Klaus Pötzelberger

The General Quantization Problem for Distributions with Regular Support

Paper

The General Quantization Problem for Distributions with Regular Support

Klaus Pötzelberger

Institut für Statistik
Wirtschaftsuniversität Wien

Forschungsberichte

Bericht 69
1999

http://statmath.wu-wien.ac.at/

The General Quantization Problem for Distributions with Regular Support

Klaus Pötzelberger

Abstract

We study the asymptotic behavior of the quantization error for general information functions and prove results for distributions P with regular support. We characterize the information functions for which the uniform distribution on the set of prototypes converges weakly to P.

AMS 1991 subject classifications. Primary 60D05; secondary 62H30, 68T10.
Key words and phrases. Quantization error, quantization dimension, k-means clustering, principal points, Theorem of Zador.

1 Introduction

The asymptotic behavior of the quantization error of a distribution is closely related to dimensional properties of its support. The Theorem of Zador [12], a landmark in the theory of optimal quantization, has been a first result in this context. It states that, subject to regularity conditions, for a distribution P on $\mathbb{R}^d$,

$$\lim_{n\to\infty} n^{2/d}\Delta_n(\|.\|^2, P) = C_{\mathfrak{L}}(d)\left(\int\left(\frac{dP}{d\mathfrak{L}^d}\right)^{d/(2+d)} d\mathfrak{L}^d\right)^{1+2/d}, \tag{1}$$

where $C_{\mathfrak{L}}(d)$ is the quantization constant and $dP/d\mathfrak{L}^d$ is the Radon-Nikodym derivative of P with respect to d-dimensional Lebesgue measure. $\Delta_n(\|.\|^2, P)$, the quantization error of P, is defined as

$$\Delta_n(\|.\|^2, P) = \inf_{|O|\le n}\int \min_{p\in O}\|x-p\|^2 dP(x). \tag{2}$$

$C_{\mathfrak{L}}(d)$ depends on the dimension d only.

Zador [13] points out that an extension to more general distributions should be possible. Especially, for the uniform distribution on the Cantor set he provides an incorrect proof of the existence of the limit (1), with d the Hausdorff dimension of the Cantor set. Furthermore, he defines a quantization dimension, which is motivated by this example. However, the limit does not exist, as has been pointed out by Graf and Luschgy [5]. Pötzelberger [9] defines the lower and the upper quantization dimensions, $\underline{dim}_Q(P)$ and $\overline{dim}_Q(P)$, as the unique real numbers s_u and $s_o \in [0, \infty]$ resp., satisfying

$$\liminf_{n\to\infty} n^{2/s}\Delta_n(\|.\|^2, P) = \begin{cases} \infty & \text{for } s < s_u, \\ 0 & \text{for } s > s_u, \end{cases}$$

$$\limsup_{n\to\infty} n^{2/s}\Delta_n(\|.\|^2, P) = \begin{cases} \infty & \text{for } s < s_o, \\ 0 & \text{for } s > s_o. \end{cases}$$

If $\underline{dim}_Q(P) = \overline{dim}_Q(P)$ the common value is called the quantization dimension of P, $dim_Q(P)$. Falconer [2] discusses various concepts of the dimension of a distribution, such as the Hausdorff and the packing dimension. See section 6 for relevant notions. Pötzelberger [9] investigates the quantization dimensions and shows how these dimensions fit into geometric measure theory. More precisely, it is shown that for distributions P on $[0,1]^d$, $\underline{dim}_Q(P) \in [dim^*_{\mathcal{H}}(P), \underline{dim}_B(P)]$ and $\overline{dim}_Q(P) \in [dim^*_{\mathcal{P}}(P), \overline{dim}_B(P)]$. Let $\mathcal{H}^s$ denote the s-dimensional Hausdorff measure. If a Borel set $E \subseteq [0,1]^d$ exists, such that $P(E) = 1$, $0 < \mathcal{H}^s(E) < \infty$, and $P << \mathcal{H}^s_{|E}$, then

$$\liminf_{n\to\infty} n^{2/s}\Delta_n(\|.\|^2, P) > 0.$$

Furthermore, if a Borel set $E \subseteq [0,1]^d$ of exact upper box-counting dimension s exists, such that $P(E) = 1$, then

$$\limsup_{n\to\infty} n^{2/s}\Delta_n(\|.\|^2, P) < \infty.$$

Distributions may be classified as follows. In general, the upper and the lower quantization dimensions differ. Let the quantization dimension of P exists and denote it by s. In this case typically $\liminf_{n\to\infty} n^{2/s}\Delta_n(\|.\|^2, P) = 0$ or $\limsup_{n\to\infty} n^{2/s}\Delta_n(\|.\|^2, P) = \infty$. The third class of distributions consists of distributions with existing quantization dimension s such that the sequence $n^{2/s}\Delta_n(\|.\|^2, P)$ is bounded and bounded away from 0. This class contains the invariant distributions on the attractor of iterated function systems, such as the Cantor distribution. Here, the quantization dimension is the similarity dimension (which coincides with the Hausdorff and the upper box-counting dimension). Finally, a class of distributions exists where $n^{2/s}\Delta_n(\|.\|^2, P)$ converges to a finite limit. According to the Theorem of Zador this class contains the distributions with nonvanishing absolutely continuous component.

In this paper we prove extensions of the Theorem of Zador. There are at least three cases where these extensions are necessary and the results are relevant.

1. The Theorem of Zador holds for distributions with support E a regular s-dimensional set. s is then necessarily integral. (1) holds, when d is replaced by s, $\mathcal{L}^d$ by the s-dimensional Hausdorff measure $\mathcal{H}^s$ and $C_{\mathcal{L}}(d)$ by a suitable constant $C_{\mathcal{H}}(s)$.

2. Pötzelberger and Strasser [10] provide a setting for classification and quantization. It includes approaches such as vector quantization, k-means clustering, principal points in the sense of Flury [4] or unsupervised learning in the theory of artificial neural nets. For a given distribution P on $\mathbb{R}^d$, a partition $\mathcal{B} = (B_1, \ldots, B_n)$ containing maximal information is sought. The information of a partition is measured by

$$I^f(\mathcal{B}, P) = \int f(\mathbb{E}(X \mid \mathcal{B}))dP, \tag{3}$$

where $f : \mathbb{R}^d \to \mathbb{R}$ is an integrable convex function and

$$\mathbb{E}(X \mid \mathcal{B}) = \sum_{i=1}^{n} \frac{1}{P(B_i)} \int_{B_i} XdP \quad I_{B_i} \tag{4}$$

is the conditional expectation of the identity $X(x) = x$, given the σ-field generated by the partition. Optimal partitions $\mathcal{B}$ correspond to optimal quantizers $O = \{p_1, \ldots, p_n\}$ and vice versa. In the case of $f(x) = \|x\|^2$, $B_i = \{x \mid argmin_k\|x - p_k\|\}$ and $p_i = \int_{B_i} XdP/P(B_i)$. The means p_i are called prototypes in case $\mathcal{B}$ is optimal.

Maximizing the information, the so-called primal problem, is equivalent to the dual problem: Find $a_1, \ldots, a_n \in \mathbb{R}^d$, which maximize

$$F(a_1, \ldots, a_n) = \int \max_{j \leq n}(\langle a_j, X\rangle - f^c(a_j))dP \tag{5}$$

where f^c denotes the conjugate convex function of f,

$$f^c(a) = \sup_x(\langle a, x\rangle - f(x)). \tag{6}$$

Then $\mathcal{B} = (B_1, \ldots, B_n)$, defined by

$$B_j = \{x \mid j = argmax_k(\langle a_k, x\rangle - f^c(a_k))\}, \tag{7}$$

is optimal for the primal problem. Examples of information functions f are

(a) $f(x) = \|x\|^2$. Here $f^c = f/4$ and the dual problem is the classical quantization problem, i.e. the problem of finding the variance-minimizing partition.

(b) $f(x) = \|x\|$. Here

$$f^c(x) = \begin{cases} 0 & \text{if } \|x\| \le 1, \\ \infty & \text{if } \|x\| > 1. \end{cases}$$

The dual problem, i.e. finding $a_1, \ldots, a_n$ with $\|a_j\| \le 1$, such that $\int \max_j \langle a_j, X\rangle dP$ is maximal, is called Kohonen's problem (cf. [6]). *(c)* Robust versions of information functions are

$$f(x) = \begin{cases} \|x\|^2/2 & \text{if } \|x\| \le c, \\ c\|x\| - c^2/2 & \text{if } \|x\| > c, \end{cases}$$

or $f(x) = 2 \log \cosh \|x\|/2$.

For twice differentiable information functions both the dimension of the support and the rank of f'' determine the asymptotic features of the quantization error. For instance, information functions with positive definite Hessian f'' lead to the same rate of convergence as the classical $f = \|.\|^2$. In Kohonen's problem the rank of f'' is $d-1$. This quantization problem is then, for instance when $P << \mathcal{L}^d$, equivalent to a $d-1$-dimensional problem.

3. Consider the approximation of P by a distribution with finite support. Let $\mathcal{B} = (B_1, \ldots, B_n)$ be a solution of the primal problem and let $O_n = \{p_1, \ldots, p_n\}$ be the set of conditional means. These prototypes represent for $f = \|.\|^2$ the distribution P in the following sense. Denote by $\mathcal{M}(P_1, P_2)$ the set of distributions on $\mathbb{R}^d \times \mathbb{R}^d$ with marginal distributions P_1 and P_2 and let $\mathcal{D}_n$ be the set of discrete distributions on $\mathbb{R}^d$ with a support of cardinality at most n. The minimal L_2-metrik ℓ_2 is defined by

$$\ell_2^2(P_1, P_2) = \inf\{\int \|x_1 - x_2\|^2 dM(x_1, x_2) \mid M \in \mathcal{M}(P_1, P_2)\}.$$

Then for $Q_n(\{p_i\}) := P(B_i)$,

$$\ell_2(P, Q_n) = \min\{\ell_2(P, Q) \mid Q \in \mathcal{D}_n\}.$$

In applications the prototypes p_i themselves, independent of their weights $P(B_i)$, are interpreted as representing the distribution P. Let $\hat{Q}_n$ be the uniform distribution on $O_n = \{p_1, \ldots, p_n\}$. We will show that in general $\hat{Q}_n$ does not converge to P. However, the squared norm may be replaced by a suitably chosen information function f, such that $\hat{Q}_n$ converges weakly to P (if $P << \mathcal{L}^d$). In that case the prototypes might be considered as generalized multivariate quantiles.

Notation. Let P denote a Borel probability measure on $\mathbb{R}^d$, $csupp(P)$ a Borel set with $P(csupp(P)) = 1$ and f a continuous convex function defined on $csupp(P)$. Furthermore, let

$X : x \mapsto x$ be the identity on $\mathbb{R}^d$. We assume that f and $\|X\|$ are P-integrable. A support of P is any Borel set E with $P(E) = 1$. We do not assume that E is closed.

For a locally finite Borel measure G and a Borel set E, $G_{|E}$ and $G_{\restriction E}$ denote the restriction to E and the conditional distribution given E, i.e. $G_{|E}(A) = G(A \cap E)$ and, if $0 < G(E) < \infty$, $G_{\restriction E} = G(A \cap E)/G(E)$. $\mathcal{L}^s$ and $\mathcal{H}^s$ are the s-dimensional Lebesgue measure and the s-dimensional Hausdorff measure. Let a partition $\mathcal{B} = (B_1, \ldots, B_n)$ into Borel sets B_i with $P(B_i) > 0$ be given. Then $O = O(\mathcal{B}) = \{p_1, \ldots, p_n\}$ with $p_i = \int_{B_i} X dP/P(B_i)$. Let $O = \{p_1, \ldots, p_n\}$ be a set of n vectors, not necessarily prototypes and let a_i be subdifferentials of f at p_i, which maximize (5). We define $I^f(O, P)$ by (5). The quantization error is

$$\Delta_n(f, O, P) = \int f dP - I^f(O, P). \tag{8}$$

It is always nonnegative, as f is convex. The quantization error of the quantization problem (f, P) is

$$\Delta_n(f, P) = \inf_{|O| \leq n} \Delta_n(f, O, P). \tag{9}$$

For the information function $f = \|.\|^2$, (2) and (9) are identical definitions of the quantization error. For measures of the form $P_{|E}$ with $0 < P(E) < \infty$, $\Delta_n(f, P_{|E})$ is an abbreviatation of $P(E)\Delta_n(f, P_{\restriction E})$.

Note that the integrability of f and $\|X\|$ implies the existence of optimal partitions and of sets of prototypes (cf. Pollard [8] or Pötzelberger and Strasser [10]). For differentiable information functions f, $\Delta_n(f, O, P)$ may be written as

$$\Delta_n(f, O, P) = \int \min_{p \in O} \{f(x) - f(p) - \langle x - p, f'(p) \rangle\} dP(x).$$

A winner of a vector x is a prototype $p \in O$, which minimizes $f(x) - f(p) - \langle x - p, f'(p) \rangle$. $B(x, r)$ denotes the open ball with center x and radius r.

The paper is organized as follows. Section 2 provides bounds and rates for the quantization error of general information functions. The results are based on an analysis of random quantizers. Section 3 provides results for regular quantization problems, i.e. when f is twice differentiable and f'' is of full rank. Nonregular problems, where f'' is not of full rank, are considered in section 4. Section 5 gives results on the convergence of the uniform distribution on the sets of prototypes. Finally, section 6 is a compilation of auxillary results, including definitions from geometric measure theory.

2 Asymptotics of the quantization error: rates and upper bounds

The first result obtained is an upper bound of the quantization error. The proof is based on an analysis of random quantizers, which is of interest of its own. A random quantizer $\tilde{O}_n$ with distribution G is a G-distributed sample, i.e. $\tilde{O}_n = \{p_1, \dots, p_n\}$ with (p_i) independent and G-distributed. $\mathbb{E}^G(\Delta_n(f, \tilde{O}_n, P))$ is by definition the quantization error of the random quantization rule G. It is an upper bound of $\Delta_n(f, P)$.

We define for a distibution G, $\alpha \in \mathbb{R}_+$, $r \in]0, \infty]$,

$$\eta_{G,\alpha,r}(x) = \inf\{G(B(x, r') \cap B(0, r))/(2r')^\alpha \quad | \; 0 < r' < 2r\}. \tag{10}$$

Theorem 1 is nontrivial only for distributions for which G, α and c exist, such that $\eta_{G,\alpha,\|x\|c}(x) > 0$ a.e. This assumption is slightly more restrictive than the existence of a strictly positive lower density. Note that for $\alpha > dim^*_{\mathcal{P}}(P)$, $\eta_{P,\alpha,r} > 0$ P-a.s.

For f differentiable on $csupp(P)$, f' denotes the gradient and $\lambda : \mathbb{R}^+ \to [0, \infty]$,

$$\lambda(t) = \sup\{\|f'(x) - f'(y)\|/\|x - y\| \quad | \; x, y \in B(0, t) \cap csupp(P), x \neq y\}, \tag{11}$$

the Lipschitz constant of f' on $B(0, t) \cap csupp(P)$.

Let G be a distribution and $y_0, \dots, y_n \in \mathbb{R}^d$. We define

$$\Delta_1(x) = f(x) - f(y_0) - \langle x - y_0, f'(y_0)\rangle,$$

$$\Delta_{n+1}(x, \{y_1, \dots, y_n\}) = \min\{f(x) - f(y_i) - \langle x - y_i, f'(y_i)\rangle \mid 0 \leq i \leq n\}.$$

Let $(y_i)_{i=1}^n$ be independent G-distributed. The expectation of $\Delta_{n+1}(x, \{y_1, \dots, y_n\})$ is denoted by $\Delta^G_{n+1}(x)$, i.e.

$$\Delta^G_{n+1}(x) = \mathbb{E}^G(\Delta_{n+1}(x, \{y_1, \dots, y_n\})).$$

THEOREM 1. *Let $f \in C^{(1)}(csupp(P))$, G a distribution on $\mathbb{R}^d$, such that $G(B(0, r)) > 0$ for all $r > 0$, for which $P(B(0, r)) > 0$ holds.*

1. *Let $x_0 \in \mathbb{R}^d \backslash \{0\}$, $c \in]1, \infty[$, $\alpha \geq dim^*_{\mathcal{P}}(P)$ and $x \in B(0, \|x_0\|)$. Then*

$$n^{2/\alpha}\Delta^G_{n+1}(x) \leq \frac{\Gamma(2/\alpha + 1)\lambda(\|x_0\|c)}{2\eta_{G,\alpha,\|x_0\|c}(x)^{2/\alpha}} + n^{2/\alpha}G(B(0, \|x_0\|c)^c)^n\Delta_1(x). \tag{12}$$

If $x \notin B(0, \|x_0\|)$, then

$$n^{2/\alpha}\Delta^G_{n+1}(x) \leq \frac{\Gamma(2/\alpha + 1)\lambda(\|x\|c)}{2\eta_{G,\alpha,\|x\|c}(x)^{2/\alpha}} + n^{2/\alpha}G(B(0, \|x_0\|c)^c)^n\Delta_1(x). \tag{13}$$

2. Let $c \in]1,\infty[$ and $\alpha \geq dim^*_{\mathcal{P}}(P)$. Then

$$\limsup_{n\to\infty} n^{2/\alpha}\Delta_n(f,P) \leq \frac{\Gamma(2/\alpha+1)}{2}\int \frac{\lambda(\|x\|c)}{\eta_{G,\alpha,\|x\|c}(x)^{2/\alpha}}dP(x). \tag{14}$$

As a corollary we obtain the "usual assumptions" (16) for $\limsup_{n\to\infty} n^{2/s}\Delta_n(\|.\|^2,P) < \infty$.

COROLLARY 1. *1. Let f'' be regular, $E \subseteq \mathbb{R}^d$ with $P(E) = 1$ and $dim_{\mathcal{P}}(E) = s$. If a measurable set $E' \supseteq E$, a σ-finite measure μ, with*

$$\inf_{x\in E}\inf_{r>0}\frac{\mu(E'\cap B(x,r))}{(2r)^s} > 0,$$

$c \in]1,\infty[$ and a μ-integrable, in $\|.\|$ nonincreasing function g with

$$\int \frac{\lambda(\|x\|c)}{g(cx)^{2/s}}dP(x) < \infty, \tag{15}$$

exist, then $\limsup_{n\to\infty} n^{2/s}\Delta_n(f,P) < \infty$.
2. In case $s = d$, $\limsup_{n\to\infty} n^{2/s}\Delta_n(\|.\|^2,P) < \infty$, if a $\delta' > 0$ exists, such that

$$\int \|x\|^{2+\delta'}dP(x) < \infty. \tag{16}$$

3. Let $f \in C^{(2)}$ with f'' regular, $P(f'' > 0) = 1$, $dim_{\mathcal{H}}(P) = dim_{\mathcal{P}}(P) = s$. If for all $\alpha > s$ a distribution G and $c \in]1,\infty[$ exist, such that

$$\int \frac{\lambda(\|x\|c)}{\eta_{G,\alpha,\|x\|c}(x)^{2/\alpha}}dP(x) < \infty, \tag{17}$$

then $dim_Q(P) = s$.

PROOF. 1. Choose $G << \mu$ with $dG/d\mu \propto g$. A constant $c_0 > 0$ exists, such that $\eta_{G,s,\|x\|c}(x) \geq c_0 g(cx)$.
2. Let $s = d$ and (16) hold. Then $E' = \mathbb{R}^d$, $\mu = \mathcal{L}^d$ and $g(x) \propto I_{\{\|x\|\leq 1\}} + \|x\|^{-(d+\delta)}I_{\{\|x\|>1\}}$ with $\delta = \delta' d$ satisfy (15). □

PROOF of THEOREM 1. Let $x_0 \neq 0$, $r_0 = \|x_0\|$, $c \in]1,\infty[$, $y_0 \in \mathbb{R}^d$ with $\Delta_1(\{y_0\},f,P) < \infty$. $y_1,\dots,y_n$ denotes a G-distributed sample of size n. Furthermore, let $N = \#\{i \geq 1 | y_i \in B(0,r_0c)\}$. N is binomially distributed with parameter n and $G(B(0,r_0c))$.
Let $x \in B(0,r_0)$. $\tilde{y} = \tilde{y}(x)$ denotes a winner for x in $\{y_1,\dots,y_n\} \cap B(0,r_0c)$, if $N > 0$. If $N = 0$, let $\tilde{y} = y_0$. Thus

$$\Delta_{n+1}(x,\{y_1,\dots,y_n\}) \leq \left(f(x) - f(\tilde{y}) - \langle x-\tilde{y}, f'(\tilde{y})\rangle\right) I_{\{N>0\}} + \Delta_1(x)I_{\{N=0\}}$$

$$\leq \frac{\lambda(r_0 c)}{2}\|x-\tilde{y}\|^2 I_{\{N>0\}} + \Delta_1(x) I_{\{N=0\}}.$$

For $N > 0$ we denote $\{y_1, \dots y_n\} \cap B(0, r_0 c)$ by $\{y_1^*, \dots, y_N^*\}$. $G_{\upharpoonright B(0,r_0 c)}$ is the conditional distribution given $B(0, r_0 c)$. Let $\alpha \geq s$ and $t > 0$. Then, given $N = n^*$ with $n^* > 0$,

$$G_{\upharpoonright B(0,r_0 c)}(\lambda(r_0 c)\|x-\tilde{y}\|^2/2 > t)$$

$$= G_{\upharpoonright B(0,r_0 c)}(\forall 1 \leq i \leq n^* : y_i^* \notin B(x, (2t/\lambda(r_0 c))^{1/2}))$$

$$= \left(1 - G_{\upharpoonright B(0,r_0 c)}(B(x, (2t/\lambda(r_0 c))^{1/2})^c)\right)^{n^*}$$

$$\leq \left(1 - \frac{\eta_{G,\alpha,r_0 c}(x)}{G(B(0,r_0 c))}(2t/\lambda(r_0 c))^{\alpha/2}\right)^{n^*}.$$

Therefore,

$$\Delta_{n+1}^G(x)$$

$$\leq \int_0^\infty \sum_{n^*=1}^{n} \binom{n}{n^*} G(B(0,r_0 c))^{n^*} G(B(0,r_0 c)^c)^{n-n^*} G_{\upharpoonright B(0,r_0 c)}(\frac{\lambda(r_0 c)}{2}\|x-\tilde{y}\|^2 > t)dt$$

$$+ G(B(0,r_0 c)^c)^n \Delta_1(x)$$

$$\leq \int_0^\infty \sum_{n^*=0}^{n} \binom{n}{n^*} G(B(0,r_0 c))^{n^*} G(B(0,r_0 c)^c)^{n-n^*} \left(1 - \frac{\eta_{G,\alpha,r_0 c}(x)}{G(B(0,r_0 c))}(2t/\lambda(r_0 c))^{\alpha/2}\right)^{n^*} dt$$

$$+ G(B(0,r_0 c)^c)^n \Delta_1(x)$$

$$= \int_0^\infty \left(1 - \eta_{G,\alpha,r_0 c}(x)(2t/\lambda(r_0 c))^{\alpha/2}\right)^n dt + G(B(0,r_0 c)^c)^n \Delta_1(x)$$

$$\leq \int_0^\infty e^{-n\eta_{G,\alpha,r_0 c}(x)(2t/\lambda(r_0 c))^{\alpha/2}} dt + G(B(0,r_0 c)^c)^n \Delta_1(x)$$

$$= \frac{1}{n^{2/\alpha}} \frac{\Gamma(2/\alpha+1)\lambda(r_0 c)}{2\eta_{G,\alpha,r_0 c}(x)^{2/\alpha}} + G(B(0,r_0 c)^c)^n \Delta_1(x).$$

For $x \notin B(0, r_0)$,

$$\Delta_{n+1}^G(x) \leq \frac{1}{n^{2/\alpha}} \frac{\Gamma(2/\alpha+1)\lambda(\|x\|c)}{2\eta_{G,\alpha,\|x\|c}(x)^{2/\alpha}} + G(B(0,\|x\|c)^c)^n \Delta_1(x)$$

may be proved along the same lines. $G(B(0,\|x\|c)^c) \leq G(B(0,r_0 c)^c)$ implies (13).

To prove *2.*, let $\|x_0\| = r_0 \neq 0$. Since $\lim_{n\to\infty} n^{2/\alpha} G(B(0,r_0 c)^c)^n = 0$,

$$\limsup_{n\to\infty} n^{2/\alpha}\Delta_n(f,P) \leq \limsup_{n\to\infty} n^{2/\alpha} \int \Delta_n^G(x) dP(x)$$

$$\leq \int_{B(0,r_0)} \frac{\Gamma(2/\alpha+1)\lambda(r_0 c)}{2\eta_{G,\alpha,r_0 c}(x)^{2/\alpha}} dP(x) + \int_{B(0,r_0)^c} \frac{\Gamma(2/\alpha+1)\lambda(\|x\|c)}{2\eta_{G,\alpha,\|x\|c}(x)^{2/\alpha}} dP(x)$$

$$= \int \frac{\Gamma(2/\alpha+1)\lambda(\|x\|c)}{2\eta_{G,\alpha,\|x\|c}(x)^{2/\alpha}} dP(x)$$

$$+ \int_{B(0,r_0)} \left(\frac{\Gamma(2/\alpha+1)\lambda(r_0 c)}{2\eta_{G,\alpha,r_0 c}(x)^{2/\alpha}} - \frac{\Gamma(2/\alpha+1)\lambda(\|x\|c)}{2\eta_{G,\alpha,\|x\|c}(x)^{2/\alpha}} \right) dP(x).$$

For $x_0 \to 0$ the second summand converges to 0. This proves (14). □

The assumption $\lambda(t) < \infty$ to achieve $\Delta_n(f, P) = O(n^{-2/s})$ with $s = dim^*_{\mathcal{P}}(P)$ can in general cases not be replaced by weaker assumptions. For instance, it is not sufficient that f is twice P-a.e. differentiable, as the following example shows. In the example is P the uniform distribution on $[0, 1[$, $d = 1 = dim^*_{\mathcal{P}}(P)$ and f is piecewise linear, therefore P-a.s. infinitely often differentiable. Furthermore, for any positive null sequence (β_n) of the form $\beta_n = \sum_{j=n+1}^{\infty} \alpha_j$ with (α_j) positive, summable and decreasing, such a function f exists which satisfies

$$\liminf_{n\to\infty} \Delta_n(f, P)/\beta_n > 0. \tag{18}$$

EXAMPLE 1. P denotes the uniform distribution on $[0, 1[$. Let $q \in]0, 1[$ and $x_i = 1 - q^i$. We choose f as continuous on $[0, 1[$ and linear on $[x_i, x_{i+1}[$, i.e. constants a_i and b_i exist, such that for $x \in [x_i, x_{i+1}[$, $f(x) = a_i(x - x_i) + b_i$. The continuity of f implies $b_{i+1} = a_i(x_{i+1} - x_i) + b_i$. We define $a_0 = 0$, $b_0 = 0$, and for $i > 0$, $a_i - a_{i-1} = 2\alpha_i/q^{2i}$, where $(\alpha_i)_{i=1}^{\infty}$ is a fixed summable and decreasing sequence of positive reals. f is convex, since (a_i) is increasing. We have $\int_0^1 f(x)dx = \sum_{i=1}^{\infty} \alpha_i < \infty$. Let $\ell_i : [0, 1[\to \mathbb{R}$, $\ell_i(x) = a_i(x - x_i) + b_i$ be the tangent on the right to f in x_i. For $r < s \leq \infty$ define ($x_\infty := 1$)

$$\underline{\delta}(r, s) = \int_{[x_r, x_s[} (f(x) - \ell_r(x))dx,$$

$$\bar{\delta}(r, s) = \int_{[x_r, x_s[} (f(x) - \ell_s(x))dx.$$

Then $\underline{\delta}(r, \infty) = \sum_{j=r+1}^{\infty} \alpha_j$ and for $s < \infty$

$$\underline{\delta}(r, s) = \sum_{j=r+1}^{s-1} \int_{[x_j, x_s[} (a_j - a_{j-1})(x - x_j)dx$$

$$= \sum_{j=r+1}^{s-1} (a_j - a_{j-1})(x_s - x_j)^2/2 = \sum_{j=r+1}^{s-1} (a_j - a_{j-1})(q^s - q^j)^2/2$$

$$= \sum_{j=r+1}^{s-1} \alpha_j q^{-2j}(q^j - q^s)^2 \geq \sum_{j=r+1}^{s-1} \alpha_j (1-q)^2, \tag{19}$$

$$\bar{\delta}(r,s) = \sum_{j=r+1}^{s-1} \int_{[x_r, x_j[} (a_{j-1} - a_j)(x - x_j)dx$$

$$= \sum_{j=r+1}^{s-1} (a_j - a_{j-1})(x_r - x_j)^2/2 = \sum_{j=r+1}^{s-1} \alpha_j q^{-2j}(q^j - q^r)^2$$

$$= \sum_{j=r+1}^{s-1} \alpha_j (1 - q^{r-j})^2 \geq \sum_{j=r+1}^{s-1} \alpha_j (1 - q^{-1})^2$$

$$\geq \sum_{j=r+1}^{s-1} \alpha_j (1-q)^2. \tag{20}$$

For a given partition $(B_i)_{i=1}^n$, $B_i = [h_i, h_{i+1}[$, $(h_1 = 0 < h_1 < ... < h_n < h_{n+1} = 1)$ let (p_i) denote the means of the intervals B_i in increasing order. p_i lies in an interval $[x_{k(i)}, x_{k(i)+1}[$ and h_i in $[x_{m(i)}, x_{m(i)+1}[$. Thus $m(i) \leq k(i) < m(i+1)$. Let $J_n = \{k(i) | i = 1, ..., n\} \cup \{m(i) | i = 2, ..., n\}$. The size of J_n is at most $2n-1$. If $p_i \neq x_{k(i)}$, then the contribution of the interval B_i to $\Delta_n(f, P)$ is

$$\int_{[h_i, h_{i+1}[} (f(x) - \ell_{k(i)}(x))dx,$$

which for $i = 2, ..., n-1$ is at least

$$\int_{[h_i, x_{m(i)+1}[} (f(x) - \ell_{k(i)}(x))dx + \int_{[x_{k(i)}, x_{k(i)+1}[} (f(x) - \ell_{k(i)}(x))dx$$

$$+ \int_{[x_{m(i+1)}, h_{i+1}[} (f(x) - \ell_{k(i)}(x))dx + \bar{\delta}(m(i)+1, k(i)) + \underline{\delta}(k(i)+1, m(i+1))$$

$$\geq \bar{\delta}(m(i)+1, k(i)) + \underline{\delta}(k(i)+1, m(i+1))$$

$$\geq \sum_{j=m(i)+2}^{k(i)-1} \alpha_j (1-q)^2 + \sum_{j=k(i)+2}^{m(i+1)-1} \alpha_j (1-q)^2.$$

In case $i = n$, the contribution of $[x_{m(i+1]}, h_{i+1}[$ becomes void and $\underline{\delta}(k(i), m(i+1))$ has to be replaced by $\underline{\delta}(k(i), \infty)$. For $i = 1$ the contribution of $[h_i, x_{m(i)+1}[$ has to be omitted and $\bar{\delta}(m(i)+1, k(i))$ has to be replaced by $\bar{\delta}(1, k(i))$. Alltogether,

$$\Delta_n(f, P) \geq \sum_{j \notin J_n \cup (J_n+1)} \alpha_j (1-q)^2 \geq (1-q)^2 \sum_{j=4n}^{\infty} \alpha_j.$$

3 Regular quantization problems with regular support

The subsequent generalization of the Theorem of Zador is formulated in terms of the Hausdorff measure. Restricted to Borel sets, d-dimensional Hausdorff measure is Lebesgue measure up to a normalizing constant. To cope for this normalizing constant we have to adapt the definition of the quantization constant. Let $C_{\mathfrak{L}}(s)$ denote the quantization constant for dimension s, $s \in \mathbb{N}$. We define

$$C_{\mathcal{H}}(s) = \frac{\pi}{8\Gamma(1+s/2)^{2/s}} C_{\mathfrak{L}}(d).$$

ASSUMPTION B. *$f \in C^{(2)}(csupp(P))$ and for all $\epsilon > 0$ a finite number of convex and compact sets E_i exist, such that $\overset{\circ}{E_i} \neq \emptyset$, $P((\cup E_i)^c) < \epsilon$ and f'' is continuous on E_i.*

THEOREM 2. *Let $s \in \mathbb{N}$, E a regular Borel set of dimension s, with $P(E) = 1$ and $P(\{x \mid$ a tangent to E at x exists $\}) = 1$. Let $p = dP/d\mathcal{H}^s_{|E}$. If f satisfies assumption B, if f'' is positive definite on a neighborhood of E and if a distribution G and a $c \in]1,\infty[$ exist with $\int \lambda(\|x\|c)\eta_{s,G,\|x\|c}(x)^{-2/s}dP(x) < \infty$, then*

$$\lim_{n\to\infty} n^{2/s}\Delta_n(f,P) = C_{\mathcal{H}}(s)\left(\int J_{f',E}(x)^{1/(2+s)}p(x)^{s/(2+s)}d\mathcal{H}^s_{|E}(x)\right)^{1+2/s}. \tag{21}$$

PROOF. We begin by establishing for a special case inequalities, which will be useful in proving estimates of the quantization error. In this special case, let the support of P be flat enough, in the following sense. Let E with $P(E) = 1$ be bounded and regular of dimension s, and let f be twice continuously differentiable on $\overline{co}(E)$, the compact and convex hull of E. Let ϵ be positive, $A \in \mathbb{R}^{d\times d}$ positive definite with

$$\frac{1}{1+\epsilon}x^\top Ax \leq \frac{f''(x)}{2} \leq (1+\epsilon)x^\top Ax \tag{22}$$

on $\overline{co}(E)$. P_{ac} denotes the absolutely continuous part of P with respect to $\mathcal{H}^s_{|E}$. Assume that tangents to E vary so little such that for P_{ac}-a.e. $x \in E$

$$\frac{1}{2^s(1+\epsilon)^{s+1}}J_{f',E}(x) \leq \left(J_{A^{1/2},E}(x)\right)^2 \leq \frac{(1+\epsilon)^{s+1}}{2^s}J_{f',E}(x) \tag{23}$$

holds. ϕ denotes a Lipschitz function from $A^{1/2}(E)$ onto a regular set $F \subseteq \mathbb{R}^s$ with

$$\frac{1}{1+\epsilon}\|x-y\| \leq \|\phi(x)-\phi(y)\| \leq (1+\epsilon)\|x-y\|. \tag{24}$$

Furthermore, let $\tilde{P}$ denote the image of P under $\psi = \phi \circ A^{1/2}$. Let O_n be a set of prototypes for (f, P). Then

$$\Delta_n(\|.\|^2, \tilde{P}) \leq \Delta_n(\|.\|^2, \tilde{P}, \psi(O_n))$$
$$= \int \min_{1\leq i\leq n} \|\tilde{x} - \psi(p_i)\|^2 d\tilde{P}(\tilde{x}) = \int \min_{1\leq i\leq n} \|\psi(x) - \psi(p_i)\|^2 dP(x)$$
$$\leq (1+\epsilon)^2 \int \min_{1\leq i\leq n} (x - p_i)^\top A(x - p_i) dP(x) \leq (1+\epsilon)^3 \Delta_n(f, P). \tag{25}$$

$J_{\psi,E}(x) = J_{A^{1/2},E}(x) J_{\phi,A^{1/2}(E)}(A^{1/2}(x))$, Lemma 7 and (23) imply

$$J_{\psi,E}(x) \geq (1/2)^{s/2} J_{f',E}(x)^{1/2} (1+\epsilon)^{-(1+3s)/2}.$$

The Theorem of Zador, the definition of the constant $C_{\mathcal{H}}(s)$ and Lemma 6 give

$$\limsup_{n\to\infty} n^{2/s} \Delta_n(f, P) \geq (1+\epsilon)^{-3} \limsup_{n\to\infty} n^{2/s} \Delta_n(\|.\|^2, \tilde{P})$$
$$= (1+\epsilon)^{-3} C_{\mathcal{L}}(s) \left(\int_F \left(\frac{d\tilde{P}}{d\mathcal{L}^s}(x) \right)^{s/(2+s)} d\mathcal{L}^s(x) \right)^{1+2/s}$$
$$= (1+\epsilon)^{-3} 2 C_{\mathcal{H}}(s) \left(\int \left(\frac{d\tilde{P}}{d\mathcal{H}^s_{|F}}(x) \right)^{s/(2+s)} d\mathcal{H}^s_{|F}(x) \right)^{1+2/s}$$
$$= (1+\epsilon)^{-3} 2 C_{\mathcal{H}}(s) \left(\int \left(J_{\psi,E}(\psi^{-1}(x))^{-1} p(\psi^{-1}(x)) \right)^{s/(2+s)} d\mathcal{H}^s_{|F}(x) \right)^{1+2/s}$$
$$= (1+\epsilon)^{-3} 2 C_{\mathcal{H}}(s) \left(\int J_{\psi,E}(x)^{2/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x) \right)^{1+2/s}$$
$$\geq (1+\epsilon)^{-6-1/s} C_{\mathcal{H}}(s) \left(\int J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x) \right)^{1+2/s}. \tag{26}$$

Analoguously, one can prove

$$\Delta_n(f, P) \leq (1+\epsilon)^3 \Delta_n(\|.\|^2, \tilde{P}) \tag{27}$$

and

$$\limsup_{n\to\infty} n^{2/s} \Delta_n(f, P) \leq (1+\epsilon)^{6+1/s} C_{\mathcal{H}}(s) \left(\int J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x) \right)^{1+2/s}. \tag{28}$$

We proceed by showing that, given the assumptions of Theorem 2, for all $\epsilon > 0$, a $m \in \mathbb{N}$, measurable sets $E_0, \ldots, E_m$, positive definite matrices $A_1, \ldots, A_m$ and Lipschitz functions $\phi_1, \ldots, \phi_m$, with $\phi_i : E_i \to F_i \subseteq \mathbb{R}^s$ exist, which satisfy:

(a) $f \in C^2(\overline{co}(E_i))$ for $1 \le i \le m$ and $f'' > 0$ on $\overline{co}(E_i)$.

(b) $\cup_{i=0}^m E_i = \mathbb{R}^d$,

(c) $\int_{E_0} \lambda(\|x\|c)\eta_{s,G,\|x\|2c}(x)^{-2/s} dP(x) < \epsilon$ and $\int_{E_0} J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s|E(x)$ $< \epsilon \int J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s|E(x)$.

(d) (22) holds for $x \in \overline{co}(E_i)$ when A is replaced by A_i, $1 \le i \le m$,

(e) (23) holds for P_{ac}-a.e. $x \in E_i$, with A replaced by A_i, $1 \le i \le m$,

(f) (24) holds for $x, y \in E_i$ and ϕ_i instead of ϕ, $1 \le i \le m$.

The existence of these objects can be established in the following way. Let $c \in]1, \infty[$ and G satisfy the assumptions of the theorem, let $\epsilon' > 0$, such that for all Borel sets B with $P(B) < \epsilon'$, $\int_B \lambda(\|x\|c)\eta_{s,G,\|x\|2c}(x)^{-2/s} dP(x) < \epsilon/3$ holds. Let $\tilde{E}_0, \tilde{E}_1, \ldots, \tilde{E}_k$ be measurable sets, where $\tilde{E}_0 = (\cup_{i\ge1}\tilde{E}_i)^c$, $P(\tilde{E}_0) < \epsilon'$, $\tilde{E}_i$ is compact and convex with nonempty interior and such that f'' is continuous on $\tilde{E}_i$. These sets may be chosen in a way that for suitable positive definite matrices A_i assumption (d) holds. Let $\tilde{E}_i^* = \{x \mid$ a tangent to $E \cap \tilde{E}_i$ exists at $x\}$. Then $P(\cup_{i\ge1}(\tilde{E}_i \setminus \tilde{E}_i^*)) = 0$. Denote by θ_x a tangent at $x \in \tilde{E}_i^*$. We identify θ_x with an orthogonal $d \times s$ matrix, the columns of which span $\theta_x - x$. A parametrization of tangents exists, for which the mappings $\Theta_i : x \mapsto \theta_x$ are measurable. According to the Theorem of Lusin $\tilde{E}_i^{**} \subsetneq \tilde{E}^*$ and continuous mappings $\tilde{\Theta}_i$ exist, such that $\Theta_i = \tilde{\Theta}_i$ on $\tilde{E}_i^{**}$ and $P_{ac}(\cup_{i\ge1}(\tilde{E}_i^* \setminus \tilde{E}_i^{**})) < \epsilon'$. Finally, $\tilde{E}_i$ may be split into a finite number of compact and convex sets $E_{i,j}, j = 1, \ldots, k_j$, and a residue set $E_{i,0}$, such that (e) and (f) hold for $\tilde{E}_i^{**} \cap E_{i,j}$ and furthermore $P(\cup_{i\ge1}E_{i,0}) < \epsilon'$. Define $m = \sum_{i=1}^k k_j$, denote the sequence of sets $\tilde{E}_i^{**} \cap E_{i,j}, 1 \le j \le k_j, 1 \le i \le k$, by $(E_l), 1 \le l \le m$ and let $E_0 = (\cup_{i\ge1}E_i)^c$. This sequence satisfies (a) - (f).

The proof of the theorem is based on the following considerations. In a first step the asymptotic behavior of the quantization error on a sufficiently flat support and for smooth enough information function f has been derived. Then the support of P is approximated by the union of sufficiently flat sets E_i on which f is smooth enough. It remains to determine the asymptotically optimal fraction of prototypes in each set E_i. Let $\tilde{\epsilon} = \epsilon^{s/4}$ and for $1 \le i \le m$ define $a_i = \int_{E_i} J_{f',E_i}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x)$ and $n_i = [n(1-\tilde{\epsilon})a_i / \sum_{j=1}^m a_j]$, $n_0 = n - \sum_{i=1}^m n_i$. Since $J_{f',E_i}(x) = J_{f',E}(x)$ for P_{ac}-a.e. $x \in E_i$, Lemma 2 implies

$$\limsup_{n\to\infty} n^{2/s} \Delta_n(f, P) \le \limsup_{n\to\infty} n^{2/s} \sum_{i=0}^m P(E_i)\Delta_{n_i}(f, P_{\restriction E_i})$$

$$\leq \limsup_{n\to\infty} \sum_{i=0}^{m} P(E_i)(n/n_i)^{2/s} \limsup_{n\to\infty} n_i^{2/s} \Delta_{n_i}(f, P_{\restriction E_i})$$

$$\leq P(E_0)\tilde{\epsilon}^{-2/s} \int_{E_0} \lambda(\|x\|c)\eta_{s,G,\|x\|2c}(x)^{-2/s} dP(x)/P(E_0)$$

$$+\sum_{i=1}^{m} P(E_i) \left(\frac{\sum_{j=1}^{m} a_j}{(1-\tilde{\epsilon})a_i}\right)^{2/s} C_{\mathcal{H}}(s) \left(\int_{E_i} J_{f',E_i}(x)^{1/(2+s)}(p(x)/P(E_i))^{s/(2+s)} d\mathcal{H}^s_{|E}(x)\right)^{1+2/s}$$

$$\leq \tilde{\epsilon}^{-2/s}\epsilon + \sum_{i=1}^{m} C_{\mathcal{H}}(s) \left(\frac{\sum_{j=1}^{m} a_j}{(1-\tilde{\epsilon})a_i}\right)^{2/s} a_i^{1+2/s}$$

$$= \epsilon^{1/2} + C_{\mathcal{H}}(s) \frac{(\sum_{j=1}^{m} a_j)^{1+2/s}}{(1-\tilde{\epsilon})^{2/s}}$$

$$\leq \epsilon^{1/2} + C_{\mathcal{H}}(s)(1-\tilde{\epsilon})^{-2/s} \left(\int J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x)\right)^{1+2/s}.$$

Thus

$$\limsup_{n\to\infty} n^{2/s}\Delta_n(f,P) \leq C_{\mathcal{H}}(s) \left(\int J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x)\right)^{1+2/s} \tag{29}$$

holds. To prove

$$\liminf_{n\to\infty} n^{2/s}\Delta_n(f,P) \geq C_{\mathcal{H}}(s) \left(\int J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x)\right)^{1+2/s}, \tag{30}$$

let for $n \in \mathbb{N}$, O_n be a set of prototypes and $(U_i)_{i=1}^m$ a sequence of open, pairwise disjoint sets with $\overline{co}(E_i) \subseteq U_i$. According to Lemma 4 a n_0 exists, such that for $n \geq n_0$ the winners for E_i lie in U_i. Let $n_i = |O_n \cap U_i|$. Then, applying Lemma 3 and statement (c),

$$\liminf_{n\to\infty} n^{2/s}\Delta_n(f,P) \geq \liminf_{n\to\infty} \sum_{i=1}^{m} P(E_i)(n/n_i)^{2/s} n_i^{2/s} \Delta_{n_i}(f, P_{\restriction E_i})$$

$$\geq \inf_{\alpha_i\geq 0, \alpha_1+\cdots+\alpha_m=1} \sum_{i=1}^{m} P(E_i)\alpha_i^{-2/s} C_{\mathcal{H}}(s) \left(\int_{E_i} J_{f',E}(x)^{1/(2+s)}(p(x)/P(E_i))^{s/(2+s)} d\mathcal{H}^s_{|E}(x)\right)^{1+2/s}$$

$$= C_{\mathcal{H}}(s) \left(\sum_{i=1}^{m} \int_{E_i} J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x)\right)^{1+2/s}$$

$$\geq C_{\mathcal{H}}(s) \left((1-\epsilon)\int J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x)\right)^{1+2/s}. \square$$

COROLLARY 2. *Let P_{ac} be the absolutely continuous part of P with respect to $\mathcal{H}^s_{|E}$ and let $\tilde{P} << P_{ac}$ with density proportional to $J^{1/s}_{E,f'}$. If the assumtions of Theorem 2 hold, then*

1. If (O_n) is asymptotically optimal for (f, P), i.e.

$$\lim_{n\to\infty} \frac{\Delta_n(f, O_n, P)}{\Delta_n(f, P)} = 1,$$

then (O_n) is asymptotically optimal for (f, P_{ac}) and for $(\|.\|^2, \tilde{P})$.

2. If $(\tilde{O}_n)$ is asymptotically optimal for (f, P_{ac}) or for $(\|.\|^2, \tilde{P})$, then a sequences (O'_n) and (k_n) with $k_n/n \to 1$ exist, such that $(O_n) = (\tilde{O}_{k_n} \cup O'_{n-k_n})$ is asymptotically optimal for (f, P).

4 Nonregular quantization problems

A result corresponding to the Theorem of Zador for nonregular quantization problems, i.e. when f'' is not of full rank, requires the following considerations. An information function f, which is continuously differentiable on *csupp*(P) induces a partition $R(x) := (f')^{-1} \circ f'(x)$ of *csupp*(P) into closed sets. Let $(P(. \mid f' = f'(x))_{x \in csupp(P)}$ be a regular conditional probability, $\eta(x) = \mathbb{E}(X \mid f' = f'(x))$ the conditional mean of X given $f' = f'(x)$, $E = \{\eta(x) \mid x \in csupp(P)\}$ and P^* the distribution of η under $x \sim P$. f is linear on each set $R(x)$, since f is convex, which implies for $f'(x) = f'(y)$,

$$\begin{aligned} 0 &\le f(x) - f(y) - \langle x - y, f'(y)\rangle \\ &= f(x) - f(y) - \langle x - y, f'(x)\rangle \\ &= -(f(y) - f(x) - \langle y - x, f'(x)\rangle) \le 0. \end{aligned}$$

The errors of both quantization problems (f, P) and (f, P^*) are equal,

$$\Delta_n(f, P) = \Delta_n(f, P^*).$$

Under assumptions, similar to those under which Theorem 2 holds, the limit of $n^{2/s}\Delta_n(f, P^*)$ ($s = dim_{\mathcal{H}}(E)$) may be obtained, even if f'' is not regular on a neighborhood of E. It is sufficient that f'' is strictly positive tangentially to E.

THEOREM 3. *Let $s \in \mathbb{N}$, E a regular set of dimension s with $P^*(E) = 1$ and $P^*(\{x \mid E$ has a tangent at $x\}) = 1$. Let $p = dP^*/d\mathcal{H}^s_{|E}$. If f satisfies assumption B and if a distribution*

G and a $c \in]1,\infty[$ *exist with* $\int \lambda(\|x\|c)\eta_{s,G,\|x\|c}(x)^{-2/s}dP^*(x) < \infty$, *then*

$$\lim_{n\to\infty} n^{2/s}\Delta_n(f,P) = C_{\mathcal{H}}(s)\left(\int J_{f',E}(x)^{1/(2+s)}p(x)^{s/(2+s)}d\mathcal{H}^s_{|E}(x)\right)^{1+2/s}. \tag{31}$$

PROOF. The inequality

$$\limsup_{n\to\infty} n^{2/s}\Delta_n(f,P) \le C_{\mathcal{H}}(s)\left(\int J_{f',E}(x)^{1/(2+s)}p(x)^{s/(2+s)}d\mathcal{H}^s_{|E}(x)\right)^{1+2/s} \tag{32}$$

follows directly from Theorem 2 and the following modification. Let for a given $\epsilon > 0$, $U \subseteq \mathbb{R}^d$ be an open set with $\int_{U^c}\lambda(\|x\|c)\eta_{s,G,\|x\|2c}(x)^{-2/s}dP(x) < \epsilon$. Let $\tilde{\epsilon} = \epsilon^{s/4}$, $n_1 = [(1-\tilde{\epsilon})n]$, $n_0 = n - n_1$ and $f_\epsilon(x) = f(x) + \epsilon\|x\|^2$. f''_ϵ is strictly positive on $\mathbb{R}^d$. Note that $E \cap U$ is regular. Therefore

$$\begin{aligned}\limsup_{n\to\infty} n^{2/s}\Delta_n(f,P) &\le \limsup_{n\to\infty} P(U^c)(n/n_0)^{2/s}\limsup_{n\to\infty} n_0^{2/s}\Delta_{n_0}(f, P\restriction U^c)\\ &\quad + \limsup_{n\to\infty} P(U)(n/n_1)^{2/s}\limsup_{n\to\infty} n_1^{2/s}\Delta_{n_1}(f_\epsilon, P\restriction U)\\ &\le \epsilon^{1/2} + C_{\mathcal{H}}(s)(1-\tilde{\epsilon})^{-2/s}\left(\int_U J_{f'_\epsilon,E}(x)^{1/(2+s)}p(x)^{s/(2+s)}d\mathcal{H}^s|E(x)\right)^{1+2/s}.\end{aligned}$$

$J_{f_\epsilon,E}(x)$ is monotone in ϵ, implying (32).
The remaining inequality

$$\liminf_{n\to\infty} n^{2/s}\Delta_n(f,P) \ge C_{\mathcal{H}}(s)\left(\int J_{f',E}(x)^{1/(2+s)}p(x)^{s/(2+s)}d\mathcal{H}^s_{|E}(x)\right)^{1+2/s} \tag{33}$$

is nontrivial only if $\int J_{f',E}(x)^{1/(2+s)}p(x)^{s/(2+s)}d\mathcal{H}^s_{|E}(x) > 0$. Let F_z denote the projection onto the null space of $f''(z)$ and $\lambda^*(z)$ the smallest strictly positive eigenvalue of $f''(z)$. If f'' is 0, we set $\lambda^*(z) = 0$. The assumptions of the theorem imply for every $\epsilon > 0$ the existence of $c_0 > 0$ and $c_1 \in]0,1[$, furthermore of measurable sets $E_1,\ldots,E_m$, with $\overline{co}(E_1),\ldots,\overline{co}(E_m)$ being pairwise disjoint, such that statements (a) to (d) hold:

(a)

$$\int_{E\setminus\cup E_i} J_{f',E}(x)^{1/(2+s)}p(x)^{s/(2+s)}d\mathcal{H}^s_{|E}(x) < \epsilon\int_{\cup E_i} J_{f',E}(x)^{1/(2+s)}p(x)^{s/(2+s)}d\mathcal{H}^s_{|E}(x),$$

(b) $\inf\{\lambda^*(x) \mid x \in \overline{co}(E_i)\} \ge c_0$, for $1 \le i \le m$,

(c) if $x, y \in E_i$ and $z \in \overline{co}(E_i)$, then $\|F_z(x-y)\| \le c_1\|x-y\|$,

(d) $\eta(\overline{co}(E_i)) \subseteq E_i$.

Let for $1 \leq i \leq m$, $O_n^i = \{p_1, \dots, p_n\}$ be a set of prototypes for $(f, P^*_{\restriction E_i})$. Then

$$\Delta_n(f, P^*_{\restriction E_i}) = \Delta(f, O_n^i, P^*_{\restriction E_i})$$

$$= \int \min_{1 \leq j \leq n} \left(f(x) - f(p_j) - \langle x - p_j, f'(p_j) \rangle \right) dP^*_{\restriction E_i}(dx).$$

Therefore a $z = z(x) \in \overline{co}(E_i)$ exists with

$$\Delta(f, O_n^i, P^*_{\restriction E_i}) = \int \min_{1 \leq j \leq n} \frac{1}{2}(x - p_j)^\top f''(z)(x - p_j) dP^*_{\restriction E_i}(dx)$$

$$\geq \frac{c_0(1 - c_1)}{2} \int \min_{1 \leq j \leq n} \|x - p_j\|^2 dP^*_{\restriction E_i}(dx) = \frac{c_0(1 - c_1)}{2} \Delta_n(\|.\|^2, P^*_{\restriction E_i}).$$

Let $\epsilon' = \epsilon c_0(1 - c_1)/2$, U_i open and pairwise disjoint neighborhoods of $\overline{co}(E_i)$ and (O_n) a sequence of prototypes for $(f, P^*_{\restriction \cup E_i})$. A n_0 exists, such that for $n \geq n_0$ the sets $O_{n_i}^i = O_n \cap U_i$ are optimal for $(f, P^*_{\restriction E_i})$. Here, $n_i = |O_n \cap U_i|$. Then, for the regular information funktion $f_{\epsilon'} = f + \epsilon' \|.\|^2$,

$$C_{\mathcal{H}}(s) \left(\int J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x) \right)^{1+2/s}$$

$$\leq (1+\epsilon)^{1+2/s} C_{\mathcal{H}}(s) \left(\int_{\cup E_i} J_{f',E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x) \right)^{1+2/s}$$

$$\leq (1+\epsilon)^{1+2/s} C_{\mathcal{H}}(s) \left(\int_{\cup E_i} J_{f'_{\epsilon'},E}(x)^{1/(2+s)} p(x)^{s/(2+s)} d\mathcal{H}^s_{|E}(x) \right)^{1+2/s}$$

$$= (1+\epsilon)^{1+2/s} \liminf_{n \to \infty} n^{2/s} \Delta_n(f_{\epsilon'}, P^*_{|\cup E_i})$$

$$\leq (1+\epsilon)^{1+2/s} \liminf_{n \to \infty} n^{2/s} \sum_{i=1}^m \Delta_{n_i}(f_{\epsilon'}, O_{n_i}^i, P^*_{|E_i})$$

$$= (1+\epsilon)^{1+2/s} \liminf_{n \to \infty} n^{2/s} \sum_{i=1}^m \left(\Delta_{n_i}(f, O_{n_i}^i, P^*_{|E_i}) + \epsilon' \Delta_{n_i}(\|.\|^2, O_{n_i}^i, P^*_{|E_i}) \right)$$

$$\leq (1+\epsilon)^{2+2/s} \liminf_{n \to \infty} \sum_{i=1}^m n^{2/s} \Delta_{n_i}(f, P^*_{|E_i})$$

$$\leq (1+\epsilon)^{2+2/s} \liminf_{n \to \infty} n^{2/s} \Delta_n(f, P^*). \square$$

5 The asymptotic distribution of prototypes

The proof of Theorem 2 is based on the following procedure. The support of P is split into small sets E_i. The major part of these sets is sufficiently "regular" and the information function is smooth enough on these sets to guarantee the convergence of $n^{2/s}\Delta_n(f, P_{\restriction E_i})$. The optimality of a set of prototypes determines the fraction of prototypes in neighborhoods of the sets E_i and therefore the asymptotic distribution of the prototypes. Recall that $\hat{Q}_n$ denotes the uniform distribution on O_n.

Let the quantization problem (f, P) be given. A sequence $(n_k)_{k=1}^{\infty}$ satisfies

ASSUMPTION C1., *if for all measurable sets E and all $\epsilon > 0$ a compact set F exists, such that for all $\epsilon' > 0$,*

$$\limsup_{k\to\infty} \frac{[\epsilon' n_k]^{2/s}\Delta_{[\epsilon' n_k]}(f, P_{|(E\setminus F)})}{n_k^{2/s}\Delta_{n_k}(f, P)} < \epsilon. \tag{34}$$

The sequence $(n_k)_{k=1}^{\infty}$ satisfies

ASSUMPTION C2., *if for all open sets U a real number $\alpha(U) \in [0,1]$ exists, such that for all subsequences (n'_k) of (n_k),*

$$\lim_{k\to\infty} \frac{n_k'^{2/s}\Delta_{n'_k}(f, P_{|U})}{n_k^{2/s}\Delta_{n_k}(f, P)} = \alpha(U). \tag{35}$$

THEOREM 4. *If the sequence $(n_k)_{k=1}^{\infty}$ satisfies C1 and C2, then a distribution Q on $\mathbb{R}^d$ exists, such that for all open sets U,*

$$Q(U) = \alpha(U)^{s/(2+s)}. \tag{36}$$

$(\hat{Q}_{n'_k})$ converges weakly to Q, for any subsequence (n'_k) of (n_k) and any sequence of prototypes $(O_{n'_k})$.

PROOF. Let $E \subseteq \mathbb{R}^d$ closed and U an arbitrary open neighborhood of E. Let $\epsilon > 0$. Furthermore, let $F \subseteq E^c$ compact and U_1, U_2 open with $E \subseteq U_1 \subseteq U$, $F \subseteq U_2$ and $U_1 \cap U_2 = \emptyset$, such that for all $\epsilon' > 0$,

$$\limsup_{k\to\infty} \frac{[\epsilon' n_k]^{2/s}\Delta_{[\epsilon' n_k]}(f, P_{|(E^c\setminus F)})}{n_k^{2/s}\Delta_{n_k}(f, P)} < \epsilon. \tag{37}$$

U_1 and U_2 may be chosen such that open sets U_3 and U_4 exist, with $\bar{U}_3 \subseteq U_1$, $\bar{U}_4 \subseteq U_2$, $\alpha(U_3) \geq \alpha(U_1)(1-\epsilon)^{1+2/s}$ and $\alpha(U_4) \geq \alpha(U_2)(1-\epsilon)^{1+2/s}$.

Let $\beta_1 = \alpha(U_1)^{s/(2+s)}$ and $\beta_2 = \alpha(U_2)^{s/(2+s)}$. To begin with, we show that $\beta_1 + \beta_2$ is approximately 1. Let $\epsilon' = \epsilon^{s/4}$. We define $n_{1,k} = [n_k(1-\epsilon')\beta_1/(\beta_1+\beta_2)]$, $n_{2,k} = [n_k(1-\epsilon')\beta_2/(\beta_1+\beta_2)]$ and $n_{3,k} = n_k - n_{1,k} - n_{2,k}$. Let $\tilde{O}_{n_k} = \tilde{O}_{n_{1,k}} \cup \tilde{O}_{n_{2,k}} \cup \tilde{O}_{n_{3,k}}$, where the sets $\tilde{O}_{n_{i,k}}$ are optimal for $(f, P_{|E})$, $(f, P_{|F})$ and $(f, P_{|(E^c \setminus F)})$ respectively. According to Lemma 2,

$$n_k^{2/s}\Delta_{n_k}(f,P) \le (n_k/n_{1,k})^{2/s} n_{1,k}^{2/s}\Delta_{n_{1,k}}(f, P_{|E})$$

$$+(n_k/n_{2,k})^{2/s} n_{2,k}^{2/s}\Delta_{n_{2,k}}(f, P_{|E}) + (n_k/n_{3,k})^{2/s} n_{3,k}^{2/s}\Delta_{n_{3,k}}(f, P_{|(E^c\setminus F)})$$

and thus

$$1 \le \left(\frac{\beta_1+\beta_2}{\beta_1(1-\epsilon')}\right)^{2/s} \beta_1^{1+2/s} + \left(\frac{\beta_1+\beta_2}{\beta_2(1-\epsilon')}\right)^{2/s} \beta_2^{1+2/s} + \frac{\epsilon}{(\epsilon')^{2/s}}$$
$$= \frac{1}{(1-\epsilon')^{2/s}}(\beta_1+\beta_2)^{1+2/s} + \epsilon^{1/2},$$

which implies $\beta_1 + \beta_2 \ge 1 - \tilde{\epsilon}$ with $1 - \tilde{\epsilon} = (1-\epsilon^{1/2})^{s/(2+s)}(1-\epsilon^{s/4})^{2/(2+s)}$.
Let (O_{n_k}) be a sequence of prototypes for (f, P) and $n^*_{i,k} = |O_{n_k} \cap U_i|$. A n^* exists (see Lemma 4) such that for $n_k \ge n^*$ all winners of U_3 are in U_1 and all winners of U_4 in U_2. Hence

$$n^{2/s}\Delta_{n_k}(f,P) \ge (n_k/n^*_{1,k})^{2/s}(n^*_{1,k})^{2/s}\Delta_{n^*_{1,k}}(f, P_{|U_3}) + (n_k/n^*_{2,k})^{2/s}(n^*_{2,k})^{2/s}\Delta_{n^*_{2,k}}(f, P_{|U_4}),$$

and therefore for all accumulation points t_i of $n^*_{i,k}/n_k$,

$$1 \ge t_1^{-2/s}\alpha(U_3) + t_2^{-2/s}\alpha(U_4)$$
$$\ge (t_1^{-2/s}\beta_1^{1+2/s} + t_2^{-2/s}\beta_2^{1+2/s})(1-\epsilon)^{1+2/s}$$
$$= (t_1+t_2)\left(\frac{t_1}{t_1+t_2}(\frac{\beta_1}{t_1})^{1+2/s} + \frac{t_2}{t_1+t_2}(\frac{\beta_2}{t_2})^{1+2/s}\right)(1-\epsilon)^{1+2/s}$$
$$\ge (t_1+t_2)\left(\frac{t_1}{t_1+t_2}\frac{\beta_1}{t_1} + \frac{t_2}{t_1+t_2}\frac{\beta_2}{t_2}\right)^{1+2/s}(1-\epsilon)^{1+2/s}.$$

Consequently, $\beta_1 + \beta_2 \le 1/(1-\epsilon)$.
$\epsilon \to 0$ implies $\tilde{\epsilon} \to 0$ and thus $\beta_1 + \beta_2 \to 1$. Let $\beta_1 + \beta_2 = 1$. $t_1^{-2/s}\beta_1^{1+2/s} + t_2^{-2/s}\beta_2^{1+2/s}$ is minimal on $0 \le t_i \le 1$, $t_1 + t_2 \le 1$ in $t_i = \beta_i$ and is then equal to 1. Thus a function $\tilde{\tilde{\epsilon}}$ exists with $\lim_{\epsilon \to 0+} \tilde{\tilde{\epsilon}}(\epsilon) = 0$ and

$$|t_1 - \beta_1| + |t_2 - \beta_2| \le \tilde{\tilde{\epsilon}}(\epsilon). \tag{38}$$

Hence we have proved that for all closed sets E, for all open $U \supseteq E$ and all $\epsilon > 0$ an open U_1 exists with $E \subseteq U_1 \subseteq U$ such that for all sequences of prototypes (O_{n_k}),

$$\limsup_{k\to\infty}\left|\frac{|O_{n_k} \cap U_1|}{n_k} - \alpha(U_1)^{s/(2+s)}\right| < \epsilon. \tag{39}$$

Assumption *C1* guarantees that $(\hat{Q}_{n_k})$ is tight. For any accumulation point μ and any subsequence (n'_k) with $\hat{Q}_{n'_k} \to^w \mu$,

$$\mu(E) - \alpha(U_1)^{s/(2+s)} \leq \mu(U_1) - \alpha(U_1)^{s/(2+s)}$$

$$\leq \liminf_{k\to\infty} \hat{Q}_{n'_k}(U_1) - \alpha(U_1)^{s/(2+s)} \leq \epsilon.$$

Let $U \supseteq E$ open, with $\mu(\partial U) = 0$, such that $\mu(U) \leq \mu(E) + \epsilon$. Then

$$\mu(E) - \alpha(U_1)^{s/(2+s)} \geq \mu(U) - \alpha(U_1)^{s/(2+s)} - \epsilon$$

$$= \lim_{k\to\infty} \hat{Q}_{n'_k}(U) - \alpha(U_1)^{s/(2+s)} - \epsilon$$

$$\geq \liminf_{k\to\infty} \hat{Q}_{n'_k}(U_1) - \alpha(U_1)^{s/(2+s)} - \epsilon \geq -2\epsilon.$$

Therefore

$$\mu(E) = \inf\{\alpha(U)^{s/(2+s)} \mid U \supseteq E, \text{open}\}.$$

In particular, the accumulation point μ is unique. In virtue of assumption *C1*, α is continuous in the following sense. For U open and $\epsilon > 0$ a compact set $E \subseteq U$ exists such that for all open sets V with $E \subseteq V \subseteq U$, $\alpha(U) \leq \alpha(V) + \epsilon$. With that $\alpha(U) = \mu(U)$ may be shown. □

COROLLARY 3. *Let the assumptions of Theorem 2 hold.*

1. *Assume* $p = dP/d\mathcal{H}^s_{|E} \neq 0$ *and let* $Q << \mathcal{H}^s_{|E}$ *with density proportional to* $J_{f',E}^{1/(2+s)} p^{s/(2+s)}$. $\hat{Q}_n$ *converges weakly to* Q *for any sequence of prototypes.*

2. *Let* $P << \mathfrak{L}^d$ *with* $dP/d\mathfrak{L}^d = p$. $(\hat{Q}_n)$ *converges weakly to* P, *if a* $c \in \mathbb{R}_+$ *exists such that* f *is a solution of the Monge-Ampère differential equation*

$$det(f'') = cp^2. \tag{40}$$

3. *Let* E *be a Borel set with* $P << \mathcal{H}^s_{|E}$. *If* Q, *defined in 1., satisfies* $\int \lambda(\|x\|c)\eta_{s,Q,\|x\|c}(x)^{-2/s} dP(x) < \infty$ *(with* $c > 1$*), then the random quantization rule defined by* Q *is optimal among random quantization rules, i.e. for all* $\tilde{O}^G_n \sim G$, $\mathbb{E}^Q(\Delta_n(f, \tilde{O}^Q_n, P)) \leq \mathbb{E}^G(\Delta_n(f, \tilde{O}^G_n, P))$.

6 Auxiliary results

LEMMA 1. *Let f_1 and f_2 be convex functions with $f_2 - f_1$ convex. Then $\Delta_n(f_1, P) \leq \Delta_n(f_2, P)$.*

PROOF. Let O^1 and O^2 be sets of prototypes, which are optimal for f_1 and f_2 respectively. Then

$$\Delta(f_1, P) = \Delta(f_1, O^1, P) = \int f_1 dP - I^{f_1}(O^1, P)$$

$$\leq \int f_1 dP - I^{f_1}(O^1, P) + \Delta(f_2 - f_1, O^2, P)$$

$$= \int f_1 dP - I^{f_1}(O^1, P) + \int (f_2 - f_1) dP - I^{f_2 - f_1}(O^2, P)$$

$$= \int f_2 dP - I^{f_2}(O^2, P) + I^{f_1}(O^2, P) - I^{f_1}(O^1, P)$$

$$\leq \Delta(f_2, O^2, P).$$

The last inequality is a consequence of $I^{f_1}(O^2, P) \leq I^{f_1}(O^1, P)$. □

LEMMA 2. *For distributions $P_1, P_2, \cdots, P_k$ and $P = \epsilon_1 P_1 + \cdots + \epsilon_k P_k$ with $\epsilon_i \in [0,1]$, $\epsilon_1 + \cdots + \epsilon_k = 1$,*

$$\Delta_n(f, P) \geq \sum_{i=1}^{k} \epsilon_i \Delta_n(f, P_i). \tag{41}$$

If $0 < n_i < n$, $n_1 + \cdots + n_k = n$, then

$$\Delta_n(f, P) \leq \sum_{i=1}^{k} \epsilon_i \Delta_{n_i}(f, P_i). \tag{42}$$

LEMMA 3. *Let $\beta_1, \ldots, \beta_k, \in]0, \infty[$,*

$$\phi(\alpha_1, \ldots, \alpha_k) := \sum_{i=1}^{k} \beta_i^{1+2/s} \alpha_i^{-2/s}.$$

The unique minimum of ϕ on the simplex $\{(\alpha_1, \ldots, \alpha_k) | \alpha_i \geq 0, \alpha_1 + \cdots + \alpha_k = 1\}$ is attained in $\alpha_i = \beta_i / \sum_{j=1}^{k} \beta_j$, where ϕ is $(\sum_{i=1}^{k} \beta_i)^{1+2/s}$.

LEMMA 4. *1. Let f'' be positive definite on csupp(P), let $F_1, \ldots, F_m$ denote compact subsets of E and U_i open neighborhoods of F_i. For every sequence of (sets of) prototypes (O_n) a n^* exists, such that for $n \geq n^*$ all winners of elements of F_i are in U_i.*

2. Let $E \subseteq \mathbb{R}^d$ be bounded and regular with $P(E) = 1$ and $P(\{x \mid$ a tangent to E at x exists $\})$ $= 1$. Furthermore, for all $x \in E$, $E \cap (f')^{-1} \circ f'(x) = \{x\}$. $F_1, \ldots, F_m$ denote compact subsets of E and U_i open neighborhoods of F_i. For every sequence of (sets of) prototypes (O_n) a n^* exists, such that for $n \geq n^*$ all winners of elements of F_i are in U_i.

Falconer [1], [2], Federer [3], Morgan [7] and Tricot [11] provide accounts of the concepts and results on geometric measure theory presented in this section.
Let $E \subseteq \mathbb{R}^d$, $\delta > 0$. A δ-cover of E is a finite or denumerably infinite family of sets $\{U_i\}$ such that $E \subseteq \cup_i U_i$ and for all i, $\mathrm{diam}(U_i) \leq \delta$. $\mathrm{diam}(A) := \sup\{\|x - y\| \mid x, y \in A\}$ denotes the diameter of A. For $s > 0$ let

$$\mathcal{H}^s_\delta(E) = \inf\{\sum_{i=1}^{\infty} \mathrm{diam}(U_i)^s \mid \{U_i\} \text{ is a } \delta\text{-cover of } E\} \tag{43}$$

and

$$\mathcal{H}^s(E) = \lim_{\delta \to 0+} \mathcal{H}^s_\delta(E). \tag{44}$$

$\mathcal{H}^s(E)$ exists in $[0, \infty]$, as $\mathcal{H}^s_\delta(E)$ is increasing in δ. $\mathcal{H}^s(E)$ is the Hausdorff measure of E. $\mathcal{H}^s$ is a Borel-regular outer measure on $\mathbb{R}^d$. Borel sets are measurable. For $s = d$ $\mathcal{H}^d = c_d^{-1}\mathcal{L}^d$, where $\mathcal{L}^d$ is Lebesgue measure and $c_d = \pi^{d/2}2^{-d}/\Gamma(1 + d/2)$ is the volume of the d-dimensional ball of diameter 1.
For every set $E \subseteq \mathbb{R}^d$ a real s, the Hausdorff dimension $dim_{\mathcal{H}}(E)$ of E, exists, such that

$$\mathcal{H}^\alpha(E) = \begin{cases} \infty & \text{for } \alpha < s, \\ 0 & \text{for } \alpha > s. \end{cases}$$

An s-set is a set E such that $0 < \mathcal{H}^s(E) < \infty$, where $s = dim_{\mathcal{H}}(E)$.
Packing measure and packing dimension are defined analogously. A finite or denumerably infinite family of disjoint balls $\{B_i\}$ with centers in E is called a δ-packing of E, if for all i, $\mathrm{diam}(B_i) \leq \delta$ holds. Define

$$\mathcal{P}^s_\delta(E) = \sup\{\sum_{i=1}^{\infty} \mathrm{diam}(B_i)^s \mid \{B_i\} \text{ is a } \delta\text{-packing of } E\} \tag{45}$$

and

$$\mathcal{P}^s_0(E) = \lim_{\delta \to 0+} \mathcal{P}^s_\delta(E). \tag{46}$$

$\mathcal{P}^s_0$ itself is not an outer measure but allows the definition of the packing measure

$$\mathcal{P}^s(E) = \inf\{\sum_{i=1}^{\infty} \mathcal{P}^s_0(E_i) \mid E \subseteq \cup_i E_i\}. \tag{47}$$

The packing dimension of a set E is the discontinuity of the mapping $\alpha \mapsto \mathcal{P}^\alpha(E)$.

Define the covering number $N(E,r)$ of a bounded set $E \subseteq \mathbb{R}^d$ as the minimal cardinality of an r-cover of E. The lower and the upper box-counting dimension of E are given by

$$\underline{dim}_B(E) = \liminf_{r\to 0+} \frac{\log N(E,r)}{\log(1/r)}, \tag{48}$$

$$\overline{dim}_B(E) = \limsup_{r\to 0+} \frac{\log N(E,r)}{\log(1/r)}. \tag{49}$$

A set E with $\overline{dim}_B(E) = s$ is a set of exact upper box-counting dimension, if $N(E,r) = O(r^{-s})$. Recall that $dim_{\mathcal{H}}(E) \leq dim_{\mathcal{P}}(E) \leq \overline{dim}_B(E)$ and $dim_{\mathcal{H}}(E) \leq \underline{dim}_B(E) \leq \overline{dim}_B(E)$.
The Hausdorff dimension $dim^*_{\mathcal{H}}(P)$ and the packing dimension $dim^*_{\mathcal{P}}(P)$ of a distribution P are defined as

$$dim^*_{\mathcal{H}}(P) = \inf\{dim_{\mathcal{H}}(E) \mid E \subseteq \mathbb{R}^d,\ \text{Borel and } P(E) = 1\}, \tag{50}$$

$$dim^*_{\mathcal{P}}(P) = \inf\{dim_{\mathcal{P}}(E) \mid E \subseteq \mathbb{R}^d,\ \text{Borel and } P(E) = 1\}. \tag{51}$$

Analoguously, the lower and the upper box-counting dimension of P are

$$\underline{dim}^*_B(P) = \inf\{\underline{dim}_B(E) \mid E \subseteq \mathbb{R}^d,\ \text{Borel and } P(E) = 1\}, \tag{52}$$

$$\overline{dim}^*_B(P) = \inf\{\overline{dim}_B(E) \mid E \subseteq \mathbb{R}^d,\ \text{Borel and } P(E) = 1\}. \tag{53}$$

Note that for $s = dim^*_{\mathcal{H}}(P)$ (and analoguously for $dim^*_{\mathcal{P}}(P)$, $\overline{dim}_B(P)$) a Borel set E exists, such that $P(E) = 1$ and $dim_{\mathcal{H}}(E) = s$ ($dim_{\mathcal{P}}(E) = s$, $\overline{dim}_B(E) = s$). The upper and lower densities of E at x are defined as $\overline{D}^s(E,x) = \limsup_{r\to 0+} \mathcal{H}^s(E \cap B(x,r))/(2r)^s$ and $\underline{D}^s(E,x) = \liminf_{r\to 0+} \mathcal{H}^s(E \cap B(x,r))/(2r)^s$.

A set $E \subseteq \mathbb{R}^d$ is called regular, if for a s and $\mathcal{H}^s$-a.e. $x \in E$, $\underline{D}^s(E,x) = \overline{D}^s(E,x) = 1$. If E is regular, then $dim_{\mathcal{H}}(E)$ equals s and is integral. Regular sets are generalized surfaces, they possess many properties of differentiable manifolds. Let $x \in \mathbb{R}^d$, θ a s-dimensional affine subspace with $x \in \theta$. For $y \in \mathbb{R}^d$ denote the projection of y onto θ by y^θ. Let $\epsilon > 0$ and $S_{\theta,x,\epsilon} = \{y \mid \ \|y - y^\theta\| \leq \epsilon \|y^\theta\|\}$. A set E has a tangent θ at x, if $\overline{D}^s(E,x) > 0$ and for all $\epsilon > 0$

$$\lim_{r\to 0+} \frac{\mathcal{H}^s(E \cap S^c_{\theta,x,\epsilon} \cap B(x,r))}{r^s} = 0. \tag{54}$$

If a tangent exists, it is unique.

LEMMA 5. *1.-3. are equivalent:*

1. *E is regular of dimension s.*

2. *At $\mathcal{H}^s$-a.e. $x \in E$ a tangent to E exists.*

3. *E is rectifiable, i.e. $E = E_0 \cup \cup_{i=1}^{\infty} E_i$ with $\mathcal{H}^s(E_0) = 0$ and for all $i > 0$ the sets E_i are Lipschitz images of closed subsets od $[0,1]^s$, i.e. closed sets $F_i \subseteq [0,1]^s$ and Lipschitz mappings $\phi_i : F_i \to \mathbb{R}^d$ exist, with $E_i = \phi_i(F_i)$.*

For a given $\epsilon > 0$ the sets F_i, E_i and the mappings ϕ_i may be chosen such that

(a) $(E_i)_{i\geq 1}$ are pairwise disjoint and

(b) ϕ_i are C^1-diffeomorphisms with Lipschitz constants of ϕ_i and of ϕ_i^{-1} in $]1-\epsilon, 1+\epsilon[$.

Let $E, F \subseteq \mathbb{R}^d$ be regular s-dimensional Borel sets and $\psi : E \to F$ differentiable and invertible. The Jacobian $J_{\psi,E}$ is the corrective factor between $\mathcal{H}^s_{|E}$ and $\mathcal{H}^s_{|F}$, i.e. for $\mathcal{H}^s_{|E}$-integrable functions g,

$$\int g J_{\psi,E} d\mathcal{H}^s_{|E} = \int g \circ \psi^{-1} d\mathcal{H}^s_{|F}.$$

$J_{\psi,E}$ is the unique nonvanishing determinante of a $s \times s$ submatrix of the approximative derivative of ψ. Let $\psi = f'$ be the gradient of f, let E have a tangent θ_x at x and let U_x denote a $d \times s$-matrix with orthogonal columns, which span $\theta_x - x$. Then $J_{\psi,E}(x) = det(U_x^\top \psi' U_x)$.

LEMMA 6. *Let $E \subseteq \mathbb{R}^d$ regular of dimension s, $P(E) = 1$, $dP/d\mathcal{H}^s_{|E} = p$ and $\psi : E \mapsto F$ bijective and differentiable. Let $\tilde{P}$ denote the law of ψ given P. Then*

$$\frac{d\tilde{P}}{d\mathcal{H}^s_{|F}}(x) = \frac{1}{J_{\psi,E}(\psi^{-1}(x))} p(\psi^{-1}(x)).$$

LEMMA 7. *Let $E \subseteq \mathbb{R}^d$ regular of dimension s, $P(E) = 1$ and $dP/d\mathcal{H}^s_{|E} = p$. Let $\epsilon > 0$, $A : E \mapsto \mathbb{R}^d$ a linear mapping and $\phi : A(E) \mapsto F$ a surjective Lipschitz mapping with*

$$\frac{1}{1+\epsilon}\|x-y\| \leq \|\phi(x) - \phi(y)\| \leq (1+\epsilon)\|x-y\|.$$

Let $\psi = \phi \circ A$ and $\tilde{P}$ the law of ψ (given P). Then $J_{\psi,E}(x) = J_{A,E}(x) J_{\phi,A(E)}(A(x))$ and $(1+\epsilon)^{-s} \leq J_{\phi,A(E)}(A(x)) \leq (1+\epsilon)^s$.

7 References

[1] K. J. FALCONER, *The Geometry of Fractal Sets*, Cambridge University Press, 1985.

[2] K. J. FALCONER, *Techniques in Fractal Geometry*, John Wiley, 1997.

[3] H. FEDERER, *Geometric Measure Theory*, Springer-Verlag, 1969.

[4] B. FLURY, *Principal points*, Biometrica, 77 (1990), pp. 33-41.

[5] S. GRAF and H. LUSCHGY, *The quantization of the Cantor distribution*, Math. Nachrichten, 183 (1997), pp. 113-133.

[6] T. KOHONEN, *Self-organization and associative memory*, Springer-Verlag, 1984.

[7] F. MORGAN, *Geometric Measure Theory*, Academic Press, 1988.

[8] D. POLLARD, *Strong consistency of k-means clustering*, Annals of Statistics, 9 (1981), pp. 135-140.

[9] K. PÖTZELBERGER, *The quantization dimension of distributions*, (1998), submitted.

[10] K. PÖTZELBERGER K. and H. STRASSER, *Data compression by unsupervised classification*, (1997), submitted.

[11] C. TRICOT, *Two definitions of fractal dimensions*, Math. Proc. Camb. Phil. Soc., 91 (1982), pp. 57-74.

[12] P. L. ZADOR, *Development and Evaluation of Procedures for Quantizing Multivariate Distributions*, Ph.D. Thesis, Stanford University, 1964.

[13] P. L. ZADOR, *Asymptotic quantization error of continuous signals and the quantization dimension*, IEEE Trans. Inform. Theory, 28 (1982), pp. 139-149.