The Annals of Probability
2002, Vol. 30, No. 2, 874–897

RATES OF CONVERGENCE FOR THE EMPIRICAL QUANTIZATION ERROR

BY SIEGFRIED GRAF AND HARALD LUSCHGY

University of Passau and University of Trier

For $n, k \in \mathbb{N}$ and $r > 0$ let $e_{n,r}(P_k)^r = \inf \frac{1}{k} \sum_{i=1}^{k} \|X_i - f(X_i)\|^r$, where the infimum is taken over all measurable maps $f : \mathbb{R}^d \to \mathbb{R}^d$ with $|f(\mathbb{R}^d)| \le n$ and $X_1, \ldots, X_k$ are i.i.d. $\mathbb{R}^d$-valued random variables. We analyse the asymptotic a.s. behaviour of the nth empirical quantization error $e_{n,r}(P_k)$.

1. Introduction. One of the central problems of data compression concerns the design of quantizers from empirical data. Then it is essential to evaluate the resulting empirical error and to optimize the quantizers.

From a probabilistic point of view, the framework of quantization can be stated as follows. Let X be a $\mathbb{R}^d$-valued random variable with distribution P. For $n \in \mathbb{N}$, let $\mathcal{F}_n$ be the set of all Borel measurable maps $f : \mathbb{R}^d \to \mathbb{R}^d$ that take at most n values. The elements of $\mathcal{F}_n$ are called n-quantizers. For each $f \in \mathcal{F}_n$, $f(X)$ gives a quantized version of X. We use the L_r-distance as measure for the deviation between X and $f(X)$. Thus if $\|\cdot\|$ denotes any norm on $\mathbb{R}^d$ and $0 < r < \infty$, we define the (minimal) nth quantization error for P by

$$e_{n,r}(P) = \inf\{(E\|X - f(X)\|^r)^{1/r} : f \in \mathcal{F}_n\} \tag{1.1}$$

under the integrability condition $E\|X\|^r < \infty$. The quantization problem is to find an n-optimal quantizer for P and the value of the nth quantization error.

In electrical engineering this problem arises in the context of coding speech and visual signals effectively. For these applications in communication and information theory we refer to Gersho and Gray (1992). In statistics quantizers may be used as models for the grouping of data. Beyond these classical applications, quantization seems to be a promising tool in some recent developments in numerical probability. See Pagès (1997) and Bally and Pagès (2000).

In practice, the underlying distribution P to be quantized is often unknown. Furthermore, even when P is known usual quantizer design algorithms for P like the Lloyd procedures are absolutely irrealistic in dimensions $d \ge 2$ since they involve at each step to compute P-integrals over complicated regions in $\mathbb{R}^d$ [see Gersho and Gray (1992)]. In any case one has to rely on empirical data to obtain good quantizers for P and estimates for $e_{n,r}(P)$.

Received May 2000; revised August 2001.
AMS 2000 subject classifications. Primary 60F15, 60E15; secondary 62H30, 94A29.
Key words and phrases. Multidimensional quantization, L_r-error, empirical measure, empirical quantization error, empirical process.

The purpose of this paper is to analyze the empirical quantization error. Let $X_1, \dots, X_k$ be a sample of k observations on P and let $P_k = k^{-1} \sum_{i=1}^{k} \delta_{X_i}$ denote the empirical measure. Then the nth empirical quantization error is given by

$$(1.2) \qquad e_{n,r}(P_k) = \inf\left\{ \left(\frac{1}{k} \sum_{i=1}^{k} \|X_i - f(X_i)\|^r \right)^{1/r} : f \in \mathcal{F}_n \right\}.$$

By the Glivenko–Cantelli theorem for the L_r-Wasserstein distance between P_k and P,

$$\sup_{n \geq 1} |e_{n,r}(P_k)^r - e_{n,r}(P)^r| \to 0 \qquad \text{a.s., } k \to \infty$$

provided $E\|X_1\|^r < \infty$ (see Section 4). We obtain a.s. rates of convergence for this "uniform strong law of large numbers." This gives some indications of how large the sample size k should be to get a good approximation of $e_{n,r}(P)$ (uniformly in n). The performance of an n-optimal quantizer $f_{n,k}(\cdot, X_1, \dots, X_k)$ for P_k, where $f_{n,k} : (\mathbb{R}^d)^{k+1} \to \mathbb{R}^d$ is assumed to be measurable, as quantizer for P is measured by

$$(1.3) \qquad \hat{e}_{n,r,k}(P) = \left(\int \|x - f_{n,k}(x, X_1, \dots, X_k)\|^r \, dP(x) \right)^{1/r}.$$

The same rates are obtained when $e_{n,r}(P_k)$ is replaced by $\hat{e}_{n,r,k}(P)$. Then we investigate the a.s. behavior of $e_{n,r}(P_k)$ when both n and $k \to \infty$. Upper bounds on probabilities of the form

$$\mathbb{P}\big(|e_{n,r}(P_k)^r - e_{n,r}(P)^r| > t/\sqrt{k}\big)$$

and related forms are given. These inequalities for the concentration of $e_{n,r}(P_k)^r$ around $e_{n,r}(P)^r$ extend those of Rhee and Talagrand (1989). The truncated version

$$(1.4) \qquad e_{n,r}^{(c)}(P) = \inf\big\{ \big(E\big(\|X - f(X)\| \wedge c\big)^r\big)^{1/r} : f \in \mathcal{F}_n \big\}$$

of $e_{n,r}(P)$ will play an interesting role. Furthermore, it is shown that

$$(1.5) \quad \lim_{k \to \infty} n(k)^{1/d} e_{n(k),r}(P_k) = Q_r(d) \left(\int h^{d/(d+r)} \, d\lambda^d \right)^{(d+r)/dr} \qquad \text{a.s.}$$

under $n(k) \to \infty$ and suitable conditions on the ratio $k/n(k)$, where $Q_r(d)$ is a constant depending on d, r and the underlying norm on $\mathbb{R}^d$ only and h is the density of the absolutely continuous part of P. The limiting result (1.5) continues to hold with $e_{n,r}(P_k)$ replaced by $\hat{e}_{n,r,k}(P)$.

The results are for probability measures P with compact support. An extension to measures with unbounded support seems to evoke great technical difficulties and remains open. The rates obtained are, in part, the true ones for nonsingular probability measures P. Continuous singular measures with compact support which are of interest, for example, in connection with fractal models need refinements of our methods. For these measures their quantization dimension [see Graf and Luschgy (2000, 2001)] should play the role of the space dimension d.

The paper is organized as follows. Section 2 contains basic facts for the $e_{n,r}(P)$-problem and the truncated problem. In Section 3 we prove the equality of $e_{n,r}(P)$ and its truncated version for truncation parameters $c = c_n$ which are asymptotically ($n \to \infty$), up to constants, as small as possible for sufficiently regular probability measures P. In particular, we settle a question of Rhee and Talagrand (1989). The main results on the empirical quantization error as outlined above are contained in Sections 4 and 5.

2. Basic properties of the quantization problem. The investigation of the quantization problem requires the concept of Voronoi partitions. If $x \in \mathbb{R}^d$ and A is a nonempty subset of $\mathbb{R}^d$, the distance from x to A is given by

$$d_A(x) = d(x, A) = \inf_{a \in A} \|x - a\|.$$

Consider a nonempty finite subset α of $\mathbb{R}^d$. The Voronoi region generated by $a \in \alpha$ is defined by

$$W(a|\alpha) = \{x \in \mathbb{R}^d : \|x - a\| = d(x, \alpha)\}. \tag{2.1}$$

The Voronoi regions $W(a|\alpha)$ are closed and star-shaped relative to a and the Voronoi diagram $\{W(a|\alpha) : a \in \alpha\}$ of α provides a covering of $\mathbb{R}^d$. A Borel measurable partition $\{A_a : a \in \alpha\}$ of $\mathbb{R}^d$ is called Voronoi partition with respect to α if

$$A_a \subset W(a|\alpha) \qquad \text{for every } a \in \alpha. \tag{2.2}$$

The quantization problem (1.1) can be formulated as an optimal location problem for n-point sets and is further equivalent to the problem of approximating P by a discrete probability with at most n supporting points. For $0 < r < \infty$ and $g : \mathbb{R}^d \to \mathbb{R}$ Borel measurable, let

$$\|g\|_r = \|g\|_{L_r(P)} = \left(\int |g|^r \, dP \right)^{1/r}.$$

For Borel probability measures P_1, P_2 on $\mathbb{R}^d$ with $\int \|x\|^r \, dP_i(x) < \infty$, define the L_r-Wasserstein (L_r-Kantorovich) distance by

$$\rho_r(P_1, P_2) = \inf \left(\int \|x - y\|^r \, d\mu(x, y) \right)^{1/r},$$

where the infimum is taken over all Borel probability measures μ on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals P_1 and P_2. Note that ρ_r is a metric if $r \geq 1$ and ρ_r^r is a metric if $r < 1$.

LEMMA 2.1. *Suppose* $\int \|x\|^r \, dP(x) < \infty$. *Then*

$$e_{n,r}(P) = \inf\{\|d_\alpha\|_r : \alpha \subset \mathbb{R}^d, 1 \le |\alpha| \le n\}$$
$$= \inf\{\rho_r(P, Q) : Q \in \mathcal{P}_n\},$$

where $\mathcal{P}_n$ *denotes the set of all discrete probabilities* Q *on* $\mathbb{R}^d$ *with* $|\operatorname{supp}(Q)| \le n$.

PROOF. See Graf and Luschgy (2000), Lemmas 3.1 and 3.4. □

A set $\alpha \subset \mathbb{R}^d$ with $1 \le |\alpha| \le n$ is called n-optimal set of centers for P of order r if

$$e_{n,r}(P) = \|d_\alpha\|_r.$$

Let $\mathcal{C}_{n,r}(P)$ denote the set of all n-optimal sets of centers. If $\alpha \in \mathcal{C}_{n,r}(P)$, $\{A_a : a \in \alpha\}$ is a Voronoi partition of $\mathbb{R}^d$ with respect to α and $f = \sum_{a \in \alpha} a \mathbb{1}_{A_a}$, then

$$e_{n,r}(P) = \bigl(E\|X - f(X)\|^r\bigr)^{1/r}$$

and hence the nearest neighbor quantizer f of α is an n-optimal quantizer. Conversely, if f is n-optimal then $f(\mathbb{R}^d) \in \mathcal{C}_{n,r}(P)$.

Under truncation we have

$$e_{n,r}^{(c)}(P) = \inf\{\|d_\alpha \wedge c\|_r : \alpha \subset \mathbb{R}^d, 1 \le |\alpha| \le n\}, \qquad 0 < c \le \infty. \tag{2.3}$$

Clearly, $e_{n,r} = e_{n,r}^{(\infty)}$. Let $\mathcal{C}_{n,r}^{(c)}(P)$ denote the set of all n-optimal sets of centers $\alpha \subset \mathbb{R}^d$ for the truncated problem, that is, $1 \le |\alpha| \le n$ and

$$e_{n,r}^{(c)}(P) = \|d_\alpha \wedge c\|_r.$$

Recall that for nonempty compact subsets A, B of $\mathbb{R}^d$ the Hausdorff metric is given by

$$H(A, B) = \max\Bigl\{\max_{a \in A} d(a, B), \max_{b \in B} d(b, A)\Bigr\}$$
$$= \sup_{x \in \mathbb{R}^d} |d(x, A) - d(x, B)|.$$

The closed (open) ball with center $a \in \mathbb{R}^d$ and radius $s \ge 0$ is denoted by $B(a, s) = \{x \in \mathbb{R}^d : \|x - a\| \le s\}$ $[\mathring{B}(a, s) = \{x \in \mathbb{R}^d : \|x - a\| < s\}]$.

PROPOSITION 2.2. *Let* $0 < c \le \infty$. *Suppose* $|\operatorname{supp}(P)| \ge n$ *and if* $c = \infty$, *suppose further* $\int \|x\|^r \, dP(x) < \infty$.

(a) *Let* $\alpha \in \mathcal{C}_{n,r}^{(c)}(P)$ *and let* $\{A_a : a \in \alpha\}$ *be a Voronoi partition of* $\mathbb{R}^d$ *with respect to* α. *Then*

$$|\alpha| = n, \ P(A_a) > 0 \qquad \textit{for every } a \in \alpha,$$

$$\beta \in \mathcal{C}_{m,r}^{(c)}\Biggl(P\Biggl(\cdot \Bigm| \bigcup_{a \in \beta} A_a\Biggr)\Biggr) \qquad \textit{for every } \beta \subset \alpha \textit{ with } |\beta| = m.$$

(b) *We have* $e^{(c)}_{n,r}(P) < e^{(c)}_{n-1,r}(P)$ $[e^{(c)}_{0,r}(P) := c]$ *and* $\mathcal{C}^{(c)}_{n,r}(P)$ *is nonempty and* H*-compact*.

PROOF. For the nontruncation case $c = \infty$, see Graf and Luschgy (2000), Theorems 4.1 and 4.12. So we assume $c < \infty$. Set $e^{(c)}_{n,r} = e^{(c)}_{n,r}(P)$.

(a) Let $\gamma = \{a \in \alpha : P(A_a) > 0\}$. Obviously we have $\gamma \in \mathcal{C}^{(c)}_{n,r}(P)$. Assume $|\gamma| < n$. Then there exists $a \in \gamma$ such that $P(A_a \setminus \{a\}) > 0$. Let $K \subset A_a \setminus \{a\}$ be compact with $P(K) > 0$. The open sets

$$O_b = \{x \in \mathbb{R}^d : \|x - a\| > \|x - b\|\} \cap \mathring{B}(b, c/2), \qquad b \in K,$$

provide a covering of K. Therefore, one can find a finite subset B of K such that

$$K \subset \bigcup_{b \in B} O_b \cap A_a.$$

Thus there exists $b \in \mathbb{R}^d$ such that

$$P(O_b \cap A_a) > 0.$$

It follows that

$$\begin{aligned}
(e^{(c)}_{n,r})^r &= \int (d_\gamma \wedge c)^r \, dP \\
&= \sum_{u \in \gamma \setminus \{a\}} \int_{A_u} (d_\gamma \wedge c)^r \, dP + \int_{A_a} (\|x - a\| \wedge c)^r \, dP(x) \\
&> \sum_{u \in \gamma \setminus \{a\}} \int_{A_u} (d_\gamma \wedge c)^r \, dP + \int_{A_a \cap O_b^c} (\|x - a\| \wedge c)^r \, dP(x) \\
&\quad + \int_{A_a \cap O_b} (\|x - b\| \wedge c)^r \, dP(x) \\
&\geq \int (d_{\gamma \cup \{b\}} \wedge c)^r \, dP \geq (e^{(c)}_{n,r})^r,
\end{aligned}$$

a contradiction.

As for the assertion concerning β, assume $\beta \notin \mathcal{C}^{(c)}_{m,r}(P(\cdot | \bigcup_{a \in \beta} A_a))$. Then there exists $\delta \subset \mathbb{R}^d$ with $1 \leq |\delta| \leq n$ and

$$\int_{\bigcup_{a \in \beta} A_a} (d_\beta \wedge c)^r \, dP > \int_{\bigcup_{a \in \beta} A_a} (d_\delta \wedge c)^r \, dP.$$

It follows that

$$\begin{aligned}
(e^{(c)}_{n,r})^r &= \int (d_\alpha \wedge c)^r \, dP \\
&> \int (d_{(\alpha \setminus \beta) \cup \delta} \wedge c)^r \, dP \geq (e^{(c)}_{n,r})^r,
\end{aligned}$$

a contradiction.

(b) *Step* 1. In the first step we assume $e_{n,r}^{(c)} < e_{n-1,r}^{(c)}$. Let $e_{n,r}^{(c)} \le b < e_{n-1,r}^{(c)}$. Since $e_{n-1,r}^{(c)} \le c$, we can find $s > 0$ such that

$$c^r P\big(B(0,s)\big) > b^r.$$

Let $S \ge s + c$ such that

$$c^r P\big(B(0,2S)^c\big) < (e_{n-1,r}^{(c)})^r - b^r.$$

(Note that s and S depend on P, r and n.) Consider the level set

$$L(b) = \big\{\alpha \subset \mathbb{R}^d : 1 \le |\alpha| \le n, \|d_\alpha \wedge c\|_r \le b\big\}.$$

Let $\alpha \in L(b)$. Then $|\alpha| = n$. Let $\alpha = \{a_1, \dots, a_n\}$ and assume without loss of generality $\|a_1\| \le \dots \le \|a_n\|$. Then $\|a_1\| \le S$. Otherwise

$$\begin{aligned} b^r &\ge \int_{B(0,s)} (d_\alpha \wedge c)^r \, dP \ge \big((S-s) \wedge c\big)^r P\big(B(0,s)\big) \\ &= c^r P\big(B(0,s)\big), \end{aligned}$$

a contradiction. Furthermore, $\|a_n\| \le 5S$ (assuming now $n \ge 2$). Otherwise

$$\|x - a_1\| \le \|x - a_n\| \mathbb{1}_{B(0,2S)}(x) + 2\|x\| \mathbb{1}_{B(0,2S)^c}(x)$$

for every $x \in \mathbb{R}^d$. If $\{A_1, \dots, A_n\}$ with $A_i = A_{a_i}$ denotes a Voronoi partition of $\mathbb{R}^d$ with respect to α and $\beta = \{a_1, \dots, a_{n-1}\}$, then

$$\begin{aligned} (e_{n-1,r}^{(c)})^r &\le \int (d_\beta \wedge c)^r \, dP \\ &\le \sum_{j=1}^{n-1} \int_{A_j} (\|x - a_j\| \wedge c)^r \, dP(x) + \int_{A_n} (\|x - a_1\| \wedge c)^r \, dP(x) \\ &\le \sum_{j=1}^{n-1} \int_{A_j} (\|x - a_j\| \wedge c)^r \, dP(x) + \int\limits_{A_n \cap B(0,2S)} (\|x - a_n\| \wedge c)^r \, dP(x) \\ &\quad + c^r P\big(A_n \cap B(0,2S)^c\big) \\ &\le \sum_{j=1}^{n} \int_{A_j} (\|x - a_j\| \wedge c)^r \, dP(x) + c^r P\big(B(0,2S)^c\big) \\ &< \int (d_\alpha \wedge c)^r \, dP + (e_{n-1,r}^{(c)})^r - b^r \\ &\le (e_{n-1,r}^{(c)})^r, \end{aligned}$$

a contradiction. We thus obtain

$$L(b) \subset \{\alpha \subset B(0,5S) : 1 \le |\alpha| \le n\}.$$

Since the latter set is H-compact and $\alpha \to \|d_\alpha \wedge c\|_r$ is H-continuous, the level set $L(b)$ is H-compact, too. This implies that $\mathfrak{C}_{n,r}^{(c)}(P) = L(e_{n,r}^{(c)})$ is H-compact.

Step 2. In the second step we prove $e_{n,r}^{(c)} < e_{n-1,r}^{(c)}$. We proceed inductively. Clearly, we have $e_{1,r}^{(c)} < c = e_{0,r}^{(c)}$. If $e_{m,r}^{(c)} < e_{m-1,r}^{(c)}$ for some $2 \le m \le n-1$, then $\mathcal{C}_{m,r}^{(c)}(P) \neq \emptyset$ by the first step. Therefore, $e_{m+1,r}^{(c)} < e_{m,r}^{(c)}$, since otherwise $\mathcal{C}_{m,r}^{(c)}(P) \subset \mathcal{C}_{m+1,r}^{(c)}(P)$ which contradicts part (a). $\square$

The diameter of a nonempty bounded subset A of $\mathbb{R}^d$ is the number

$$\operatorname{diam}(A) = \sup\{\|a-b\| : a, b \in A\}.$$

For $A \subset \mathbb{R}^d$ and $\varepsilon \ge 0$, let $A^\varepsilon = \{x \in \mathbb{R}^d : d(x, A) \le \varepsilon\}$. The following lemma provides a further necessary condition for n-optimality in case $\operatorname{supp}(P)$ is compact.

LEMMA 2.3. *Let $A \subset \mathbb{R}^d$ be a compact set and $0 < c \le \infty$. Suppose* $\operatorname{supp}(P) \subset A$ *and* $|\operatorname{supp}(P)| \ge n$. *If $\alpha \in \mathcal{C}_{n,r}^{(c)}(P)$, then $\alpha \subset A^D$, where $D =$* $\operatorname{diam}(A)$.

PROOF. By Proposition 2.2(a), we have $P(W(a|\alpha)) > 0$ and $a \in \mathcal{C}_{1,r}^{(c)}(P(\cdot|W(a|\alpha)))$ for all $a \in \alpha$. Assume $d(a, A \cap W(a|\alpha)) > \operatorname{diam}(A \cap W(a|\alpha))$ for some $a \in \alpha$. Let $y \in \operatorname{supp}(P(\cdot|W(a|\alpha))$. Then $P(W(a|\alpha) \cap \mathring{B}(y, c)) > 0$ and $\|x - y\| < \|x - a\|$ for all $x \in A \cap W(a|\alpha)$. Hence

$$\begin{aligned}
&\int_{W(a|\alpha)} (\|x-y\| \wedge c)^r \, dP(x) \\
&\qquad = \int_{W(a|\alpha) \cap \mathring{B}(y,c)} \|x-y\|^r \, dP(x) + c^r P\big(W(a|\alpha) \cap \mathring{B}(y,c)^c\big) \\
&\qquad < \int_{W(a|\alpha) \cap \mathring{B}(y,c)} (\|x-a\| \wedge c)^r \, dP(x) + c^r P\big(W(a|\alpha) \cap \mathring{B}(y,c)^c\big) \\
&\qquad = \int_{W(a|\alpha)} (\|x-a\| \wedge c)^r \, dP(x),
\end{aligned}$$

a contradiction. This proves the lemma. $\square$

Next we give bounds for $|e_{n,r}(P)^r - e_{n,r}(Q)^r|$. For a compact set $A \subset \mathbb{R}^d$, let

$$\mathcal{G}_{n,r} = \mathcal{G}_{n,r}(A) = \{d_\alpha^r : \alpha \subset A^D, 1 \le |\alpha| \le n\} \tag{2.4}$$

and

$$\mathcal{B}_n = \mathcal{B}_n(A) = \left\{ \bigcup_{a \in \alpha} B(a, s) : \alpha \subset A^D,\ 1 \le |\alpha| \le n,\ s \le 2D \right\}, \tag{2.5}$$

where $D = \operatorname{diam}(A)$. For $\mathcal{G} \subset L_1(P) \cap L_1(Q)$, let

$$\|P - Q\|_{\mathcal{G}} = \sup\left\{ \left| \int g \, dP - \int g \, dQ \right| : g \in \mathcal{G} \right\}.$$

LEMMA 2.4. *Let* $A = \text{supp}(P)$ *and suppose* A *is compact. Let* Q *be a probability measure on* $\mathbb{R}^d$ *with* $\text{supp}(Q) \subset A$. *Then*

$$|e_{n,r}(P)^r - e_{n,r}(Q)^r| \le \|P - Q\|_{\mathcal{G}_{n,r}} \le (2D)^r \|P - Q\|_{\mathcal{B}_n}.$$

Furthermore,

$$\sup_{n\ge 1} |e_{n,r}(P)^r - e_{n,r}(Q)^r| \le r(2D)^{r-1}\rho_1(P, Q) \qquad \text{if } r \ge 1.$$

PROOF. For $n \in \mathbb{N}$, let

$$\mathcal{A}_n = \{\alpha \subset A^D : 1 \le |\alpha| \le n\}.$$

Choose $\beta \in \mathcal{C}_{n,r}(P) \cap \mathcal{A}_n$. If $|A| \ge n$, this is possible in view of Lemma 2.3. If $|A| < n$, choose $\beta = A$. One obtains

$$\begin{aligned} e_{n,r}(Q)^r - e_{n,r}(P)^r &\le \int d_\beta^r \, dQ - \int d_\beta^r \, dP \\ &\le \sup\left\{\left|\int d_\alpha^r \, dQ - \int d_\alpha^r \, dP\right| : \alpha \in \mathcal{A}_n\right\}. \end{aligned}$$

Now choose $\gamma \in \mathcal{C}_{n,r}(Q) \cap \mathcal{A}_n$. If $|\text{supp}(Q)| \ge n$, this is possible again by Lemma 2.3. If $|\text{supp}(Q)| < n$, choose $\gamma = \text{supp}(Q)$. One gets

$$\begin{aligned} e_{n,r}(P)^r - e_{n,r}(Q)^r &\le \int d_\gamma^r \, dP - \int d_\gamma^r \, dQ \\ &\le \sup\left\{\left|\int d_\alpha^r \, dP - \int d_\alpha^r \, dQ\right| : \alpha \in \mathcal{A}_n\right\}. \end{aligned}$$

This gives

$$|e_{n,r}(P)^r - e_{n,r}(Q)^r| \le \|P - Q\|_{\mathcal{G}_{n,r}}.$$

Let $\alpha \in \mathcal{A}_n$. Then $\max_{x\in A} d_\alpha(x) \le 2D$ and hence

$$\begin{aligned} \left|\int d_\alpha^r \, dP - \int d_\alpha^r \, dQ\right| &= r\left|\int_0^{2D} P(d_\alpha > t)t^{r-1}\, dt - \int_0^{2D} Q(d_\alpha > t)t^{r-1}\, dt\right| \\ &\le r\int_0^{2D} |P(d_\alpha \le t) - Q(d_\alpha \le t)| t^{r-1}\, dt \\ &\le r\|P - Q\|_{\mathcal{B}_n} \int_0^{2D} t^{r-1}\, dt \\ &= (2D)^r \|P - Q\|_{\mathcal{B}_n}. \end{aligned}$$

Thus

$$\|P - Q\|_{\mathcal{G}_{n,r}} \le (2D)^r \|P - Q\|_{\mathcal{B}_n}.$$

Furthermore, let μ be a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals P and Q. Since

$$|u^r - v^r| \le r \max\{u^{r-1}, v^{r-1}\}|u - v| \qquad \text{if } r \ge 1,\ u, v \ge 0,$$

we have

$$\begin{aligned}\left|\int d_\alpha^r \, dP - \int d_\alpha^r \, dQ\right| &= \left|\int_{A\times A} \big(d_\alpha^r(x) - d_\alpha^r(y)\big) \, d\mu(x, y)\right| \\ &\le \int_{A\times A} |d_\alpha^r(x) - d_\alpha^r(y)| \, d\mu(x, y) \\ &\le r(2D)^{r-1} \int |d_\alpha(x) - d_\alpha(y)| \, d\mu(x, y) \\ &\le r(2D)^{r-1} \int \|x - y\| \, d\mu(x, y).\end{aligned}$$

This implies

$$\|P - Q\|_{\mathcal{G}_{n,r}} \le r(2D)^{r-1} \rho_1(P, Q) \qquad \text{if } r \ge 1. \qquad \square$$

In the sequel the quantization problem with respect to the L_∞-distance serves as upper bound for the L_r-problem. For $n \in \mathbb{N}$ and a probability P with compact support, let

$$e_{n,\infty}(P) = \inf\{\operatorname{ess\,sup} \|X - f(X)\| : f \in \mathcal{F}_n\}.$$

For $g : \mathbb{R}^d \to \mathbb{R}$ Borel measurable, set

$$\|g\|_\infty = \|g\|_{L_\infty(P)} = \inf\{c \ge 0 : |g| \le c \ P\text{-a.s.}\}.$$

It is easy to check that

$$e_{n,\infty}(P) = \inf\{\|d_\alpha\|_\infty : \alpha \subset \mathbb{R}^d,\ 1 \le |\alpha| \le n\}. \tag{2.6}$$

Since $\|d_\alpha\|_\infty = \max\{d_\alpha(x) : x \in \operatorname{supp}(P)\}$, $e_{n,\infty}(P)$ depends only on the support of P. For a nonempty compact set $A \subset \mathbb{R}^d$, define

$$e_{n,\infty}(A) = \inf\Big\{\max_{x\in A} d_\alpha(x) : \alpha \subset \mathbb{R}^d,\ 1 \le |\alpha| \le n\Big\} \tag{2.7}$$

and let $\mathcal{C}_{n,\infty}(A)$ denote the set of all sets $\alpha \subset \mathbb{R}^d$ with $1 \le |\alpha| \le n$ for which the infimum in (2.7) is attained. For $m \ge 0$ and $\beta \in \mathcal{C}_{m,\infty}(A)$ ($\beta = \varnothing$ if $m = 0$) let

$$\mathcal{G}_{n,m,r} = \mathcal{G}_{n,m,r}(A, \beta) = \big\{d_{\alpha\cup\beta}^r : \alpha \subset A^D, 1 \le |\alpha| \le n\big\}. \tag{2.8}$$

LEMMA 2.5. *Let $m \ge 0$. In the situation of Lemma 2.4 we have*

$$e_{n+m,r}(P)^r - e_{n,r}(Q)^r \le \|P - Q\|_{\mathcal{G}_{n,m,r}}.$$

PROOF. Choose $\alpha \in \mathcal{C}_{n,r}(Q)$ such that $\alpha \subset A^D$. Then

$$\begin{aligned}
e_{n+m,r}(P)^r - e_{n,r}(Q)^r &\le \int d_{\alpha\cup\beta}^r \, dP - \int d_\alpha^r \, dQ \\
&\le \int d_{\alpha\cup\beta}^r \, dP - \int d_{\alpha\cup\beta}^r \, dQ \\
&\le \|P - Q\|_{\mathcal{G}_{n,m,r}}. \qquad \square
\end{aligned}$$

Having in mind the definition of $\hat{e}_{n,r,k}(P)$ [see (1.3)], the following bounds for the loss of a quantizer mismatch are useful.

LEMMA 2.6. *Suppose $\int \|x\|^r \, dP(x) < \infty$ and let Q be a probability measure on $\mathbb{R}^d$ with $\int \|x\|^r \, dQ(x) < \infty$. Let $\alpha \in \mathcal{C}_{n,r}(Q)$. Then*

$$0 \le \left(\int d_\alpha^r \, dP\right)^{1/r} - e_{n,r}(P) \le 2\rho_r(P, Q) \qquad \text{if } r \ge 1.$$

In the situation of Lemma 2.4 *and if $\alpha \subset A^D$, then we have*

$$\int d_\alpha^r \, dP - e_{n,r}(P)^r \le 2\|P - Q\|_{\mathcal{G}_{n,r}}.$$

PROOF. We have

$$\begin{aligned}
0 &\le \left(\int d_\alpha^r \, dP\right)^{1/r} - e_{n,r}(P) \\
&\le \left|\left(\int d_\alpha^r \, dP\right)^{1/r} - \left(\int d_\alpha^r \, dQ\right)^{1/r}\right| + |e_{n,r}(P) - e_{n,r}(Q)|.
\end{aligned}$$

Lemma 2.1 shows that $|e_{n,r}(P) - e_{n,r}(Q)| \le \rho_r(P, Q)$. The Minkowski inequality implies that, for every probability measure μ on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals P and Q,

$$\begin{aligned}
\left|\left(\int d_\alpha^r \, dP\right)^{1/r} - \left(\int d_\alpha^r \, dQ\right)^{1/r}\right| &\le \left(\int |d_\alpha(x) - d_\alpha(y)|^r \, d\mu(x, y)\right)^{1/r} \\
&\le \left(\int \|x - y\|^r \, d\mu(x, y)\right)^{1/r}.
\end{aligned}$$

Hence

$$\left|\left(\int d_\alpha^r \, dP\right)^{1/r} - \left(\int d_\alpha^r \, dQ\right)^{1/r}\right| \le \rho_r(P, Q)$$

and thus the first assertion.

The second assertion easily follows from Lemma 2.4. $\square$

The main result concerning the asymptotics of the nth quantization error as $n \to \infty$ reads as follows. Let P_a denote the absolutely continuous part of P and let $U([0, 1]^d)$ denote the uniform distribution on the unit cube.

THEOREM 2.7. *Suppose* $\int \|x\|^{r+\delta}\, dP(x) < \infty$ *for some* $\delta > 0$. *Let* $Q_r(d) = \inf_{n\geq 1} n^{1/d} e_{n,r}(U([0,1]^d))$. *Then* $Q_r(d) > 0$ *and*

$$\lim_{n\to\infty} n^{1/d} e_{n,r}(P) = Q_r(d)\left(\int h^{d/(d+r)}\, d\lambda^d\right)^{(d+r)/dr} < \infty,$$

where $h = dP_a/d\lambda^d$.

PROOF. See Graf and Luschgy (2000), Theorem 6.2. $\square$

Little is known about the true value of the constant $Q_r(d)$ except for the l_∞-norm where $Q_r(d) = \frac{1}{2}(\frac{d}{d+r})^{1/r}$. Some geometric considerations lead to

$$Q_1(2) = \frac{2 + 3\log(\sqrt{3})}{3^{7/4}\sqrt{2}}, \qquad Q_2(2) = \left(\frac{5}{18\sqrt{3}}\right)^{1/2}$$

for the l_2-norm and $Q_r(2) = \frac{1}{\sqrt{2}}(\frac{2}{2+r})^{1/r}$ for the l_1-norm [see Graf and Luschgy (2000), Section 8].

3. Truncation. Here we discuss the equality of $e_{n,r}$ and its truncated version for small truncation parameters. Under this equality we improve the bound of Lemma 2.4.

We will need the following lemma. It provides a generalization of Lemma 2.2 in Hochbaum and Steele (1982).

LEMMA 3.1. *Let* $A \subset \mathbb{R}^d$ *be compact*, $\operatorname{supp}(P) \subset A$ *and* $0 < c \leq \infty$. *There exists a constant* $C < \infty$ *depending only on* A *and* r (*and the underlying norm but not on* c) *such that*

$$e_{n,r}^{(c)}(P)^r - e_{n+1,r}^{(c)}(P)^r \leq C n^{-1-r/d}.$$

PROOF. Let $e_{n,r}^{(c)} = e_{n,r}^{(c)}(P)$. We may assume without loss of generality that $n \geq 2$ and $|\operatorname{supp}(P)| \geq n+1$. Let $\alpha \in \mathcal{C}_{n+1,r}^{(c)}(P)$ and let $\{A_a : a \in \alpha\}$ be a Voronoi partition of $\mathbb{R}^d$ with respect to α. Then

$$(e_{n+1,r}^{(c)})^r = \sum_{a\in\alpha} \int_{A_a} (\|x-a\| \wedge c)^r\, dP(x),$$

$P(A_a) > 0$ for all $a \in \alpha$ and $|\alpha| = n+1$ [see Proposition 2.2]. Note that

$$\left|\left\{a \in \alpha : \int_{A_a} (\|x-a\| \wedge c)^r\, dP(x) > \frac{4(e_{n+1,r}^{(c)})^r}{n+1}\right\}\right| \leq \frac{n+1}{4}$$

and

$$\left|\left\{a \in \alpha : P(A_a) \geq \frac{4}{n+1}\right\}\right| \leq \frac{n+1}{4}.$$

This implies that

$$\beta = \left\{a \in \alpha : \int_{A_a} (\|x-a\| \wedge c)^r \, dP(x) \leq \frac{4(e_{n+1,r}^{(c)})^r}{n+1} \text{ and } P(A_a) \leq \frac{4}{n+1}\right\}$$

satisfies $|\beta| \geq (n+1)/2$. Since $\beta \subset A^D$ with $D = \operatorname{diam}(A)$ (see Lemma 2.3), there exists $a_1, a_2 \in \beta$, $a_1 \neq a_2$ such that

$$\|a_1 - a_2\| \leq C_1 n^{-1/d}$$

for some constant C_1 depending on A. If $\gamma = \alpha \setminus \{a_1\}$, then

$$\begin{aligned}(e_{n,r}^{(c)})^r &\leq \int (d\gamma \wedge c)^r \, dP = \sum_{a\in\gamma} \int_{A_a} (\|x-a\| \wedge c)^r \, dP(x) + \int_{A_{a_1}} (d_\gamma \wedge c)^r \, dP \\ &= (e_{n+1,r}^{(c)})^r + \int_{A_{a_1}} [(d_\gamma \wedge c)^r - (\|x-a_1\| \wedge c)^r] \, dP(x).\end{aligned}$$

Since

$$d_\gamma(x) \leq \|x-a_2\| \leq \|x-a_1\| + \|a_1-a_2\|$$

and hence

$$\begin{aligned}d_\gamma(x) \wedge c &\leq (\|x-a_1\| + \|a_1-a_2\|) \wedge c \\ &\leq \|x-a_1\| \wedge c + \|a_1-a_2\| \wedge c \\ &\leq \|x-a_1\| \wedge c + \|a_1-a_2\|\end{aligned}$$

we deduce

$$\begin{aligned}&(e_{n,r}^{(c)})^r - (e_{n+1,r}^{(c)})^r \\ &\quad \leq (2^r-1) \int_{A_{a_1}} (\|x-a_1\| \wedge c)^r \, dP(x) + 2^r \|a_1-a_2\|^r P(A_{a_1}) \\ &\quad \leq \frac{4(2^r-1)(e_{n+1,r}^{(c)})^r}{n+1} + \frac{4 \cdot 2^r C_1^r n^{-r/d}}{n+1}.\end{aligned}$$

Note that

$$e_{n+1,r}^{(c)} \leq e_{n+1,\infty}(A) \leq C_2 n^{-1/d}$$

for some constant C_2 depending on A. This yields the lemma. □

For the rest of this section let $A = \operatorname{supp}(P)$, $D = \operatorname{diam}(A)$ and suppose A is compact.

LEMMA 3.2.

(a) *Let* $c > 0$. *The following assertions are equivalent.*

(i) $e_{n,r}(P) = e_{n,r}^{(c)}(P)$.

(ii) $\mathcal{C}_{n,r}(P) \subset \mathcal{C}_{n,r}^{(c)}(P)$ *and* $\sup\{\|d_\alpha\|_\infty : \alpha \in \mathcal{C}_{n,r}(P)\} \le c$.

(iii) *There exists* $\alpha \in \mathcal{C}_{n,r}^{(c)}(P)$ *such that* $\|d_\alpha\|_\infty \le c$.

(b) $e_{n,r}(P) = e_{n,r}^{(c)}(P)$ *for all* $c \ge 2D$.

The simple proof is given for the readers convenience.

PROOF. (a) (i) $\Rightarrow$ (ii). Let $\alpha \in \mathcal{C}_{n,r}(P)$. Then

$$e_{n,r}(P)^r = \int d_\alpha^r \, dP \ge \int (d_\alpha \wedge c)^r \, dP \ge e_{n,r}^{(c)}(P)^r = e_{n,r}(P)^r.$$

Hence $\alpha \in \mathcal{C}_{n,r}^{(c)}(P)$ and $d_\alpha = d_\alpha \wedge c$ P-a.s. This gives $\|d_\alpha\|_\infty \le c$.

(ii) $\Rightarrow$ (iii) is obvious since $\mathcal{C}_{n,r}(P) \neq \varnothing$.

(iii) $\Rightarrow$ (i). We have

$$e_{n,r}^{(c)}(P)^r = \int (d_\alpha \wedge c)^r \, dP = \int d_\alpha^r \, dP \ge e_{n,r}(P)^r.$$

(b) We may assume $|A| \ge n$. Let $\alpha \in \mathcal{C}_{n,r}^{(c)}(P)$. Then by Lemma 2.3, $\alpha \subset A^D$. Hence $\|d_\alpha\|_\infty \le 2D \le c$ and the assertion follows from (a). $\square$

Lemma 3.2 shows that $c = e_{n,\infty}(A)$ is the smallest possible value for $e_{n,r}(P) = e_{n,r}^{(c)}(P)$ to hold. Note that the order of convergence to zero of $e_{n,\infty}(A)$ is $n^{-1/d}$ if $\lambda^d(A) > 0$. Here λ^d denotes the d-dimensional Lebesgue measure.

PROPOSITION 3.3. *Suppose there is a constant* $C_1 > 0$ *such that*

$$P\big(B(a,s)\big) \ge C_1 s^d \qquad \text{for all } a \in A \text{ and } 0 < s \le D. \tag{3.1}$$

Then there exists a constant $C_2 < \infty$ *depending on* C_1, A *and* r (*and the underlying norm*) *such that* $c_n = C_2 n^{-1/d}$ *satisfies*

$$e_{n,r}(P) = e_{n,r}^{(c_n)}(P).$$

PROOF. Let the truncation parameter $c = c_n > 0$ be unspecified and let $\alpha = \alpha_n \in \mathcal{C}_{n,r}^{(c)}(P)$ with $\alpha \subset A^D$ (see Lemma 2.3). Set

$$\delta = \delta_n = \max_{x \in A} d_\alpha(x).$$

Choose $y = y_n \in A$ and $a = a_n \in \alpha$ such that $\delta = d_\alpha(y) = \|y - a\|$. For every $b \in \alpha$ we have $\|y - b\| \ge \|y - a\|$. For every $x \in B(y, \delta/2)$ and $b \in \alpha$ this yields

$$\|x - b\| \ge \|y - b\| - \|x - y\| \ge \|y - a\| - \|x - y\| \ge \delta/2$$

and hence

$$d_\alpha(x) \ge \delta/2, \qquad x \in B(y, \delta/2).$$

Let $\beta = \beta_n = \alpha \cup \{y\}$. Then

$$d_\beta(x) = \|x - y\|, \qquad x \in B(y, \delta/2).$$

Using condition (3.1) and $\delta \le 2D$ we deduce

$$\begin{aligned}
e_{n,r}^{(c)}(P)^r &- e_{n+1,r}^{(c)}(P)^r \\
&\ge \int (d_\alpha \wedge c)^r \, dP - \int (d_\beta \wedge c)^r \, dP \\
&\ge \int_{B(y,\delta/2)} [(d_\alpha \wedge c)^r - (d_\beta \wedge c)^r] \, dP \\
&\ge \big((\delta/2) \wedge c\big)^r P\big(B(y,\delta/2)\big) - \int_{B(y,\delta/2)} (\|x-y\| \wedge c)^r \, dP(x) \\
&= r \int_0^{(\delta/2)\wedge c} P\big(B(y,s)\big) s^{r-1} \, ds \\
&\ge C_1 r \int_0^{(\delta/2)\wedge c} s^{d+r-1} \, ds \\
&= \frac{C_1 r}{d+r} \big((\delta/2) \wedge c\big)^{d+r}.
\end{aligned}$$

By Lemma 3.1, this yields

$$\left(\frac{\delta}{2} \wedge c\right)^{d+r} \le \frac{C(d+r)}{C_1 r} n^{-(d+r)/d}$$

and thus

$$\frac{\delta}{2} \wedge c \le K n^{-1/d},$$

where $K = (C(d+r)/C_1 r)^{1/(d+r)}$. Now let $C_2 = 2K$ and $c = c_n = C_2 n^{-1/d}$. Then $\delta \le c$. Thus the assertion follows from Lemma 3.2(a). $\square$

Notice that the preceding proposition implies, for $\alpha_n \in \mathcal{C}_{n,r}(P)$,

$$H(\alpha_n, A) \le C_2 n^{-1/d}. \tag{3.2}$$

In fact, by Lemma 3.2(a) we have

$$\max_{x \in A} d_{\alpha_n}(x) \le C_2 n^{-1/d}.$$

If $a \in \alpha_n$, then $A \cap W(a|\alpha_n) \ne \varnothing$ (cf. Proposition 2.2). Taking $y \in A \cap W(a|\alpha_n)$ one gets

$$d(a, A) \le \|a - y\| = d_{\alpha_n}(y) \le \max_{x \in A} d_{\alpha_n}(x).$$

The following corollary settles a question of Rhee and Talagrand (1989), Remark, page 195. It extends their Proposition C to arbitrary dimensions and arbitrary norms (and any $r > 0$, and more general supports). The proof of Rhee and Talagrand for the case $d = 2$ and the l_2-norm on $\mathbb{R}^2$ is purely geometric and breaks down for higher dimensions and non-Euclidean norms.

COROLLARY 3.4. *Let $P = h\lambda^d$. Suppose the support A of P is a finite union of compact convex sets and $h \geq u > 0$ λ^d-a.s. on A. Then there exists a constant $C < \infty$ depending on u, A and r (and the underlying norm) such that $c_n = Cn^{-1/d}$ satisfies*

$$e_{n,r}(P) = e_{n,r}^{(c_n)}(P).$$

PROOF. We have

$$P\big(B(a,s)\big) \geq u\lambda^d\big(B(a,s) \cap A\big).$$

Let $A = \bigcup_{j=1}^m A_j$ with compact convex sets $A_j \subset \mathbb{R}^d$ and notice that $\lambda^d(A_j) > 0$ for every j. By a simple geometric argument,

$$\lambda^d\big(B(a,s) \cap A_j\big) \geq K_j s^d$$

for all $a \in A_j$, $0 < s \leq D$ and some constant $K_j > 0$ depending on A_j [cf. Graf and Luschgy (2000), Example 12.7].This implies for $a \in A_i$

$$\lambda^d\big(B(a,s) \cap A\big) \geq \lambda^d\big(B(a,s) \cap A_i\big) \geq K_i s^d \geq \min_{1 \leq j \leq m} K_j s^d.$$

Thus condition (3.1) is satisfied which gives the assertion. $\square$

Note that condition (3.1) is satisfied for various singular distributions P like Hausdorff measures on surfaces of convex sets, on compact differentiable manifolds or self-similar sets [cf. Graf and Luschgy (2000), Section 12].

Under the condition $e_{n,r} = e_{n,r}^{(c)}$, the bound of Lemma 2.4 can be improved. Let

$$\mathcal{G}_{n,r}^{(c)} = \mathcal{G}_{n,r}^{(c)}(A) = \big\{(d_\alpha \wedge c)^r : \alpha \subset A^D, 1 \leq |\alpha| \leq n\big\}. \tag{3.3}$$

LEMMA 3.5. *Suppose $e_{n,r}(P) = e_{n,r}^{(c)}(P)$. Let Q be a probability measure on $\mathbb{R}^d$ with* supp$(Q) \subset A$. *Then*

$$|e_{n,r}(P)^r - e_{n,r}(Q)^r| \leq \|P - Q\|_{\mathcal{G}_{n,r}^{(c)}}.$$

PROOF. Let

$$\mathcal{A}_n = \big\{\alpha \subset A^D : 1 \leq |\alpha| \leq n\big\}.$$

Choose $\beta \in \mathcal{C}_{n,r}(P) \cap \mathcal{A}_n$ (cf. the proof of Lemma 2.4). Then by Lemma 3.2(a), $\max_{x \in A} d_\beta(x) \le c$ and hence

$$\begin{aligned} e_{n,r}(Q)^r - e_{n,r}(P)^r &\le \int d_\beta^r \, dQ - \int d_\beta^r \, dP \\ &= \int (d_\beta \wedge c)^r \, dQ - \int (d_\beta \wedge c)^r \, dP \\ &\le \sup\left\{ \left| \int (d_\alpha \wedge c)^r \, dQ - \int (d_\alpha \wedge c)^r \, dP \right| : \alpha \in \mathcal{A}_n \right\}. \end{aligned}$$

Now choose $\gamma \in \mathcal{C}_{n,r}(Q) \cap \mathcal{A}_n$. Then

$$\begin{aligned} e_{n,r}(P)^r - e_{n,r}(Q)^r &= e_{n,r}^{(c)}(P)^r - \int d_\gamma^r \, dQ \\ &\le \int (d_\gamma \wedge c)^r \, dP - \int (d_\gamma \wedge c)^r \, dQ \\ &\le \sup\left\{ \left| \int (d_\alpha \wedge c)^r \, dP - \int (d_\alpha \wedge c)^r \, dQ \right| : \alpha \in \mathcal{A}_n \right\}. \end{aligned}$$

This yields the lemma. □

4. Uniform strong law of large numbers with rates. Let $X_1, X_2, \dots$ be i.i.d. $\mathbb{R}^d$-valued random variables with distibution P and let $P_k = \frac{1}{k}\sum_{i=1}^k \delta_{X_i}$ be the empirial measure of $X_1, \dots, X_k$. The empirical version of $e_{n,r}(P)^r$ is given by

$$e_{n,r}(P_k)^r = \frac{1}{k} \inf\left\{ \sum_{i=1}^k d_\alpha^r(X_i) : \alpha \subset \mathbb{R}^d, 1 \le |\alpha| \le n \right\}$$

(see Lemma 2.1). Set

$$Y_{k,r} = Y_{k,r}(P) = \sup_{n \ge 1} |e_{n,r}(P_k)^r - e_{n,r}(P)^r|$$

and

$$Z_{k,r} = Z_{k,r}(P) = \sup_{n \ge 1} |e_{n,r}(P_k) - e_{n,r}(P)|$$

and suppose $\int \|x\|^r \, dP(x) < \infty$. Let $r \ge 1$. Then by Lemma 2.1,

$$Z_{k,r} \le \rho_r(P_k, P). \tag{4.1}$$

It follows from the Glivenko–Cantelli theorem for ρ_r that

$$Z_{k,r} \to 0 \qquad \text{a.s.}, \ k \to \infty.$$

Since

$$Y_{k,r} \le r \max\{e_{1,r}(P_k)^{r-1}, e_{1,r}(P)^{r-1}\} Z_{k,r},$$

this implies

$$Y_{k,r} \to 0 \qquad \text{a.s.}, \ k \to \infty.$$

Let $r < 1$. It follows from Lemma 2.1 that

$$Y_{k,r} \le \rho_r(P_k, P)^r \tag{4.2}$$

and thus

$$Y_{k,r} \to 0 \qquad \text{a.s.}$$

Since

$$Z_{k,r} \le \frac{1}{r} \max\{e_{1,r}(P_k)^{1-r}, e_{1,r}(P)^{1-r}\} Y_{k,r},$$

this yields

$$Z_{k,r} \to 0 \qquad \text{a.s.}$$

In this section we determine a.s. rates of convergence to zero for $Y_{k,r}$ and $Z_{k,r}$. Let us first settle some measurability questions.

LEMMA 4.1. *Let $\mathcal{G} \subset L_1(P)$ be a class of continuous functions. Then $e_{n,r}(P_k)$ and $\|P_k - P\|_{\mathcal{G}}$ are measurable. If $\int \|x\|^r \, dP(x) < \infty$, then $\rho_r(P_k, P)$ is measurable.*

PROOF. The measurability of $\rho_r(P_k, P)$ follows from a functional description of ρ_r [cf. Rachev and Rüschendorf (1998), Corollary 2.5.2 and Lemma 8.4.34]. The other assertions are obvious. $\square$

For positive random variables V_n and numbers $a_n > 0$ we write $V_n = O(a_n)$ a.s. if there is a constant $C < \infty$ such that $\limsup_{n\to\infty} V_n/a_n \le C$ a.s.

THEOREM 4.2. *Suppose* $\operatorname{supp}(P)$ *is compact. The constants occuring in the sequel depend on $D = \operatorname{diam}(\operatorname{supp}(P))$, d and r.*

(a) *Let $d/1 \wedge r > 2$. There exist constants $C_1, C_2 \in (0, \infty)$ such that*

$$\mathbb{P}(k^{1\wedge r/d} Y_{k,r} > C_1) \le 5 \exp(-C_2 k^{(d-2\wedge 2r)/d})$$

for every $k \ge 1$. In particular

$$Y_{k,r} = O(k^{-1\wedge r/d}) \qquad \text{a.s.}, \ k \to \infty.$$

(b) *Let $d/1 \wedge r = 2$, that is, $d = 2$, $r \ge 1$ or $d = 1$, $r = 1/2$. There exist constants $C_3 \in (0, \infty)$ and $C_4 \in (1, \infty)$ such that*

$$\mathbb{P}\left(\frac{k^{1/2}}{Lk} Y_{k,r} > C_3\right) \le 5 \exp\bigl(-C_4 (Lk)^2\bigr)$$

for every $k \ge 1$, where $Lk = \log(k \vee e)$. In particular,

$$Y_{k,r} = O(k^{-1/2} \log k) \qquad \text{a.s.}$$

(c) *Let* $d/1 \wedge r < 2$, *that is*, $d = 1$ *and* $r > 1/2$. *Then*

$$Y_{k,r} = O\left(\left(\frac{\log\log k}{k}\right)^{1/2}\right) \qquad a.s.$$

Part (a) of the preceding theorem gives the exact a.s. order of convergence of $Y_{k,r}$ in case $r \leq 1$ and $P_a \neq 0$. In fact, since $e_{n,r}(P_k) = 0$ for $n \geq k$ and thus

$$Y_{k,r} = \sup_{n \leq k-1} |e_{n,r}(P_k)^r - e_{n,r}(P)^r| \vee e_{k,r}(P)^r,$$

it follows from Theorem 2.7 that

$$\liminf_{k\to\infty} k^{r/d} Y_{k,r} \geq \liminf_{k\to\infty} k^{r/d} e_{k,r}(P)^r > 0.$$

For the same reason part (b) gives the exact a.s. rate up to the $\log k$ term if $r \leq 1$ and $P_a \neq 0$. For $r = 1$ the $\log k$ term can probably be sharpened to $(\log k)^{1/2}$. However, this remains open.

Theorem 4.2 is a consequence of an entropy estimate for the function class $\bigcup_{n\geq 1} \mathcal{G}_{n,r}$ with respect to the $L_\infty(P)$-norm and known inequalities for empirical processes. For a subset B of a normed space and $\varepsilon > 0$, the covering number $N(\varepsilon, B)$ is the minimal number of balls of radius ε with center belonging to B needed to cover B. The entropy is $\log N(\varepsilon, B)$.

LEMMA 4.3. *We have*

$$\log N\left(\varepsilon, \bigcup_{n\geq 1} \mathcal{G}_{n,r}, \|\cdot\|_{L_\infty(P)}\right) \leq C\varepsilon^{-d/1\wedge r}$$

for every $\varepsilon > 0$ *and a constant* C *depending on* D, d *and* r *only*.

PROOF. Set $\mathcal{G}_r = \bigcup_{n=1}^\infty \mathcal{G}_{n,r}$. Let $\beta = \beta(\varepsilon) \subset A^D$ be a set of centers of balls of radius $\delta = \varepsilon/r(2D)^{r-1}$ if $r \geq 1$ and $\delta = \varepsilon^{1/r}$ if $r < 1$ that cover A^D. Let α be a finite subset of A^D. For every $a \in \alpha$ there is $b_a \in \beta$ such that $\|a - b_a\| \leq \delta$. Set $\gamma = \{b_a : a \in \alpha\}$. Then

$$\|d_\alpha - d_\gamma\|_\infty \leq H(\alpha, \gamma) \leq \delta$$

and hence

$$\|d_\alpha^r - d_\gamma^r\|_\infty \leq r(2D)^{r-1}\|d_\alpha - d_\gamma\|_\infty \leq \varepsilon \qquad \text{if } r \geq 1$$

and

$$\|d_\alpha^r - d_\gamma^r\|_\infty \leq \|d_\alpha - d_\gamma\|_\infty^r \leq \varepsilon \qquad \text{if } r < 1.$$

We deduce that the $L_\infty(P)$-balls with centers $\{d_\gamma^r : \gamma \subset \beta, \gamma \neq \varnothing\}$ and radius ε cover $\mathcal{G}_r$. Thus $N(\varepsilon, \mathcal{G}_r, \|\cdot\|_\infty)$ is bounded by $2^{|\beta|}$. We have $N(\varepsilon, \mathcal{G}_r, \|\cdot\|_\infty) = 1$

for $\varepsilon \geq (2D)^r$ [$\geq \operatorname{diam}(\mathcal{G}_r)$]. Furthermore, $\beta(\varepsilon)$ can be taken to satisty $|\beta(\varepsilon)| \leq K\varepsilon^{-d/1\wedge r}$ for all $0 < \varepsilon \leq (2D)^r$ and a constant K depending only an D, d and r. This proves the assertion with $C = K \log 2$. $\square$

PROOF OF THEOREM 4.2. Let $\mathcal{G}_r = \bigcup_{n\geq 1} \mathcal{G}_{n,r}(A)$ with $A = \operatorname{supp}(P)$. It follows from Lemma 2.4 that

$$Y_{k,r} \leq \|P_k - P\|_{\mathcal{G}_r} \qquad \text{a.s.}$$

(a) and (b). We have

$$\mathbb{P}\big(Y_{k,r} > (2D)^r t\big) \leq \mathbb{P}\big(k^{1/2}\|P_k - P\|_{\mathcal{G}_r/(2D)^r} > k^{1/2}t\big)$$

for all $t \geq 0$, where $\mathcal{G}_r/(2D)^r = \{g/(2D)^r : g \in \mathcal{G}_r\}$. Since all members g of $\mathcal{G}_r/(2D)^r$ satisfy $0 \leq g \leq 1$ on supp(P) and $N(\varepsilon, \mathcal{G}_r/(2D)^r, \|\cdot\|_\infty) = N((2D)^r\varepsilon, \mathcal{G}_r, \|\cdot\|_\infty)$, the assertions follow from Lemma 4.3 and a result of Alexander [(1984), Corollary 2.2 of Correction with $\varepsilon = 1/2$ and $\Psi(M, n, \alpha) = 2M^2$].

(c) The assertion follows from Lemma 4.3 and the LIL for $\|P_k - P\|_{\mathcal{G}_r}$ [cf. Alexander (1984), Theorem 2.12]. For $r \geq 1$, part (c) can also be deduced from the LIL for empirical distribution functions. In fact, let F denote the distribution function of P and let F_k denote the empirical distribution function of $X_1, \dots, X_k$. Let the underlying norm be the absolute value. Then

$$\rho_1(P_k, P) = \int |F_k(x) - F(x)|\, dx \leq D \sup_{x\in\mathbb{R}} |F_k(x) - F(x)| \qquad \text{a.s.}$$

Therefore, by Lemma 2.4,

$$Y_{k,r} \leq r(2D)^r \sup_{x\in\mathbb{R}} |F_k(x) - F(x)| \qquad \text{a.s.}$$

and the assertion follows from the LIL for $\sup_{x\in\mathbb{R}} |F_k(x) - F(x)|$. $\square$

REMARK. It would be of interest to know whether one can improve the exponent in the entropy bound of Lemma 4.3 in case $r > 1$. This would give better order bounds in Theorem 4.2 for $r > 1$.

The a.s. order of $Z_{k,r}$ for the uniform distribution on $[0, 1]^d$ with $d \geq 3$ can be deduced from a result of Talagrand (1994).

THEOREM 4.4. *Let $P = U([0, 1]^d)$ and $r \geq 1$.*

(a) *Let $d \geq 3$. There exists constants $C_1, C_2 \in (0, \infty)$ depending on d and r (and the underlying norm) such that*

$$\mathbb{P}(k^{1/d} Z_{k,r} > C_1) \leq C_2 k^{-2}$$

for every $k \geq 1$. In particular

$$Z_{k,r} = O(k^{-1/d}) \qquad \text{a.s., } k \to \infty.$$

(b) *Let $d = 1$. Then*

$$Z_{k,r} = O\left(\left(\frac{\log\log k}{k}\right)^{1/2}\right) \qquad a.s.$$

Part (a) of the preceding theorem gives the exact a.s order of convergence of $Z_{k,r}$ for every $r \geq 1$. This follows from Theorem 2.7. The case $d = 2$ remains open. For $r < 1$ we have nothing to add beside the fact that the upper bounds on the rate of $Y_{k,r}$ in Theorem 4.2 also provide upper bounds for $Z_{k,r}$.

PROOF OF THEOREM 4.4. By (4.1),

$$Z_{k,r} \leq \rho_r(P_k, P) \leq K_1 \rho_{r,l_2}(P_k, P),$$

where ρ_{r,l_2} denotes the L_r-Wasserstein metric with respect to the l_2-norm and K_1 is a constant that depends on the underlying norm on $\mathbb{R}^d$ only. The assumptions of Theorem 1.1 of Talagrand (1994) are fulfilled for the function $\varphi : \mathbb{R}^d \to \mathbb{R}$, $\varphi(x) = K_2 \|x\|_{l_2}^r$ with a suitable constant K_2 depending on r provided $r \geq \log 4/\log(5/4)$. It follows from this theorem that

$$\mathbb{P}\big(\rho_{r,l_2}(P_k, P) > K_3 K_2^{-1/r} k^{-1/d}\big) \leq K_3 k^{-2}$$

for some constant K_3 depending only on d. The above inequality holds for $r \geq \log 4/\log(5/4)$ and thus for all $r \geq 1$. This proves (a).

(b) Let F_k denote the kth empirical distribution function and let the underlying norm be the absolute value. Then

$$\rho_r(P_k, P) = \left(\int_0^1 |F_k^{-1}(t) - t|^r \, dt\right)^{1/r} \leq \sup_{t \in (0,1)} |F_k^{-1}(t) - t| = \sup_{x \in [0,1]} |F_k(x) - x|.$$

The assertion follows from (4.1) and Smirnov's LIL for $\sup_{x \in [0,1]} |F_k(x) - x|$. $\square$

REMARK. Let

$$\hat{Y}_{k,r} = \sup_{n \geq 1} |\hat{e}_{n,r,k}(P)^r - e_{n,r}(P)^r|$$

and

$$\hat{Z}_{k,r} = \sup_{n \geq 1} |\hat{e}_{n,r,k}(P) - e_{n,r}(P)|.$$

Then Lemma 2.6 and the above proofs show that Theorems 4.2 and 4.4 remain valid with $Y_{k,r}$ and $Z_{k,r}$ replaced by $\hat{Y}_{k,r}$ and $\hat{Z}_{k,r}$, respectively.

5. Concentration inequalities and asymptotics. Throughout let $A = \operatorname{supp}(P)$ and suppose A is compact. By Lemma 2.5, we have for $t \geq 0$ and $m \geq 0$,

$$\mathbb{P}\big(e_{n,r}(P_k)^r - e_{n+m,r}(P)^r < -t/\sqrt{k}\big) \leq \mathbb{P}(\sqrt{k}\|P_k - P\|_{\mathcal{G}_{n,m,r}} > t).$$

Then a straightforward modification of the arguments for the case $r = 1$ in Rhee and Talagrand (1989), Theorem A, gives the following inequalities. [Part (a) is an immediate consequence of Hoeffding's (1963) inequality.]

THEOREM 5.1. *The constants occuring in the sequel depend on the support A of P and r. Let $k, n \in \mathbb{N}$ and $0 \leq m \leq n - 1$.*

(a) *There exists a constant $C_1 \in (0, \infty)$ such that, for all $t \geq 0$,*

$$\mathbb{P}\big(e_{n,r}(P_k)^r - e_{n-m,r}(P)^r > t/\sqrt{k}\big) \leq \exp\big(-C_1 t^2 (1+m)^{2r/d}\big).$$

(b) *Let $d/1 \wedge r > 2$. There exist constants $C_2, C_3 \in (0, \infty)$ such that, for all $t \geq C_2 n^{1/2 - 1\wedge r/d}$,*

$$\mathbb{P}\big(e_{n,r}(P_k)^r - e_{n+m,r}(P)^r < -t/\sqrt{k}\big) \leq 5 \exp\big(-C_3 t^2 (1+m)^{2r/d}\big).$$

(c) *Let $d/1 \wedge r = 2$. There exist constants $C_4, C_5 \in (0, \infty)$ such that, for all $t \geq C_4(1 + \log(n/m \vee 1))$,*

$$\mathbb{P}\big(e_{n,r}(P_k)^r - e_{n+m,r}(P)^r < -t/\sqrt{k}\big) \leq 5 \exp\big(-C_5 t^2 (1+m)^{2r/d}\big).$$

(d) *Let $d/1 \wedge r < 2$. There exist constants $C_6, C_7 \in (0, \infty)$ such that, for all $t \geq C_6 (m \vee 1)^{1/2 - 1\wedge r/d}$,*

$$\mathbb{P}\big(e_{n,r}(P_k)^r - e_{n+m,r}(P)^r < -t/\sqrt{k}\big) \leq 5 \exp\big(-C_7 t^2 (1+m)^{2r/d}\big).$$

In the situation of Proposition 3.3 the preceding theorem can be improved considerably.

THEOREM 5.2. *Suppose $P = h\lambda^d$, A is a finite union of compact convex sets and the density satisfies $h \geq u > 0$ λ^d a.s. on A. Then there exist constants $C_1, C_2 \in (0, \infty)$-depending on u, A and r such that for all $t \geq C_1 n^{1/2 - 1\wedge r/d}$,*

$$\mathbb{P}\big(|e_{n,r}(P_k)^r - e_{n,r}(P)^r| > t/\sqrt{k}\big) \leq 6 \exp\big(-C_2 t^2 n^{2r/d}\big).$$

PROOF. Since the truncation problem for $c_n = Cn^{-1/d}$ is solved for any dimension d, any norm and any $r > 0$ (see Corollary 3.4) and thus by Lemma 3.5,

$$|e_{n,r}(P_k)^r - e_{n,r}(P)^r| \leq \|P_k - P\|_{\mathcal{G}_{n,r}^{(c_n)}},$$

the assertion follows along the lines of Rhee and Talagrand (1989), Theorem B. $\square$

Let $n = n(k)$ vary with k. For instance, in case where the hypotheses of Theorem 5.2 are satisfied we obtain from the Borel–Cantelli lemma

$$|e_{n,r}(P_k)^r - e_{n,r}(P)^r| = O\left(\left(\frac{\log k}{k}\right)^{1/2}\right) \qquad \text{a.s.}$$

if $d/1 \wedge r = 2$ and $n = n(k)$ is arbitrary and even

$$|e_{n,r}(P_k)^r - e_{n,r}(P)^r| = O(k^{-1/2}) \qquad \text{a.s.}$$

if $d/1 \wedge r = 2$ and $n(k) = \Omega((\log k)^{1/1\vee r})$, that is, $\liminf n(k)/(\log k)^{1/1\vee r} > 0$. The choice $n(k) = k$ shows that the first bound is sharp up to the $(\log k)^{1/2}$ term and the second bound is sharp in case $r \le 1$ and $P_a \ne 0$.

Now we deduce conditions for

$$|e_{n(k),r}(P_k)^r - e_{n(k),r}(P)^r| = o(n(k)^{-r/d}) \qquad \text{a.s.}, \ k \to \infty$$

to hold which in case $n(k) \to \infty$ is equivalent to

$$n(k)^{1/d} e_{n(k),r}(P_k) \to Q_r(P) \qquad \text{a.s.},$$

where

$$Q_r(P) = Q_r(d)\left(\int h^{d/(d+r)}\, d\lambda^d\right)^{(d+r)/dr} < \infty \tag{5.1}$$

and $h = dP_a/d\lambda^d$ (see Theorem 2.7). Recall that $Q_r(P) > 0$ if $P_a \ne 0$.

THEOREM 5.3. *Let $n(k) \to \infty$.*

(a) *We have*

$$\limsup_{k\to\infty} n(k)^{1/d} e_{n(k),r}(P_k) \le Q_r(P) \qquad \textit{a.s.}$$

(b) *If $n(k) = o(k^{d/(d+2(1\vee r)-2)})$, then*

$$n(k)^{1/d} e_{n(k),r}(P_k) \to Q_r(P) \qquad \textit{a.s.}, \ k \to \infty.$$

(c) *Let $P = U([0,1]^d)$, $d \ge 3$ and $r \ge 1$. If $n(k) = o(k)$, then*

$$n(k)^{1/d} e_{n(k),r}(P_k) \to Q_r(d) \qquad \textit{a.s.}, \ k \to \infty.$$

PROOF. (a) Let $0 < \delta < 1$ and $m = [\delta n]$. We write

$$e_{n,r}(P_k)^r = e_{n,r}(P_k)^r - e_{n-m,r}(P)^r + e_{n-m,r}(P)^r.$$

By Theorem 2.7,

$$\lim_{n\to\infty} n^{r/d} e_{n-m,r}(P)^r = (1-\delta)^{-r/d} Q_r(P)^r.$$

For $\varepsilon > 0$ and $t = \varepsilon k^{1/2} n^{-r/d}$, one obtains by Theorem 5.1(a)

$$\begin{aligned}\mathbb{P}\big(n^{r/d}\big(e_{n,r}(P_k)^r - e_{n-m,r}(P)^r\big) > \varepsilon\big) &= \mathbb{P}\big(e_{n,r}(P_k)^r - e_{n-m,r}(P)^r > t/\sqrt{k}\big)\\ &\le \exp(-C_1\varepsilon^2\delta^{2r/d}k)\end{aligned}$$

and so by the Borel–Cantelli lemma,

$$\limsup_{k\to\infty} n^{r/d}\big(e_{n,r}(P_k)^r - e_{n-m,r}(P)^r\big) \le 0 \qquad \text{a.s.}$$

It follows that

$$\limsup_{k\to\infty} n^{r/d} e_{n,r}(P_k)^r \le (1-\delta)^{-r/d} Q_r(P)^r \qquad \text{a.s.}$$

Letting δ tend to zero gives the assertion.

(b) For $0 < \delta < 1$ let $m = [\delta n]$. We write

$$e_{n,r}(P_k)^r = e_{n,r}(P_k)^r - e_{n+m,r}(P)^r + e_{n+m,r}(P)^r.$$

By Theorem 2.7, we have

$$\lim_{n\to\infty} n^{r/d} e_{n+m,r}(P)^r = (1+\delta)^{-r/d} Q_r(P)^r.$$

Let $\varepsilon > 0$ and let $t = \varepsilon k^{1/2} n^{-r/d}$. We will apply Theorem 5.1. Since $n = o(k^{d/(d+2(1\vee r)-2)})$ and $2(1\vee r) - 2 = 2r - 2(1\wedge r)$, we have, given any constant $C < \infty$, $t \ge C n^{1/2 - 1\wedge r/d}$ for k large enough. In case $d/1\wedge r = 2$, this yields $t \ge C_4(1 + \log(2/\delta)) \ge C_4(1 + \log(n/m \vee 1))$ and in case $d/1\wedge r < 2$, we obtain $t \ge C_6(\delta/2)^{1/2-1\wedge r/d} n^{1/2-1\wedge r/d} \ge C_6(m\vee 1)^{1/2-1\wedge r/d}$ for k large enough. Therefore, by Theorem 5.1(b)–(d),

$$\begin{aligned}\mathbb{P}\big(n^{r/d}\big(e_{n,r}(P_k)^r - e_{n+m,r}(P)^r\big) < -\varepsilon\big) &= \mathbb{P}\big(e_{n,r}(P_k)^r - e_{n+m,r}(P)^r < -t/\sqrt{k}\big)\\ &\le 5\exp(-C_8\varepsilon^2\delta^{2r/d}k),\end{aligned}$$

where $C_8 = \min\{C_3, C_5, C_7\}$. This implies

$$\liminf_{k\to\infty} n^{r/d}\big(e_{n,r}(P_k)^r - e_{n+m,r}(P)^r\big) \ge 0 \qquad \text{a.s.}$$

and hence

$$\liminf_{k\to\infty} n^{r/d} e_{n,r}(P_k)^r \ge (1+\delta)^{-r/d} Q_r(P)^r \qquad \text{a.s.}$$

Letting δ tend to zero gives the assertion in view of part (a).

(c) follows immediately from Theorem 4.4(a) and Theorem 2.7. □

REMARK. (a) For $d = 2$, $r = 1$, uniform distributions P on compact convex sets and the l_2-norm the limiting result of part (b) has been obtained by Zemel (1985) under the restriction $n(k) = o(k/\log k)$. For $d = 1$, $r = 2$ and distributions P with compact support and smoth densities a weak (in-probability) version of the limiting result (b) is contained in Wong [(1984), Theorem 2] under the restriction $n(k) = o((k/\log k)^{1/3})$. Theorem 5.3 shows that the $\log k$ term can be eliminated in both cases.

(b) Part (c) of Theorem 5.3 remains valid with $e_{n(k),r}(P_k)$ replaced by $\hat{e}_{n,r,k}(P)$. The same holds for the limiting result of part (b) but under stronger restrictions on $n(k)$. For instance, if $d/1 \wedge r > 2$ and $n(k) = o(k^{1/1\wedge r})$, then

$$\lim_{k\to\infty} n(k)^{1/d}\hat{e}_{n,r,k}(P) = Q_r(P) \qquad \text{a.s.,}\ k \to \infty.$$

This can easily be deduced from the remark following the proof of Theorem 4.4.

Acknowledgment. The authors thank the referees for their helpful comments that improved the presentation of the paper.

REFERENCES

ALEXANDER, K. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12** 1041–1067. [Correction (1987). *Ann. Probab*. **15** 428–430.]

BALLY, V. and PAGÈS, G. (2000). A quantization algorithm for solving multidimensional optimal stopping problems. Technical report, Univ. Paris 6.

GERSHO, A. and GRAY, R. M. (1992). *Vector Quantization and Signal Compression*. Kluwer, Boston.

GRAF, S. and LUSCHGY, H. (2000). *Foundations of Quantization for Probability Distributions. Lecture Notes in Math.* **1730**. Springer, Berlin.

GRAF, S. and LUSCHGY, H. (2001). The quantization dimension of self-similar distributions. *Math. Nachrichten*. To appear.

HOCHBAUM, D. and STEELE, J. M. (1982). Steinhaus's geometric location problem for random samples in the plane. *Adv. Appl. Probab.* **14** 56–67.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

PAGÈS, G. (1997). A space vector quantization method for numerical integration. *J. Comput. Appl. Math.* **89** 1–38.

RACHEV, S. T. and RÜSCHENDORF, L. (1998). *Mass Transportation Problems* **I**, **II**. Springer, New York.

RHEE, W. T. and TALAGRAND, M. (1989). A concentration inequality for the k-median problem. *Math. Oper. Res.* **14** 189–202.

TALGRAND, M. (1994). The transportation cost from the uniform measure to the empirical measure in dimension ≥ 3. *Ann. Probab.* **22** 919–959.

WONG, M. A. (1984). Asymptotic properties of univariate sample k-means clusters. *J. Classification* **1** 255–270.

ZEMEL, E. (1985). Probabilistic analysis of geometric location problems. *SIAM. J. Algebraic Discrete Methods* **6** 189–200.

DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE
UNIVERSITY OF PASSAU
D-94030 PASSAU
GERMANY
E-MAIL: graf@fmi.uni-passau.de

DEPARTMENT IV, MATHEMATICS
UNIVERSITY OF TRIER
D-54286 TRIER
GERMANY
E-MAIL: luschgy@uni-trier.de