mason. An empirical process approach to the uniform consistency of kernel-type function estimators. Journal of Theoretical Probability, 13(1):1–37, 2000.

[14] E. Ginè and A. Guillou. On consistency of kernel density estimators for randomly censored data: rates holding uniformly over adaptive intervals. Annales de l'Institut Henri Poincare (B) Probability and Statistics, 37(4):503 – 522, 2001.

[15] M. Hallin, D. Paindaveine, and M. Šiman. Multivariate quantiles and multiple-output regression quantiles: From L_1-optimization to halfspace depth. The Annals of Statistics, 38(2):635–669, 04 2010.

[16] X. He and G. Wang. Convergence of depth contours for multivariate datasets. The Annals of Statistics, 25(2):495–504, 1997.

[17] J. Kiefer. On bahadur's representation of sample quantiles. The Annals of Mathematical Statistics, 38(5):1323–1342, 1967.

[18] J. Kiefer. Deviations between the sample quantile process and the sample df. In Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969), pages 299–319. Cambridge Univ. Press, London, 1970.

[19] V. I. Koltchinskii. M-estimation, convexity and quantiles. The Annals of Statistics, 25(2):435–477, 04 1997.

[20] J. Komlos, P. Major, and G. Tusnady. An approximation of partial sums of independent rv's and the sample df. i. Z. Wahrsch. Verw. Gebiete, 32:111–131, 1975.

[21] L. Kong and I. Mizera. Quantile tomography: using quantiles with multivariate data. Statistica Sinica, 22(4):1589–1610, 2012.

[22] G. Koshevoy and K. Mosler. Zonoid trimming for multivariate distributions. The Annals of Statistics, 25(5):1998–2017, 10 1997.

[23] R. Y. Liu and K. Singh. A quality index based on data depth and multivariate rank tests. Journal of the American Statistical Association, 88(421):252–260, 1993.

[24] Regina Y. Liu. On a notion of data depth based on random simplices. The Annals of Statistics, 18(1):405–414, 03 1990.

[25] D. Nolan. The excess-mass ellipsoid. Journal of Multivariate Analysis, 39(2):348–371, 1991.

[26] D. Nolan. Asymptotics for multivariate trimming. Stochastic Processes and their Applications, 42(1):157–169, 1992.

[27] W. Polonik. Measuring mass concentrations and estimating density contour clusters-an excess mass approach. The Annals of Statistics, 23(3):855–881, 06 1995.

[28] W. Polonik. Minimum volume sets and generalized quantile processes. Stochastic Processes and their Applications, 69(1):1 – 24, 1997.

[29] D. Rainer and M. Karl. Weighted-mean trimming of multivariate data. Journal of Multivariate Analysis, 102(3):405 – 421, 2011.

[30] R. Serfling. Generalized quantile processes based on multivariate depth functions, with applications in nonparametric multivariate analysis. Journal of Multivariate Analysis, 83(1):232 – 247, 2002.

[31] G. R. Shorack and J. A. Wellner. Empirical processes with applications to statistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.

[32] M. Talagrand. Sharper bounds for gaussian and empirical processes. The Annals of Probability, 22(1):28–76, 01 1994.

[33] J. W. Tukey. Mathematics and the picturing of data. In Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2, pages 523–531. Canad. Math. Congress, Montreal, Que., 1975.

[34] Y. Zuo and R. Serfling. General notions of statistical depth function. The Annals of Statistics, 28(2):461–482, 04 2000.

[35] Y. Zuo and R. Serfling. Structural properties and convergence results for contours of sample statistical depth functions. The Annals of Statistics, 28(2):483–499, 04 2000.