Optimal quantization for uniform distributions on Cantor-like sets

Wolfgang Kreitmeier[⋆]

Abstract In this paper, the problem of optimal quantization is solved for uniform distributions on some higher dimensional, not necessarily self-similar $N-$adic Cantor-like sets. The optimal codebooks are determined and the optimal quantization error is calculated. The existence of the quantization dimension is characterized and it is shown that the quantization coefficient does not exist. The special case of self-similarity is also discussed. The conditions imposed are a separation property of the distribution and strict monotonicity of the first N quantization error differences. Criteria for these conditions are proved and as special examples modified versions of classical fractal distributions are discussed.

Keywords optimal quantization · quantization dimension · quantization coefficient · self-similar probabilities

Mathematics Subject Classification (2000) Primary 28A80 · Secondary 62H30

1 Introduction

The optimal quantization of probability distributions was first discussed by electrical engineers in the 1940's as a problem arising out of signal processing and data compression. The aim is to determine the best approximation of a $d-$dimensional probability distribution v by a discrete one, containing at most n supporting points. The approximation will always be generated by a set α on $\mathbb{R}^d$, consisting of at most n points. This set α is called codebook and with $r > 0$ and the Euclidean norm $\| \cdot \|$, the error distance is defined by

$$\Psi_{\alpha,r}(v) = \int \min_{a \in \alpha} \| x - a \|^r \, dv(x). \tag{1}$$

Subsequently, the $n-$optimal quantization error for v of order r is denoted by

$$V_{n,r}(v) = \inf\{\Psi_{\beta,r}(v) : \beta \subset \mathbb{R}^d, \mathrm{card}(\beta) \le n\}, \tag{2}$$

W. Kreitmeier
Fakultät für Informatik und Mathematik, Universität Passau, Innstraße 33, 94032 Passau, Germany
Tel.: +49(0)851/509-3014
Fax: +49(0)851/509-3137
E-mail: wolfgang.kreitmeier@uni-passau.de

where card denotes cardinality. The set α is called $n-$optimal set (or $n-$optimal codebook) for v of order r, if $\Psi_{\alpha,r}(v) = V_{n,r}(v)$. The problem of optimal quantization is to determine for every integer $n \in \mathbb{N} = \{1,2,..\}$ all $n-$optimal sets and to calculate the optimal quantization error $V_{n,r}(v)$.

A good overview to the theory of optimal quantization and its historical development is presented by Gray and Neuhoff [9]. For a comprehensive mathematical treatment of the problem, the reader is referred to Graf and Luschgy [5], [7]. A more applied approach to the subject, stressing aspects of information theory and signal processing, can be found in the book of Gersho and Gray [3].

Although under weak assumptions an $n-$optimal codebook exists for every integer n (cf. [5], Theorem 4.12), the non-convex nature of the $n-$optimal quantization problem makes it difficult to determine the $n-$optimal codebooks and the optimal quantization error. For nonsingular distributions, exact solutions were found only in some special cases (cf. [5], chapter 4.4, 5.2).

In the singular case, Graf and Luschgy [4] solved the quantization problem for the classical self-similar Cantor distribution. Subsequently, for more generalized Cantor distributions, which are not necessarily self-similar, among other results the optimal codebooks were determined by Kesseböhmer and Zhu [15]. Kreitmeier [17] extended this codebook result to a class of Cantor distributions, including the classical one studied by Graf and Luschgy.

Because exact solutions are difficult to find, the aim is to get high-rate asymptotics for the optimal quantization error. To this end, we consider the sequence $\left(\frac{r\log(n)}{-\log(V_{n,r}(v))} \right)$, which was introduced by Zador [23]. If n tends to infinity, the strict positive limit (if it exists) is called quantization dimension and denoted by $D_r(v)$. If the quantization dimension exists, the question arises whether or not the sequence $\left(n^{\frac{r}{D_r(v)}} V_{n,r}(v) \right)$ has a strict positive and finite limit. If yes, it is called quantization coefficient and denoted by $Q_r(v)$.

If the distribution v is absolute continous with respect to the d-dimensional Lebesgue measure and $\int x^{r+\delta} dv(x) < \infty$ for some $\delta > 0$, then the quantization dimension exists and equals d. The quantization coefficent exists under these conditions as well. This result goes back to Zador [22] , respectively Bucklew and Wise [1]. A complete proof was given by Graf and Luschgy (cf. [5], Theorem 6.2).

In the case of singular distributions, the situation is different. For self-similar distributions, satisfying the so-called open set condition, Graf and Luschgy [6] have shown that the quantization dimension exist. Pötzelberger [21] proved the existence of the quantization coefficient, if the self-similar distribution satisfies the strong separation condition and a non-arithmetic condition. Using other methods, Graf and Luschgy have extended this result by replacing the strong separation condition with the open set condition (cf. [8], Theorem 4.1). For singular distributions which are not self-similar, Lindsay gave an example for the non-existence of the quantization dimension (cf. [19], Example 5.5). Later on, the existence of the quantization dimension for distributions on (not necessarily self-similar) Cantor-like sets was systematically studied and characterized by Kesseböhmer and Zhu [15], Kreitmeier [17] and Zhu [24].

In this paper, we will solve for natural $N \geq 2$ the quantization problem for uniform distributions on one- or higher dimensional, not necessarily self-similar $N-$adic Cantor-like sets.

In section 3, for every $n \geq N$ the $n-$optimal codebooks are determined and a formula for the $n-$optimal quantization error is proved (cf. Theorem 1). The existence of the quantization dimension is characterized (cf. Proposition 4) and under weak assumptions we show that the quantization coefficient does not exist (cf. Proposition 5). Although the formula for the

quantization dimension is a special case of the results of Zhu [24], the class of distributions considered in this present paper is not completely embraced by the one studied in [24] (cf. section 3.2). In the self-similar case, we get a more explicit formula for the optimal quantization errors and determine the set of accumulation points of the sequence $\left(n^{\frac{r}{D_r(\nu)}} V_{n,r}(\nu) \right)$. In our proofs, the conditions imposed are a separation property of the uniform distribution on the Cantor-like set and strict monotonicity of the first N differences of optimal quantization errors. Both conditions will be analyzed in detail in section 3.3. To this end, we need some results about optimal quantization for Borel probabilities, uniformly concentrated on a disjoint union of balls with equal radius. These facts are denoted and proved in section 2 and interesting on its own. To show the applicability of our results from section 3, we discuss in section 5 the optimal quantization of the uniform distribution on modified versions of several classical fractal probabilities, e.g. Cantor distribution and uniform distribution on Sierpinski gasket. For the whole paper, we denote ω as a finite subset of $\mathbb{R}^d$ with $\operatorname{card}(\omega) = N$.

2 Uniformly separated probabilities

In this section, we study probability distributions which are uniformly concentrated on a disjoint union of N balls with equal radius. If the radius is small enough, then for $n \geq N$ the $n-$optimal codebooks have a separation property (cf. Definition 2), studied below in detail. Moreover, we derive an upper bound for the difference between the optimal quantization errors of the uniformly separated probability on the balls and the uniform distribution on the ball centers. Let $d \in \mathbb{N}$ and α be a nonempty finite subset of $\mathbb{R}^d$. We call

$$W(a|\alpha) := \left\{ x \in \mathbb{R}^d : \| x - a \| = \min_{b \in \alpha} \| x - b \| \right\}$$

the Voronoi cell of a with respect to α. Let $x \in \mathbb{R}^d$ and $l > 0$. Let

$$B(x,l) := \{ z \in \mathbb{R}^d : \| x - z \| \leq l \}$$

be the closed ball with center x and radius $l > 0$. Let ν be a Borel probability measure on $\mathbb{R}^d$. Recall $\omega \subset \mathbb{R}^d$ with $\operatorname{card}(\omega) = N \geq 2$.

Definition 1 The distribution ν is called **uniformly** $(l, \omega)-$ **separated**, respectively $(l, \omega)-$separated in short notation, if the balls $(B(x,l))_{x \in \omega}$ are pairwise disjoint and $\nu(B(x,l)) = \frac{1}{N}$ for every $x \in \omega$.

Let $n \in \mathbb{N}$ and $r > 0$. The set of all n-optimal codebooks for ν of order r is denoted by $C_{n,r}(\nu)$. Let $\alpha \in C_{n,r}(\nu)$. For $x \in \omega$ let

$$\underline{\alpha}(x) = \alpha \cap B(x,l)$$

and

$$\overline{\alpha}(x) = \{ a \in \alpha : W(a|\alpha) \cap B(x,l) \neq \emptyset \}.$$

Clearly, $\underline{\alpha}(x) \subseteq \overline{\alpha}(x)$. With

$$d_{\min}(\omega) := \min\{ \| x - y \| : x, y \in \omega; x \neq y \}$$

we define the minimal distance in ω.

2.1 Properties of the optimal codebooks

If $n \geq N$, the calculation of the $n-$optimal quantization error of v simplifies considerably, if $\{\overline{\alpha}(x) : x \in \omega\}$ becomes a partition of α, i.e. if

$$\overline{\alpha}(x) \cap \overline{\alpha}(y) = \emptyset \qquad (3)$$

for every $x, y \in \omega$ with $x \neq y$. Due to the importance of this property we make the following definition.

Definition 2 Let n be an integer and $n \geq N$. The (l, ω)-separated distribution v is called **separately quantizing**, if relation (3) holds for every $n-$optimal codebook $\alpha \in C_{n,r}(v)$.

Let $B, C \subset \mathbb{R}^d$ be arbitrary sets and

$$d(B, C) := \inf\{\| x - y \| : x \in B, y \in C\}$$

be the distance between these two sets. For $x \in \mathbb{R}^d$ we simply denote $d(x, B)$ instead of $d(\{x\}, B)$. With $\text{conv}(B)$ we indicate the convex hull of B, with $\overline{B}$ its topological closure.

For any Borel measurable set $A \subset \mathbb{R}^d$ with $v(A) > 0$, let $v(\cdot|A) = \frac{v(A \cap \cdot)}{v(A)}$ be the conditional distribution of v w.r.t. A. The support of v is denoted by $\text{supp}(v)$. The following Proposition states a sufficient condition for v to be separately quantizing and a formula for the optimal quantization error.

Proposition 1 *Let $n \in \mathbb{N}$ and $n \geq N$. Let $r \geq 1$. Assume that v is $(l, \omega)-$separated.*

(a) Let v be separately quantizing. Then $\overline{\alpha}(x) = \underline{\alpha}(x)$ for every $x \in \omega$. The $n-$optimal quantization error can be written as

$$V_{n,r}(v) = \sum_{x \in \omega} \sum_{a \in \underline{\alpha}(x)} v(W(a|\alpha) \cap B(x, l)) V_{1,r}(v(\cdot|W(a|\alpha) \cap B(x, l))).$$

For the special case $n = N$,

$$V_{N,r}(v) = \frac{1}{N} \sum_{x \in \omega} V_{1,r}(v(\cdot|B(x, l))).$$

(b) If $d_{\min}(\omega) > 10l$, then v is separately quantizing. Moreover, if $r > 1$, then additionally to (a),

$$\left(1 + \frac{1}{N}\right) V_{N,r}(v) < V_{N-1,r}(v). \qquad (4)$$

Proof Let $n \geq N$, $r \geq 1$ and $\alpha \in C_{n,r}(v)$.
(a) Let $x \in \omega$. First we will show that $\overline{\alpha}(x) \subset \underline{\alpha}(x)$. To this end let $a \in \overline{\alpha}(x)$. From the definition we get

$$W(a|\alpha) \cap B(x, l) \neq \emptyset.$$

Using property (3) we deduce

$$W(a|\alpha) \cap \bigcup_{z \in \omega \setminus \{x\}} B(z, l) = \emptyset.$$

Due to $v(\bigcup_{z \in \omega} B(z, l)) = 1$ and [5], Theorem 4.1 we obtain

$$0 < v(W(a|\alpha)) = v(W(a|\alpha) \cap B(x, l)), \qquad (5)$$

which yields

$$\overline{\mathrm{conv}(\mathrm{supp}(\nu(\cdot|W(a|\alpha))))} \subset B(x,l). \tag{6}$$

According to [5], Theorem 4.1., we have $\{a\} \in C_{1,r}(\nu(\cdot|W(a|\alpha)))$. The application of [5], Remark 4.6 (a) yields $a \in \overline{\mathrm{conv}(\mathrm{supp}(\nu(\cdot|W(a|\alpha))))}$. From (6) we get $a \in B(x,l)$. Hence, $a \in \underline{\alpha}(x)$. As already stated, the converse inclusion $\underline{\alpha}(x) \subset \overline{\alpha}(x)$ is obvious. Thus, we have $\overline{\alpha}(x) = \underline{\alpha}(x)$.

Next, we will prove the two equations for the $n-$optimal quantization error. From [5], Lemma 3.3 and [5], Theorem 4.2 we get

$$V_{n,r}(\nu) = \sum_{a \in \alpha} \nu(W(a|\alpha))V_{1,r}(\nu(\cdot|W(a|\alpha))).$$

With (3) and (5) we obtain

$$\begin{aligned}
V_{n,r}(\nu) &= \sum_{x \in \omega} \sum_{a \in \overline{\alpha}(x)} \nu(W(a|\alpha))V_{1,r}(\nu(\cdot|W(a|\alpha))) \\
&= \sum_{x \in \omega} \sum_{a \in \overline{\alpha}(x)} \nu(W(a|\alpha) \cap B(x,l))V_{1,r}(\nu(\cdot|W(a|\alpha) \cap B(x,l))).
\end{aligned}$$

From $\overline{\alpha}(x) = \underline{\alpha}(x)$ for every $x \in \omega$ we deduce the first equation in Proposition 1. If $n = N$ we have $\mathrm{card}(\underline{\alpha}(x)) = \mathrm{card}(\overline{\alpha}(x)) = 1$ for every $x \in \omega$. Hence, we get from the first equation,

$$V_{N,r}(\nu) = \frac{1}{N} \sum_{x \in \omega} V_{1,r}(\nu(\cdot|B(x,l))).$$

(b) Let $d_{\min}(\omega) > 10l$.

(b1) We will show that $\overline{\alpha}(x) \cap \overline{\alpha}(y) = \emptyset$ for every $x,y \in \omega$ with $x \neq y$.

To this end, let

$$\omega_1 = \{x \in \omega : d(B(x,l),\alpha) > 2l\} \tag{7}$$

and $\omega_2 = \omega \backslash \omega_1$. We will prove that

$$\overline{\alpha}(x) \subset B(x,5l) \text{ for every } x \in \omega_2. \tag{8}$$

Let $x \in \omega_2$. Then $y \in B(x,l)$ and $a \in \alpha$ exist, with

$$\| y - a \| = d(B(x,l),\alpha) \leq 2l.$$

Let $b \in \overline{\alpha}(x)$. A $y' \in B(x,l)$ exists, with

$$\| y' - b \| = d(y',\alpha) \leq \| y' - a \|.$$

Thus, we get

$$\begin{aligned}
\| x - b \| &\leq \| x - y' \| + \| y' - b \| \\
&\leq l + \| y' - a \| \\
&\leq l + \| y' - y \| + \| y - a \| \\
&\leq l + 2l + 2l = 5l,
\end{aligned}$$

which yields $b \in B(x,5l)$. Thus, relation (8) is proved.

Due to $d_{\min}(\omega) > 10l$ we obtain

$$\overline{\alpha}(x) \cap \overline{\alpha}(y) = \emptyset \text{ for every } x,y \in \omega_2, x \neq y. \tag{9}$$

The set ω_2 is nonempty because otherwise we would have

$$\Psi_{\omega,r}(v) \le l^r < (2l)^r < \Psi_{\alpha,r}(v) = V_{n,r}(v),$$

which contradicts $\mathrm{card}(\omega) = N \le n$. Assertion (b1) is proved, if $\omega_1 = \emptyset$. We proceed indirectly and assume that $\omega_1 \ne \emptyset$. From definition (7) of ω_1 we get

$$V_{n,r}(v) = \Psi_{\alpha,r}(v) > \sum_{x \in \omega_1} \frac{(2l)^r}{N} + \sum_{x \in \omega_2} \int_{B(x,l)} d(z, \overline{\alpha}(x))^r dv(z). \tag{10}$$

Two cases appear.

1. $\mathrm{card}(\omega_1) \le \mathrm{card}\left(\alpha \setminus \bigcup_{x \in \omega_2} \overline{\alpha}(x)\right)$.

Let $\beta = \omega_1 \cup \bigcup_{x \in \omega_2} \overline{\alpha}(x)$. Obviously

$$\begin{aligned}
\mathrm{card}(\beta) &\le \mathrm{card}(\omega_1) + \mathrm{card}(\cup_{x \in \omega_2} \overline{\alpha}(x)) \\
&\le \mathrm{card}(\alpha \setminus \cup_{x \in \omega_2} \overline{\alpha}(x)) + \mathrm{card}(\cup_{x \in \omega_2} \overline{\alpha}(x)) \\
&= \mathrm{card}(\alpha).
\end{aligned} \tag{11}$$

With (10) and $\omega_1 \ne \emptyset$ we have

$$\Psi_{\beta,r}(v) \le \sum_{x \in \omega_1} \frac{l^r}{N} + \sum_{x \in \omega_2} \int_{B(x,l)} d(z, \overline{\alpha}(x))^r dv(z) < \Psi_{\alpha,r}(v).$$

Together with (11), this contradicts the optimality of α.

2. $\mathrm{card}(\omega_1) > \mathrm{card}\left(\alpha \setminus \bigcup_{x \in \omega_2} \overline{\alpha}(x)\right)$.

Let

$$\omega_{2,s} = \{x \in \omega_2 : \mathrm{card}(\overline{\alpha}(x)) = 1\}$$

and

$$\omega_{2,p} = \{x \in \omega_2 : \mathrm{card}(\overline{\alpha}(x)) > 1\} = \omega_2 \setminus \omega_{2,s}.$$

We will derive in this second case a contradiction as well. To this end, we construct an alternative codebook β by taking points from $\bigcup_{x \in \omega_{2,p}} \overline{\alpha}(x)$ and replacing them by ω_1. We will show that β does not contain more points than α, but generates a quantization error less than the one of α.

From (9), we obtain

$$\mathrm{card}(\alpha) = \mathrm{card}\left(\alpha \setminus \bigcup_{x \in \omega_2} \overline{\alpha}(x)\right) + \sum_{x \in \omega_2} \mathrm{card}(\overline{\alpha}(x)).$$

On the other hand, $\mathrm{card}(\alpha) = \mathrm{card}(\omega_1) + \mathrm{card}(\omega_2) + (n - N)$. Thus, we deduce

$$\begin{aligned}
0 &< \mathrm{card}(\omega_1) - \mathrm{card}\left(\alpha \setminus \bigcup_{x \in \omega_2} \overline{\alpha}(x)\right) \\
&= \sum_{x \in \omega_2} \mathrm{card}(\overline{\alpha}(x)) - \mathrm{card}(\omega_2) - (n - N) \\
&= \sum_{x \in \omega_{2,p}} \mathrm{card}(\overline{\alpha}(x)) + \mathrm{card}(\omega_{2,s}) - \mathrm{card}(\omega_2) - (n - N) \\
&= \sum_{x \in \omega_{2,p}} \mathrm{card}(\overline{\alpha}(x)) - \mathrm{card}(\omega_{2,p}) - (n - N) \\
&\le \sum_{x \in \omega_{2,p}} (\mathrm{card}(\overline{\alpha}(x)) - 1).
\end{aligned}$$

Now, let $N_{2,p} = \mathrm{card}(\omega_{2,p})$ and $\{x_1,..,x_{N_{2,p}}\} = \omega_{2,p}$ an arbitrary enumeration of $\omega_{2,p}$. Based on this enumeration, we recursively define the mapping

$$\varphi : \omega_2 \longrightarrow \mathbb{N}_0$$

by

(i) $\varphi(x) = 0$ for all $x \in \omega_{2,s}$,
(ii) for $i \in \{1,..,N_{2,p}\}$ and with $\sum_{j=1}^{0} \varphi(x_j) := 0$ it is

$$\varphi(x_i) = \min\left(\mathrm{card}(\overline{\alpha}(x_i)) - 1, \mathrm{card}(\omega_1) - \mathrm{card}(\alpha \setminus \bigcup_{x \in \omega_2} \overline{\alpha}(x)) - \sum_{j=1}^{i-1} \varphi(x_j) \right)$$

Obviously, φ has the following properties

(iii) $\varphi(x) \in \{0,..,\mathrm{card}(\overline{\alpha}(x)) - 1\}$ for every $x \in \omega_{2,p}$,
(iv) $\sum_{x \in \omega_2} \varphi(x) = \mathrm{card}(\omega_1) - \mathrm{card}\left(\alpha \setminus \bigcup_{x \in \omega_2} \overline{\alpha}(x)\right)$.

With this mapping φ, we take for every $x \in \omega_2$ a

$$\beta(x) \in C_{\mathrm{card}(\overline{\alpha}(x)) - \varphi(x), r}(\nu(\cdot | B(x,l)))$$

and define the already mentioned codebook

$$\beta = \omega_1 \cup \bigcup_{x \in \omega_2} \beta(x).$$

Now, we prove that $\Psi_{\beta,r}(\nu) < V_{n,r}(\nu)$.

At first we get the upper bound

$$\Psi_{\beta,r}(\nu) \leq \sum_{x \in \omega_1} \frac{l^r}{N} + \sum_{x \in \omega_2} \int_{B(x,l)} d(z, \beta(x))^r d\nu(z) \tag{12}$$
$$= \sum_{x \in \omega_1} \frac{l^r}{N} + \frac{1}{N} \sum_{x \in \omega_2} V_{\mathrm{card}(\beta(x)),r}(\nu(\cdot | B(x,l))).$$

From (10) and (12), we obtain

$$V_{n,r}(\nu) - \Psi_{\beta,r}(\nu) > \sum_{x \in \omega_2} \int_{B(x,l)} d(z, \overline{\alpha}(x))^r d\nu(z)$$
$$- \frac{1}{N} \sum_{x \in \omega_2} V_{\mathrm{card}(\beta(x)),r}(\nu(\cdot | B(x,l))) + \sum_{x \in \omega_1} \frac{l^r(2^r - 1)}{N}$$
$$= \frac{1}{N} \sum_{x \in \omega_2} \Psi_{\overline{\alpha}(x),r}(\nu(\cdot | B(x,l)))$$
$$- \frac{1}{N} \sum_{x \in \omega_2} V_{\mathrm{card}(\beta(x)),r}(\nu(\cdot | B(x,l))) + \sum_{x \in \omega_1} \frac{l^r(2^r - 1)}{N}.$$

From definition (2) of the optimal quantization error, we get

$$V_{n,r}(\nu) - \Psi_{\beta,r}(\nu) \tag{13}$$
$$> \sum_{x \in \omega_1} \frac{l^r(2^r - 1)}{N}$$
$$+ \frac{1}{N} \sum_{x \in \omega_2} \left(V_{\text{card}(\overline{\alpha}(x)),r}(\nu(\cdot|B(x,l))) - V_{\text{card}(\beta(x)),r}(\nu(\cdot|B(x,l)))\right)$$
$$= \sum_{x \in \omega_1} \frac{l^r(2^r - 1)}{N}$$
$$+ \frac{1}{N} \sum_{x \in \{z \in \omega_2 : \varphi(z) \neq 0\}} \left(V_{\text{card}(\overline{\alpha}(x)),r}(\nu(\cdot|B(x,l)))\right.$$
$$\left. - V_{\text{card}(\overline{\alpha}(x)) - \varphi(x),r}(\nu(\cdot|B(x,l)))\right).$$

From property (iii) and (iv) of φ, we deduce

$$\text{card}(\{z \in \omega_2 : \varphi(z) \neq 0\}) \leq \sum_{x \in \omega_2} \varphi(x)$$
$$= \text{card}(\omega_1) - \text{card}\left(\alpha \setminus \bigcup_{x \in \omega_2} \overline{\alpha}(x)\right) \leq \text{card}(\omega_1).$$

Inequality (13) thus transforms into

$$V_{n,r}(\nu) - \Psi_{\beta,r}(\nu) \tag{14}$$
$$> \frac{1}{N}\left(\text{card}(\omega_1)l^r(2^r - 1) + \sum_{x \in \{z \in \omega_2 : \varphi(z) \neq 0\}} (0 - V_{1,r}(\nu(\cdot|B(x,l))))\right)$$
$$\geq \frac{1}{N}\left(\text{card}(\omega_1)l^r(2^r - 1) - \sum_{x \in \{z \in \omega_2 : \varphi(z) \neq 0\}} \Psi_{\{x\},r}(\nu(\cdot|B(x,l)))\right)$$
$$\geq \frac{\text{card}(\omega_1)}{N}\left(l^r(2^r - 1) - l^r\right) \geq 0.$$

Property (iv) of the mapping φ and (9) implies

$$\text{card}(\beta) \leq \text{card}(\omega_1) + \sum_{x \in \omega_2} \text{card}(\beta(x))$$
$$= \text{card}(\omega_1) + \sum_{x \in \omega_2} (\text{card}(\overline{\alpha}(x)) - \varphi(x))$$
$$= \text{card}(\omega_1) + \text{card}\left(\bigcup_{x \in \omega_2} \overline{\alpha}(x)\right) - \left(\text{card}(\omega_1) - \text{card}(\alpha \setminus \bigcup_{x \in \omega_2} \overline{\alpha}(x))\right)$$
$$= \text{card}(\alpha).$$

Thus, (14) contradicts the optimality of α. Because each case ends in a contradiction, we have proved assertion (b1).

(b2) We will prove inequality (4).

Our approach is to determine two distinguished points from ω for $\alpha \in C_{N-1,r}(v)$. Using the partitions of ω into the sets ω_1, $\omega_{2,s}$ and $\omega_{2,p}$, we can deduce an upper bound for $V_{N,r}(v)$, respectively a lower bound for $V_{N-1,r}(v)$. This enables us to prove inequality (4).

Let $r > 1$ and $\alpha \in C_{N-1,r}(v)$. Due to $\text{card}(\alpha) = N - 1$ there exist two different points $x_1, x_2 \in \omega$ and $a \in \alpha$ with $d(x_1, a) = d(x_1, \alpha)$, respectively $d(x_2, a) = d(x_2, \alpha)$. For every $z \in B(x_1, l)$ we have

$$d(z, \alpha) \geq d(x_1, \alpha) - d(x_1, z) \geq d(x_1, \alpha) - l,$$

respectively $d(z, \alpha) \geq d(x_2, \alpha) - l$ for $z \in B(x_2, l)$. Hence,

$$\int_{B(x_1, l)} d(z, \alpha)^r dv(z) + \int_{B(x_2, l)} d(z, \alpha)^r dv(z) \tag{15}$$
$$\geq \frac{1}{N} \left((\max(0, d(x_1, \alpha) - l))^r + (\max(0, d(x_2, \alpha) - l))^r \right).$$

Obviously,

$$x^r + y^r \geq 2 \left(\frac{1}{2} \right)^r (x + y)^r \tag{16}$$

for arbitrary $x, y \geq 0$ with $x + y > 0$. From $l < \frac{1}{2} d_{\min}(\omega)$ and (16) we deduce for inequality (15) that

$$\int_{B(x_1, l)} d(z, \alpha)^r dv(z) + \int_{B(x_2, l)} d(z, \alpha)^r dv(z) \tag{17}$$
$$\geq \frac{2}{N} \left(\frac{1}{2} \right)^r (d(x_1, a) + d(x_2, a) - 2l)^r$$
$$\geq \frac{2}{N} \left(\frac{1}{2} d_{\min}(\omega) - l \right)^r.$$

With statement (9), which is also valid if $\text{card}(\alpha) = N - 1$, we get

$$\text{card}(\alpha) \geq \sum_{x \in \omega_2} \text{card}(\overline{\alpha}(x)).$$

The definition of $\omega_{2,p}$ and $\omega_{2,s}$ yields

$$\sum_{x \in \omega_2} \text{card}(\overline{\alpha}(x)) = \sum_{x \in \omega_{2,s}} \text{card}(\overline{\alpha}(x)) + \sum_{x \in \omega_{2,p}} \text{card}(\overline{\alpha}(x))$$
$$= \text{card}(\omega_{2,s}) + \text{card}(\omega_{2,p}) + \sum_{x \in \omega_{2,p}} (\text{card}(\overline{\alpha}(x)) - 1)$$
$$\geq \text{card}(\omega_2) + \text{card}(\omega_{2,p}).$$

Thus, we have

$$\text{card}(\omega_2) + \text{card}(\omega_{2,p}) \leq \text{card}(\alpha)$$
$$= N - 1 = \text{card}(\omega_2) + \text{card}(\omega_1) - 1,$$

which yields

$$\text{card}(\omega_{2,p}) \leq \text{card}(\omega_1) - 1. \tag{18}$$

Using (17), we conclude that

$$V_{N-1,r}(\nu) = \sum_{x \in \omega_{2,s} \setminus \{x_1,x_2\}} \int_{B(x,l)} d(z,\alpha)^r d\nu(z) \tag{19}$$

$$+ \sum_{x \in \omega_{2,p} \setminus \{x_1,x_2\}} \int_{B(x,l)} d(z,\alpha)^r d\nu(z)$$

$$+ \sum_{x \in \omega_1 \setminus \{x_1,x_2\}} \int_{B(x,l)} d(z,\alpha)^r d\nu(z)$$

$$+ \sum_{x \in \{x_1,x_2\}} \int_{B(x,l)} d(z,\alpha)^r d\nu(z)$$

$$\geq \sum_{x \in \omega_{2,s} \setminus \{x_1,x_2\}} \frac{1}{N} V_{1,r}(\nu(\cdot|B(x,l)))$$

$$+ \sum_{x \in \omega_{2,p} \setminus \{x_1,x_2\}} \frac{1}{N} V_{\text{card}(\overline{\alpha}(x)),r}(\nu(\cdot|B(x,l)))$$

$$+ \sum_{x \in \omega_1 \setminus \{x_1,x_2\}} \frac{1}{N}(2l)^r + \frac{2}{N}\left(\frac{1}{2}d_{\min}(\omega) - l\right)^r$$

$$\geq \frac{1}{N}\left(\sum_{x \in \omega_{2,s} \setminus \{x_1,x_2\}} V_{1,r}(\nu(\cdot|B(x,l))) \right.$$

$$\left. + (\text{card}(\omega_1) - 2)(2l)^r + 2(\frac{1}{2}d_{\min}(\omega) - l)^r \right).$$

From (a) we get

$$V_{N,r}(\nu) = \frac{1}{N} \sum_{x \in \omega} V_{1,r}(\nu(\cdot|B(x,l))),$$

which implies

$$V_{N,r}(\nu) \leq \frac{1}{N}\left(\sum_{x \in \omega_{2,s}} V_{1,r}(\nu(\cdot|B(x,l))) + (\text{card}(\omega_{2,p}) + \text{card}(\omega_1))l^r \right). \tag{20}$$

With (19) and (20) we deduce

$$N\left(V_{N-1,r}(\nu) - (1 + \frac{1}{N})V_{N,r}(\nu) \right)$$

$$\geq \sum_{x \in \omega_{2,s} \setminus \{x_1,x_2\}} V_{1,r}(\nu(\cdot|B(x,l))) + (\text{card}(\omega_1) - 2)(2l)^r + 2\left(\frac{1}{2}d_{\min}(\omega) - l\right)^r$$

$$- \left(1 + \frac{1}{N}\right)\left(\sum_{x \in \omega_{2,s}} V_{1,r}(\nu(\cdot|B(x,l))) + (\text{card}(\omega_{2,p}) + \text{card}(\omega_1))l^r \right)$$

$$\geq (\text{card}(\omega_1) - 2)(2l)^r + 2\left(\frac{1}{2}d_{\min}(\omega) - l\right)^r$$

$$- \left(\frac{\text{card}(\omega_{2,s})l^r}{N} + (1 + \frac{1}{N})(\text{card}(\omega_{2,p}) + \text{card}(\omega_1))l^r \right) - 2 \cdot (2l)^r.$$

Due to $d_{\min}(\omega) > 10l$, the right hand side is strict greater than

$$\left((\operatorname{card}(\omega_1) - 2)2^r + 2 \cdot 4^r - \frac{\operatorname{card}(\omega_{2,s})}{N} \right.$$
$$\left. - \left(1 + \frac{1}{N} \right) (\operatorname{card}(\omega_{2,p}) + \operatorname{card}(\omega_1)) - 2 \cdot 2^r \right) l^r$$
$$= ((\operatorname{card}(\omega_1) - 2)2^r + 2 \cdot 4^r - 1 - \operatorname{card}(\omega_1) - \operatorname{card}(\omega_{2,p}) - 2 \cdot 2^r) l^r.$$

Due to (18) and $r > 1$ we obtain the lower bound

$$((\operatorname{card}(\omega_1) - 2)2^r + 2 \cdot 4^r - 2\operatorname{card}(\omega_1) - 2 \cdot 2^r) l^r$$
$$= ((\operatorname{card}(\omega_1))2^r - 4 \cdot 2^r + 2 \cdot 4^r - 2 \cdot \operatorname{card}(\omega_1)) l^r > 0,$$

which proves inequality (4). $\quad\blacksquare$

Remark 1 The sufficient condition $d_{\min}(\omega) > 10l$ in Proposition 1 (b) ensuring the separation property (3) of the optimal codebooks needs only the constant value 10 and does neither depend on ω nor the norm exponent r. It is also independent of the dimension d.

For $x \in \mathbb{R}^d$ we denote δ_x as the dirac measure in x.

Remark 2 One may conjecture that the condition $d_{\min}(\omega) > 10l$ in Proposition 1 (b) can be weakened, i.e. Proposition 1 (b) is probably still true for a smaller constant than 10. But we cannot drop the condition completely. For example, let $d = 1$ and $v = \frac{1}{2}(\delta_0 + \delta_1)$ be the uniform distribution on the points 0 and 1. Let $\omega = \{0, \frac{7}{10}\}$ and $l = \frac{3}{10}$. Hence, v is a $(l, \omega)-$separated measure. Clearly, $C_{2,r}(v) = \{\{0,1\}\}$ and $W(0|\{0,1\}) =]-\infty, \frac{1}{2}]$. Thus, we obtain

$$W(0|\{0,1\}) \cap B(0,l) = B(0,l) \neq \emptyset$$

and

$$W(0|\{0,1\}) \cap B\left(\frac{7}{10}, l \right) = \left[\frac{4}{10}, \frac{5}{10} \right] \neq \emptyset,$$

which implies that v is not separately quantizing.

2.2 Approximation of the optimal quantization error

For $d \in \mathbb{N}$ denote U_ω as the uniform distribution on ω. It is well known (see e.g. [5], p. 57 ff) that the $n-$optimal quantization error depends continuously from the distribution it refers to, if the distance between distributions is defined by a suitable metric. Moreover, this continuity is uniform in n.

In our special case, for every $\varepsilon > 0$ an $l > 0$ exists such that for every $(l, \omega)-$separated distribution v and $n \in \mathbb{N}$, we have $|V_{n,r}(v) - V_{n,r}(U_\omega)| < \varepsilon$. The following Proposition 2 refines this abstract approximation result by presenting an upper bound for the optimal quantization error difference, independent of n and tending to zero if l tends to zero. For an arbitrary set $C \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$ let

$$1_C(x) := \begin{cases} 1 & \text{, if } x \in C \\ 0 & \text{, otherwise} \end{cases}$$

be the characteristic function on C. We denote with

$$\operatorname{diam}(C) = \sup\{\| x - y \| : x, y \in C\}$$

the diameter of C.

Proposition 2 *Let $n \in \mathbb{N}$ and $r > 1$. Let ν be an (l, ω)-separated probability measure. Then*

$$|V_{n,r}(\nu) - V_{n,r}(U_\omega)| \leq \begin{cases} lr(l + \mathrm{diam}(\omega))^{r-1}, & \text{if } n < N \\ l^r, & \text{if } n \geq N \end{cases} .$$

Proof Let $n \geq N$. Then $V_{n,r}(U_\omega) = 0$ and

$$|V_{n,r}(\nu) - V_{n,r}(U_\omega)| = V_{n,r}(\nu) \leq \Psi_{\omega,r}(\nu) \leq l^r.$$

If $n < N$, we can identify two cases.

1. $V_{n,r}(\nu) \geq V_{n,r}(U_\omega)$.
With an $\alpha \in C_{n,r}(U_\omega)$ we obtain

$$V_{n,r}(\nu) - V_{n,r}(U_\omega) \leq \int d(y, \alpha)^r d\nu(y) - \int d(y, \alpha)^r dU_\omega(y) \tag{21}$$

$$= \sum_{x \in \omega} \left(\int_{B(x,l)} d(y, \alpha)^r d\nu(y) - \frac{1}{N} d(x, \alpha)^r \right).$$

Due to $\nu(B(x,l)) = \frac{1}{N}$, for every $x \in \omega$ the right hand side of (21) equals

$$\sum_{x \in \omega} \int_{B(x,l)} (d(y, \alpha)^r - d(x, \alpha)^r) \, d\nu(y) \tag{22}$$

$$\leq \sum_{x \in \omega} \int_{B(x,l)} (d(y, \alpha)^r - d(x, \alpha)^r) \, 1_{\{z \in B(x,l): d(z,\alpha) \geq d(x,\alpha)\}}(y) d\nu(y).$$

Now let $x \in \omega$ and $y \in B(x,l)$ such that $d(y, \alpha) \geq d(x, \alpha)$.
Choose a $\xi(y) \in [d(x, \alpha), d(y, \alpha)]$ such that

$$d(y, \alpha)^r - d(x, \alpha)^r = r\xi(y)^{r-1}(d(y, \alpha) - d(x, \alpha)) \tag{23}$$

From [5], Remark 4.6 (a) we deduce

$$\alpha \subset \overline{\mathrm{conv}(\mathrm{supp}(U_\omega))}.$$

On the other hand we have

$$\mathrm{diam}\left(\overline{\mathrm{conv}(\mathrm{supp}(U_\omega))}\right) = \mathrm{diam}(\omega),$$

which yields

$$\max(d(y, \alpha), d(x, \alpha)) \leq d(x, \alpha) + d(x, y)$$
$$\leq \mathrm{diam}(\omega) + l.$$

Equation (23) implies

$$d(y, \alpha)^r - d(x, \alpha)^r \leq r(l + \mathrm{diam}(\omega))^{r-1} d(x, y) \leq rl(l + \mathrm{diam}(\omega))^{r-1}.$$

Thus, we derive for (22) the upper bound

$$\sum_{x \in \omega} \int_{B(x,l)} rl(l + \mathrm{diam}(\omega))^{r-1} d\nu(y) = rl(l + \mathrm{diam}(\omega))^{r-1}.$$

Using (21) we get
$$V_{n,r}(v) - V_{n,r}(U_\omega) \le rl(l + \text{diam}(\omega))^{r-1}.$$

2. $V_{n,r}(v) < V_{n,r}(U_\omega)$.
Let $\alpha \in C_{n,r}(v)$. Similar to the first case, we obtain

$$V_{n,r}(U_\omega) - V_{n,r}(v)$$
$$\le \sum_{x \in \omega} \int_{B(x,l)} (d(x,\alpha)^r - d(y,\alpha)^r) dv(y)$$
$$\le \sum_{x \in \omega} \int_{B(x,l)} (d(x,\alpha)^r - d(y,\alpha)^r) 1_{\{z \in B(x,l): d(x,\alpha) \ge d(z,\alpha)\}}(y) dv(y).$$

Now we consider all $x \in \omega$ and $y \in B(x,l)$ with $d(y,\alpha) \le d(x,\alpha)$. As in the first case we deduce
$$0 < V_{n,r}(U_\omega) - V_{n,r}(v) \le rl(l + \text{diam}(\omega))^{r-1}.$$
The combination of both cases yields the assertion.

3 Uniform distributions on Cantor-like sets

Let $r > 1$. Recall $N \ge 2$ and the finite set $\omega = \{x_1, .., x_N\} \subset \mathbb{R}^d$, consisting of N different points $x_i \in \mathbb{R}^d$. Note that we do not make any restrictions on the dimension $d \in \mathbb{N}$. Let $(c_k) \subset]0,1[$ be a sequence of contraction ratios and $S_{k,i}$ contractive similitudes for $k \in \mathbb{N}$ and $i \in \{1, .., N\}$ with fixpoint x_i and contraction ratio c_k independent of i. We assume that $\sup_{k \in \mathbb{N}} c_k = \bar{c} < 1$ and that the mappings $S_{k,i}$ satisfy the *open set condition*, i.e. a bounded open set $V \subset \mathbb{R}^d$ exists such that for every $k \in \mathbb{N}$ and $i \in \{1, .., N\}$ we have

$$S_{k,i}(V) \subset V \tag{24}$$

and

$$S_{k,i}(V) \cap S_{k,j}(V) = \emptyset$$

for every $j \in \{1, .., N\} \setminus \{i\}$. Now we can construct the N-adic Cantor-like set E and the uniform distribution μ on it. To this end, let $m \in \mathbb{N}$, recall $\overline{V}$ as the topological closure of V in $\mathbb{R}^d$ and define for any $k \in \mathbb{N}$ the set

$$\mathscr{D}_k^m(\overline{V}) = \{S_{m,i_1} \circ ... \circ S_{k+m-1,i_k}(\overline{V}) : (i_1, .., i_k) \in \{1, .., N\}^k\}.$$

With $E_k^m(\overline{V}) = \bigcup_{F \in \mathscr{D}_k^m(\overline{V})} F$, we define the Cantor-like set $E^m(\overline{V}) = \bigcap_{k \in \mathbb{N}} E_k^m(\overline{V})$. By repeated subdivision, (cf. [2], p. 9) we get a unique Borel probability measure $\mu^{(m)}$ on $\mathbb{R}^d$, defined by $\mu^{(m)}(F) = N^{-k}$ for every $F \in \mathscr{D}_k^m(\overline{V})$. Note that $\mu^{(m)}$ is supported on $E^m(\overline{V})$. We write shortly $\mathscr{D}_k^1(\overline{V}) = \mathscr{D}_k(\overline{V})$, respectively $E_k^1(\overline{V}) = E_k(\overline{V})$, $E^1(\overline{V}) = E$ and $\mu^{(1)} = \mu$. We denote by $\{1, .., N\}^*$ the set of all words on the alphabet $1, .., N$, including the empty word $\emptyset$. With $\sigma = \sigma_1 ... \sigma_k \in \{1, .., N\}^*$, we call k the length of σ and write $|\sigma|$ for it. The empty word $\emptyset$ has length 0.
We denote $S_\sigma^m = S_{m,\sigma_1} \circ \cdots \circ S_{k+m-1,\sigma_k}$ respectively shortly $S_\sigma^1 = S_\sigma$. Moreover we write $\pi_k^m = \prod_{l=m}^{k+m-1} c_l$ respectively shortly $\pi_k^1 = \pi_k$.

Lemma 1 *Let m, k be integers and $\sigma \in \{1, .., N\}^*$ with length k. Then we have*

$$\mu^{(m)}(\cdot | S_\sigma^m(E^m(\overline{V}))) = \mu^{(m+k)} \circ (S_\sigma^m)^{-1}.$$

Proof The assertion follows from the construction of $\mu^{(m)}$ respectively $\mu^{(m+k)}$.

3.1 Solution of the optimal quantization problem

To solve the quantization problem for uniform distributions on Cantor-like sets, we assume that for every $m \geq 1$, the distribution $\mu^{(m)}$ satisfies a separation condition which incorporates that $\mu^{(m)}$ is separately quantizing (cf. Definition 2). Moreover we need a monotonicity condition for the first N differences of the optimal quantization errors of $\mu^{(m)}$. Let us define these two conditions.

(a) **Separate quantization** (SQ): we say that μ satisfies the separate quantization condition, if a set $\kappa = \{y_1,..,y_N\} \subset \mathbb{R}^d$ with $\mathrm{card}(\kappa) = N$ exists, such that for every $m \in \mathbb{N}$ an $l_m > 0$ exists, with
  (a1) $B(y_i, l_m) \supset S_{m,i}(E^m(\overline{V}))$ for every $i \in \{1,..,N\}$,
  (a2) the distribution $\mu^{(m)}$ being an (l_m, κ)-separated probability measure (cf. Definition 1) and separately quantizing (cf. Definition 2).
(b) **Monotonicity of quantization error differences** (MQED): we say that μ satisfies the monotonicity condition for the quantization error differences, if for every $m \in \mathbb{N}$ and $n \in \{1,..,N-1\}$ the relation

$$V_{n,r}(\mu^{(m)}) - V_{n+1,r}(\mu^{(m)}) > V_{n+1,r}(\mu^{(m)}) - V_{n+2,r}(\mu^{(m)})$$

does hold.

Remark 3 If the condition (SQ) is satisfied, also the *strong separation condition* is satisfied, i.e.

$$S_{m,i}(E^m(\overline{V})) \cap S_{m,j}(E^m(\overline{V})) = \emptyset \tag{25}$$

for every $m \in \mathbb{N}$ and $i,j \in \{1,..,N\}$ with $i \neq j$.

Proposition 3 *Let $m,n \in \mathbb{N}$ with $n \geq N$ and $\alpha \in C_{n,r}(\mu^{(m)})$. If condition (SQ) holds, then $n_1,..,n_N \in \mathbb{N}$ exist with $\sum_{i=1}^N n_i = n$ and*

$$V_{n,r}(\mu^{(m)}) = \frac{1}{N} \sum_{i=1}^N c_m^r V_{n_i,r}(\mu^{(m+1)}). \tag{26}$$

Moreover, for each $i \in \{1,..,N\}$ an n_i-optimal set $\beta_i \in C_{n_i,r}(\mu^{(m+1)})$ exists such that

$$\alpha = \bigcup_{i=1}^N S_{m,i}(\beta_i).$$

Proof Let $\alpha \in C_{n,r}(\mu^{(m)})$. Applying condition (SQ), we find with $\kappa = \{y_1,..,y_N\}$ that

$$\alpha = \bigcup_{i=1}^N \overline{\alpha}(y_i)$$

with $\overline{\alpha}(y_i) \cap \overline{\alpha}(y_j) = \emptyset$ for every $i,j \in \{1,..,N\}$ with $i \neq j$.
Define $n_i = \mathrm{card}(\overline{\alpha}(y_i)) > 0$ for every $i \in \{1,..,N\}$. From [5], Theorem 4.1 and 4.2 we deduce

$$V_{n,r}(\mu^{(m)}) = \sum_{i=1}^N \mu^{(m)}\Big(\bigcup_{a \in \overline{\alpha}(y_i)} W(a|\alpha) \Big) \cdot V_{n_i,r}\Big(\mu^{(m)}(\cdot| \bigcup_{a \in \overline{\alpha}(y_i)} W(a|\alpha))\Big). \tag{27}$$

Again by (SQ) and Lemma 1 we obtain

$$\mu^{(m)}(\cdot \mid \bigcup_{a \in \overline{\alpha}(y_i)} W(a|\alpha)) = \mu^{(m)}(\cdot \mid B(y_i, l_m)) \tag{28}$$

$$= \mu^{(m)}(\cdot \mid S_{m,i}(E^m(\overline{V}))) = \mu^{(m+1)} \circ S_{m,i}^{-1}$$

respectively

$$\mu^{(m)}(\bigcup_{a \in \overline{\alpha}(y_i)} W(a|\alpha)) = \frac{1}{N} \tag{29}$$

for every $i \in \{1,..,N\}$. The application of (28) and (29) to (27) yields

$$V_{n,r}(\mu^{(m)}) = \frac{1}{N} \sum_{i=1}^{N} V_{n_i,r}(\mu^{(m+1)} \circ S_{m,i}^{-1}).$$

From [5], Lemma 3.2 (a) and the definition of $S_{m,i}$ we get

$$V_{n,r}(\mu^{(m)}) = \frac{1}{N} \sum_{i=1}^{N} c_m^r V_{n_i,r}(\mu^{(m+1)}).$$

Define $\beta_i = S_{m,i}^{-1}(\overline{\alpha}(y_i))$ for every $i \in \{1,..,N\}$. From [5], Theorem 4.1 we obtain

$$S_{m,i}(\beta_i) \in C_{n_i,r}(\mu^{(m)}(\cdot \mid \bigcup_{a \in \overline{\alpha}(y_i)} W(a|\alpha))).$$

Applying (28), we deduce $S_{m,i}(\beta_i) \in C_{n_i,r}(\mu^{(m+1)} \circ S_{m,i}^{-1})$. Finally, [5], Lemma 3.2 (a) implies $\beta_i \in C_{n_i,r}(\mu^{(m+1)})$.

Remark 4 If condition (SQ) is satisfied, we get as a direct consequence of Proposition 3 for every $m \in \mathbb{N}$ the identities

$$V_{N,r}(\mu^{(m)}) = c_m^r \cdot V_{1,r}(\mu^{(m+1)})$$

respectively

$$V_{N+1,r}(\mu^{(m)}) = \frac{c_m^r}{N} \left((N-1)V_{1,r}(\mu^{(m+1)}) + V_{2,r}(\mu^{(m+1)}) \right).$$

As an application, we derive the equation

$$\frac{c_m^r}{N} \left(V_{1,r}(\mu^{(m+1)}) - V_{2,r}(\mu^{(m+1)}) \right) \tag{30}$$

$$= \frac{c_m^r}{N} \left(\frac{1}{c_m^r} V_{N,r}(\mu^{(m)}) - \frac{N}{c_m^r} V_{N+1,r}(\mu^{(m)}) + (N-1)V_{1,r}(\mu^{(m+1)}) \right)$$

$$= \frac{1}{N} V_{N,r}(\mu^{(m)}) - V_{N+1,r}(\mu^{(m)}) + \frac{c_m^r(N-1)}{N} \frac{1}{c_m^r} V_{N,r}(\mu^{(m)})$$

$$= V_{N,r}(\mu^{(m)}) - V_{N+1,r}(\mu^{(m)}),$$

which will be useful in the sequel.

Remark 5 In the self-similar case (cf. section 4), Graf and Luschgy (cf. [5], Lemma 14.10) have shown that an $n_0 \geq N$ exists such that identity (26) remains true for $n \geq n_0$ under the strong separation condition, which is, according to Remark 3, weaker than our condition (SQ). Nevertheless, n_0 will be greater than N in many cases.

For any $\sigma \in \{1,..,N\}^*$ and $n < |\sigma|$, we define

$$\sigma|_n = \begin{cases} 0 & \text{, if } n = 0 \\ \sigma_1...\sigma_n & \text{, otherwise} \end{cases}$$

as the restriction of σ to n. For any $\tau \in \{1,..,N\}^*$, we define the order relation $\sigma \prec \tau$ if and only if $|\sigma| \leq |\tau|$ and $\tau|_{|\sigma|} = \sigma$. We call σ a predecessor of τ, if $\sigma \prec \tau$. A finite set $\Gamma \subset \{1,..,N\}^*$ is called a finite antichain, if any two different words $\sigma, \tau \in \Gamma$ are incomparable, i.e. neither $\sigma \prec \tau$ nor $\tau \prec \sigma$. A finite antichain Γ is called maximal if any word $\tau \in \{1,..,N\}^*$ with $|\tau| \geq \max\{|\sigma| : \sigma \in \Gamma\}$ has a predecessor in Γ. By an iterative approach of Proposition 3 it is straightforward to see that the following statement holds.

Corollary 1 *Let $m,n \in \mathbb{N}$ with $n \geq N$. Assume that condition (SQ) holds and let α be an $n-$optimal set for $\mu^{(m)}$ of order $r > 1$. Then a maximal finite antichain $\Gamma \subset \{1,..,N\}^*$ and $n_\sigma \in \{1,..,N\}$ for every $\sigma \in \Gamma$ exists, with $\sum_{\sigma \in \Gamma} n_\sigma = n$ and*

$$V_{n,r}(\mu^{(m)}) = \sum_{\sigma \in \Gamma} \frac{(\pi_{|\sigma|})^r}{N^{|\sigma|}} V_{n_\sigma,r}(\mu^{(m+|\sigma|)}). \tag{31}$$

Moreover, for each $\sigma \in \Gamma$ an n_σ-optimal set β_σ for $\mu^{(m+|\sigma|)}$ of order r exists such that

$$\alpha = \bigcup_{\sigma \in \Gamma} S_\sigma^m(\beta_\sigma).$$

Remark 6 By one more application of Proposition 3, the statement of Corollary 1 does also hold with $n_\sigma \in \{1,..,N-1\}$ for every $\sigma \in \Gamma$.

For an integer k and $\Gamma \subset \{1,..,N\}^*$ let Γ_k be the subset of Γ consisting of all words with length k.

Remark 7 If the maximal finite antichain Γ from Proposition 3 contains only words of length k, i.e. $\Gamma = \Gamma_k$, then Γ consists of exactly N^k words.

For $\sigma = \sigma_1...\sigma_{k-1}\sigma_k \in \{1,..,N\}^*$ we denote $\sigma^- = \sigma_1...\sigma_{k-1}$. We say that σ is equivalent to $\tau \in \{1,..,N\}^*$, if $\sigma^- = \tau^-$.

Remark 8 Let Γ and β_σ, $\sigma \in \Gamma$ as in Corollary 1. Conversely to Remark 6, we can cluster the sets β_σ. We will denote exactly this. Let $l := \max\{|\sigma| : \sigma \in \Gamma\}$ and assume that $l \geq 2$. The equivalence relation described above induces the partition $\Gamma_{l,1},..,\Gamma_{l,w}$ of Γ_l with an integer $1 \leq w \leq N$. Now let $i \in \{1,..,w\}$ and $\tau \in \Gamma_{l,i}$. Because Γ is a maximal finite antichain, we get $\operatorname{card}(\Gamma_{l,i}) = N$. It is easy to see that

$$\bigcup_{\sigma \in \Gamma_{l,i}, \tau^- = \sigma^-} S_{m+l-1,\sigma_l}(\beta_\sigma)$$

is independent of the choice of $\tau \in \Gamma_{l,i}$ and an l_i-optimal set for $\mu^{(m+l-1)}$ with

$$N \leq l_i \leq N \cdot \max_{\sigma \in \Gamma_{l,i}} \operatorname{card}(\beta_\sigma).$$

Remark 9 In the self-similar case (cf. section 4), Graf and Luschgy (cf. [5], Corollary 14.2) have shown that the right hand side in equation (31) is greater or equal to $V_{n,r}(\mu^{(m)})$, if condition (SQ) is dropped completely.

If we additionally assume that the monotonicity condition (MQED) holds, we can sharpen identity (31) to a formula for the optimal quantization error, which depends only on the first N optimal quantization errors.

Theorem 1 *Let $r > 1$ and $n \geq N$. Let $k \geq 1$ be an integer and $i \in \{1,..,N-1\}$ such that $iN^k \leq n < (i+1)N^k$. Assume that condition (SQ) and (MQED) holds. Let α be an $n-$optimal set for μ of order r. Then for every $\sigma \in \{1,..,N\}^*$ with length k a set β_σ exists such that*

$$\alpha = \bigcup_{\sigma \in \{1,..,N\}^*, |\sigma|=k} S_\sigma(\beta_\sigma)$$

and $n - iN^k$ of the sets β_σ are $(i+1)-$optimal sets for $\mu^{(k+1)}$ of order r, while $(i+1)N^k - n$ of the sets β_σ are $i-$optimal sets for $\mu^{(k+1)}$ of order r. Moreover,

$$V_{n,r}(\mu) \tag{32}$$
$$= \frac{(\pi_k)^r}{N^k} \left(((i+1)N^k - n) \cdot V_{i,r}(\mu^{(k+1)}) + (n - iN^k) \cdot V_{i+1,r}(\mu^{(k+1)}) \right).$$

Proof To prove Theorem 1 we will use Corollary 1. We divide the proof into two steps.

1. We will show that condition (MQED) ensures that the set Γ in Corollary 1 is identical with Γ_k and that n_σ can take only two different adjacent values for every $\sigma \in \Gamma$.

Let $\Gamma \subset \{1,..,N\}^*$ be as in Corollary 1. Let l be the number of all appearing lengths in Γ and denote $\Gamma = \Gamma_{k_1} \cup \cdots \cup \Gamma_{k_l}$ with $k_1 < \cdots < k_l$. We study all appearing (sub)cases.

Case 1: $l = 1$.

Case 1.1: $k_1 = k$.

Case 1.1.1: $\max_{\sigma \in \Gamma} n_\sigma - \min_{\sigma \in \Gamma} n_\sigma \leq 1$.
We will show that this case is the only one which can happen and all remaining cases end in a contradiction or can be transformed into this case.

Case 1.1.2: $\max_{\sigma \in \Gamma} n_\sigma - \min_{\sigma \in \Gamma} n_\sigma > 1$.
Let $\sigma_0 \in \Gamma$ with $n_{\sigma_0} = \min_{\sigma \in \Gamma} n_\sigma$ and $\sigma_1 \in \Gamma$ with $n_{\sigma_1} = \max_{\sigma \in \Gamma} n_\sigma$. According to Corollary 1 let $\beta_\sigma \in C_{n_\sigma,r}(\mu^{(k+1)})$ for every $\sigma \in \Gamma$ such that $\alpha = \bigcup_{\sigma \in \Gamma} S_\sigma(\beta_\sigma)$ is an $n-$optimal set for μ or order r. Let $\beta'_{\sigma_0} \in C_{n_{\sigma_0}+1,r}(\mu^{(k+1)})$ and $\beta'_{\sigma_1} \in C_{n_{\sigma_1}-1,r}(\mu^{(k+1)})$. For $\sigma \in \Gamma$ we define

$$\gamma_\sigma = \begin{cases} \beta_\sigma, & \text{if } \sigma \in \Gamma \backslash \{\sigma_0, \sigma_1\} \\ \beta'_{\sigma_0}, & \text{if } \sigma = \sigma_0 \\ \beta'_{\sigma_1}, & \text{if } \sigma = \sigma_1 \end{cases} \quad .$$

Define $\alpha' = \bigcup_{\sigma \in \Gamma} S_\sigma(\gamma_\sigma)$. Clearly, $n = \text{card}(\alpha) = \text{card}(\alpha')$. By definition (1) and the construction of μ we obtain

$$\Psi_{\alpha',r}(\mu) = \int \min_{a \in \alpha'} \| x - a \|^r \, d\mu(x)$$
$$\leq \sum_{\sigma \in \Gamma} \frac{1}{N^k} \int \min_{a \in S_\sigma(\gamma_\sigma)} \| x - a \|^r \, d\mu(\cdot \mid S_\sigma(E))(x)$$

Applying Lemma 1, we get

$$\Psi_{\alpha',r}(\mu) \leq \sum_{\sigma \in \Gamma} \frac{1}{N^k} \int \min_{a \in S_\sigma(\gamma_\sigma)} \| x - a \|^r \, d\mu^{(k+1)} \circ S_\sigma^{-1}(x) \qquad (33)$$

$$= \sum_{\sigma \in \Gamma} \frac{(\pi_k)^r}{N^k} \int \min_{a \in \gamma_\sigma} \| x - a \|^r \, d\mu^{(k+1)}(x)$$

$$= \sum_{\sigma \in \Gamma} \frac{(\pi_k)^r}{N^k} V_{\mathrm{card}(\gamma_\sigma)}(\mu^{(k+1)}).$$

From Corollary 1, we deduce

$$V_{n,r}(\mu) = \sum_{\sigma \in \Gamma} \frac{(\pi_k)^r}{N^k} V_{\mathrm{card}(\beta_\sigma)}(\mu^{(k+1)}). \qquad (34)$$

The combination of (33) and (34) together with the definition of γ_σ implies that

$$V_{n,r}(\mu) - \Psi_{\alpha',r}(\mu) \geq \frac{(\pi_k)^r}{N^k} \left(V_{n_{\sigma_0},r}(\mu^{(k+1)}) + V_{n_{\sigma_1},r}(\mu^{(k+1)}) \right.$$
$$\left. - V_{n_{\sigma_0}+1,r}(\mu^{(k+1)}) - V_{n_{\sigma_1}-1,r}(\mu^{(k+1)}) \right).$$

Due to the choice of σ_0 and σ_1, we have

$$n_{\sigma_0} < n_{\sigma_0} + 1 \leq n_{\sigma_1} - 1 < n_{\sigma_1}.$$

Condition (MQED) then yields $V_{n,r}(\mu) - \Psi_{\alpha',r}(\mu) > 0$, which contradicts the optimality of α. Thus Case 1.1.2 cannot happen.

Case 1.2: $k_1 = k' \neq k$.
Corollary 1 ensures $n = \sum_{\sigma \in \Gamma} n_\sigma \leq N^{k'+1}$ respectively $n \geq N^{k'}$.
Because $N^k \leq n < N^{k+1}$, only the case $k' = k - 1$ with $n_\sigma = N$ for every $\sigma \in \Gamma$ can happen. But in this case, we can apply Proposition 3 and will end into Case 1.1.1 with $\Gamma = \Gamma_k$ and $n_\sigma = 1$ for every $\sigma \in \Gamma$.

Case 2: $l > 1$.

Case 2.1: a $\sigma_1 \in \Gamma_{k_1}$ exists, with $n_{\sigma_1} < N$.

Case 2.1.1: a $\sigma_l \in \Gamma_{k_l}$ exists, with $n_{\sigma_l} > 1$.
We proceed similar to Case 1.1.2. Let $\beta_\sigma \in C_{n_\sigma,r}(\mu^{(|\sigma|+1)})$ for every $\sigma \in \Gamma$, such that $\alpha = \bigcup_{\sigma \in \Gamma} S_\sigma(\beta_\sigma)$ is an n−optimal set for μ or order r.
Let $\beta'_{\sigma_1} \in C_{n_{\sigma_1}+1,r}(\mu^{(k_1+1)})$ and $\beta'_{\sigma_l} \in C_{n_{\sigma_l}-1,r}(\mu^{(k_l+1)})$. For $\sigma \in \Gamma$ we define

$$\gamma_\sigma = \left\{ \begin{array}{l} \beta_\sigma, \text{ if } \sigma \in \Gamma \backslash \{\sigma_1, \sigma_l\} \\ \beta'_{\sigma_1}, \text{ if } \sigma = \sigma_1 \\ \beta'_{\sigma_l}, \text{ if } \sigma = \sigma_l \end{array} \right. .$$

and $\alpha' = \bigcup_{\sigma \in \Gamma} S_\sigma(\gamma_\sigma)$. By the same arguments as in Case 1.1.2 we obtain

$$V_{n,r}(\mu) - \Psi_{\alpha',r}(\mu) \geq \sum_{\sigma \in \Gamma} \frac{(\pi_{|\sigma|})^r}{N^{|\sigma|}} \left(V_{\mathrm{card}(\beta_\sigma)}(\mu^{(|\sigma|+1)}) - V_{\mathrm{card}(\gamma_\sigma)}(\mu^{(|\sigma|+1)}) \right)$$

$$\geq \frac{(\pi_{k_1})^r}{N^{k_1}} \left(V_{n_{\sigma_1},r}(\mu^{(k_1+1)}) - V_{n_{\sigma_1}+1,r}(\mu^{(k_1+1)}) \right)$$

$$- \frac{(\pi_{k_l})^r}{N^{k_l}} \left(V_{n_{\sigma_l}-1,r}(\mu^{(k_l+1)}) - V_{n_{\sigma_l},r}(\mu^{(k_l+1)}) \right)$$

By iterative application of equation (30) from Remark 4 and condition (MQED), we deduce

$$
\begin{aligned}
&V_{n,r}(\mu) - \Psi_{\alpha',r}(\mu) \\
&= \frac{(\pi_{k_1})^r}{N^{k_1}} \left(V_{n_{\sigma_1},r}(\mu^{(k_1+1)}) - V_{n_{\sigma_1}+1,r}(\mu^{(k_1+1)}) \right. \\
&\qquad \left. - \frac{(\pi_{k_l}^{k_1+1})^r}{N^{k_l-k_1}} \left(V_{n_{\sigma_l}-1,r}(\mu^{(k_l+1)}) - V_{n_{\sigma_l},r}(\mu^{(k_l+1)}) \right) \right) \\
&\geq \frac{(\pi_{k_1})^r}{N^{k_1}} \left(V_{N-1,r}(\mu^{(k_1+1)}) - V_{N,r}(\mu^{(k_1+1)}) \right. \\
&\qquad \left. - \frac{c_{k_1+1}^r \cdot \ldots \cdot c_{k_l}^r}{N^{k_l-k_1-1}N} \left(V_{1,r}(\mu^{(k_1+1)}) - V_{2,r}(\mu^{(k_1+1)}) \right) \right) \\
&\geq \frac{(\pi_{k_1})^r}{N^{k_1}} \left(V_{N-1,r}(\mu^{(k_1+1)}) - V_{N,r}(\mu^{(k_1+1)}) \right. \\
&\qquad \left. - \left(V_{N,r}(\mu^{(k_1+1)}) - V_{N+1,r}(\mu^{(k_1+1)}) \right) \right) > 0,
\end{aligned}
$$

which contradicts the optimality of α. Thus, Case 2.1.1 does not appear.

Case 2.1.2: $n_\sigma = 1$ for every $\sigma \in \Gamma_{k_l}$.

Case 2.1.2.1: $l = 2$.

Case 2.1.2.1.1: $n_\sigma \geq N - 1$ for every $\sigma \in \Gamma_{k_1}$.
From Remark 8 we know that the sets β_σ, $\sigma \in \Gamma_{k_2}$, which are optimal for $\mu^{(k_2+1)}$, can be clustered to sets β'_σ, $\sigma \in \Gamma'$ with $\Gamma' = \{\sigma^- : \sigma \in \Gamma_{k_2}\}$, which are N-optimal for $\mu^{(k_2)}$. If $k_2 - 1 = k_1$ we end into case 1.1.1. If $k_2 - 1 > k_1$ we proceed as in case 2.1.1 and end in a contradiction.

Case 2.1.2.1.2: a $\sigma \in \Gamma_{k_1}$ exists, with $n_\sigma < N - 1$.
First, we cluster Γ_{k_2} to Γ' as described in Case 2.1.2.1.1. If $k_2 - 1 = k_1$, we proceed as in Case 1.1.2 and end in a contradiction. If $k_2 - 1 > k_1$, we adopt the arguments of Case 2.1.1, yielding a contradiction as well.

Case 2.1.2.2: $l > 2$.
Due to $k_2 - 1 > k_1$, we proceed as in Case 2.1.2.1.2 and end in a contradiction.

Case 2.2: $n_\sigma = N$ for every $\sigma \in \Gamma_{k_1}$.

Case 2.2.1: $l = 2$.

Case 2.2.1.1: $n_\sigma \leq 2$ for every $\sigma \in \Gamma_{k_2}$.

Case 2.2.1.1.1: $n_\sigma = 2$ for a $\sigma \in \Gamma_{k_2}$.
We subdivide Γ_{k_1} according to Remark 6 into $\Gamma'' \subset \{1,..,N\}^*$ such that
$\Gamma_{k_1} = \{\sigma^- : \sigma \in \Gamma''\}$. For every $\tau \in \Gamma_{k_1}$ then the N-optimal set β_τ is divided into 1-optimal sets $\beta_{\tau 1},..,\beta_{\tau N}$. If $k_2 - 1 = k_1$ we end in case 1.1.1. If $k_2 - 1 > k_1$ we proceed as in case 2.1.1 and end in a contradiction.

Case 2.2.1.1.2: $n_\sigma = 1$ for every $\sigma \in \Gamma_{k_2}$.
First we subdivide Γ_{k_1} into Γ'' like in case 2.2.1.1.1. If $k_2 - 1 = k_1$ then we end in case 1.1.1. If $k_2 - 1 > k_1$, we cluster Γ_{k_2} to Γ' as described in Case 2.1.2.1.1. If $k_2 - 1 = k_1 + 1$ and $N = 2$, then we proceed as in Case 1.1.2. Otherwise, we adopt the arguments of Case 2.1.1. Hence, both cases end in a contradiction.

Case 2.2.1.2: a $\sigma \in \Gamma_{k_2}$ exists, with $n_\sigma > 2$.

After subdivision of Γ_{k_1} into Γ'' like in case 2.2.1.1.1, we proceed, depending on the value of $k_2 - k_1$, like in case 1.1.2 respectively 2.1.1 and both ways end in a contradiction.

Case 2.2.2: $l > 2$.

Due to $k_2 - 1 > k_1$ we can proceed as in Case 2.2.1.1.2 respectively Case 2.1.1 and end in a contradiction.

2. We finish the proof.

From Corollary 1 and 1. we deduce for $n \in [iN^k, (i+1)N^k[$ that two integers m_i, m_{i+1} exist such that

$$
\begin{aligned}
V_{n,r}(\mu) &= \sum_{\sigma \in \Gamma_k} \frac{(\pi_k)^r}{N^k} V_{n_\sigma, r}(\mu^{(k+1)}) \\
&= \frac{(\pi_k)^r}{N^k} \left(m_i \cdot V_{i,r}(\mu^{(k+1)}) + m_{i+1} \cdot V_{i+1,r}(\mu^{(k+1)}) \right)
\end{aligned}
$$

with $i \cdot m_i + (i+1) \cdot m_{i+1} = n$. Applying Remark 7 we obtain $m_i + m_{i+1} = N^k$. Thus we get $m_i = (i+1)N^k - n$, respectively $m_{i+1} = n - iN^k$, which proves equation (32). The representation of the optimal sets follows from above and Corollary 1.

Remark 10 Although one may conjecture that conditions (SQ) and (MQED) in Theorem 1 can be weakened, they cannot be dropped completely. The example of one-dimensional Cantor distributions (cf. section 5.1) will demonstrate that (SQ) can not even be replaced by the strong separation condition (cf. (25)). Moreover, a discussion of modified versions of the Sierpinski gasket (cf. section 5.2) shows that condition (MQED) cannot be dropped completely either.

Remark 11 The special case of one-dimensional dyadic homogeneous Cantor distributions ($d = 1, N = 2$) was investigated by Kesseböhmer and Zhu [15] and Kreitmeier [17]. Although not explicitly stated, both conditions (SQ) and (MQED) were necessary in these papers as well. Lemma 3.6 in [15], respectively Proposition 3.11 in [17] indemnifies (SQ). The arguments ensuring that condition (MQED) is satisfied are somewhat hidden, but can be found in Lemma 3.3 - 3.5 of [15], respectively Lemma 4.1 - 4.3 of [17]. The appropriate version of Theorem 1 is stated as Proposition 3.7 in [15], respectively Theorem 4.4 in [17].

3.2 Quantization dimension and quantization coefficient

As defined in writing on the topic (see e.g. [5], Definition 11.1), we call

$$
\underline{D}_r(\mu) := \liminf_{n \to \infty} \frac{r \log(n)}{-\log(V_{n,r}(\mu))}
$$

the lower and

$$
\overline{D}_r(\mu) := \limsup_{n \to \infty} \frac{r \log(n)}{-\log(V_{n,r}(\mu))}
$$

the upper quantization dimension of μ of order r. If the two numbers agree, their common value $D_r(\mu) = \underline{D}_r(\mu) = \overline{D}_r(\mu)$ is called quantization dimension of μ of order r. First introduced by Zador [23], the quantization dimension for singular distributions was studied by

Graf and Luschgy (cf. [5], [6]) in the self-similar case. They have shown that the quantization dimension always exists in this case, if the open set condition is satisfied. Lindsay has shown (cf. [19], Example 5.5) that the quantization dimension needs not to exist for singular distributions that are not self-similar. Kesseböhmer and Zhu [15] studied the quantization dimension for homogeneous singular distributions on Cantor-like sets in the 1-dimensional case. To this end, they determined the structure of the optimal sets under the restriction that $\bar{c} = \sup_{i \in \mathbb{N}} c_i \leq \frac{1}{4}$. Kreitmeier [17] extended this codebook result and the related existence characterization of the quantization dimension to the case $\bar{c} \leq \frac{1}{3}$. Several authors (cf. [11], [12], [20]) studied Hausdorff and packing dimension of even more generalized Cantor-like sets as introduced by us in section 3. They need a boundary distortion condition, which turns in our setting into

$$\inf_{i \in \mathbb{N}} c_i > 0. \tag{35}$$

Hua et.al. (cf. [13], section 3.3) were able to determine Hausdorff and packing dimension of Cantor-like sets without using (35). But they needed (in our terms) the condition

$$\lim_{i \to \infty} \frac{\log(c_i)}{\log(\pi_i)} = 0.$$

The inquiries for quantization dimension of generalized one-dimensional Cantor sets (cf. [15], Theorem 1.6 (3), respectively [17], Proposition 5.1) abstained from (35). Zhu [24] also characterized the quantization dimension for a large class of singular distributions on higher dimensional Cantor-like sets. Due to this more general approach, he needed the boundary condition (35) there. In our setting (cf. Proposition 4), we can drop (35), but are restricted to the conditions (SQ) and (MQED).

Lemma 2 *Assume that condition (SQ) and (MQED) are satisfied. Then strict positive constants $0 < M_1 \leq M_2 < \infty$ exist such that for every $n \in \mathbb{N}$ with $n \in [iN^k, (i+1)N^k[$, the relation*

$$M_1 \cdot (\pi_k)^r \leq V_{n,r}(\mu) \leq M_2 \cdot (\pi_k)^r$$

does hold.

Proof Immediate consequence of Theorem 1 with $M_1 = \min_{l \in \{1,..,N\}} V_{l,r}(\mu) > 0$ and $M_2 = \max_{l \in \{1,..,N\}} V_{l,r}(\mu) < \infty$.

Proposition 4 *Assume that condition (SQ) and (MQED) holds. Then*

$$\underline{D}_r(\mu) = \liminf_{n \to \infty} \frac{-\log(N)}{\frac{1}{n} \sum_{i=1}^n \log(c_i)} \leq \limsup_{n \to \infty} \frac{-\log(N)}{\frac{1}{n} \sum_{i=1}^n \log(c_i)} = \overline{D}_r(\mu). \tag{36}$$

Proof From the definition of $\underline{D}_r(\mu)$ and $n \in [N^{k(n)}, N^{k(n)+1}[$, we deduce $V_{N^k,r}(\mu) \geq V_{n,r}(\mu)$ respectively

$$\begin{aligned}
\underline{D}_r(\mu) &= \liminf_{n \to \infty} \frac{r \log(n)}{-\log(V_{n,r}(\mu))} \\
&\leq \liminf_{k \to \infty} \frac{r \log(N^{k+1})}{-\log(V_{N^k,r}(\mu))} \\
&= \liminf_{k \to \infty} \frac{kr \log(N) + r \log(N)}{-\log(V_{N^k,r}(\mu))}.
\end{aligned}$$

Because $(V_{n,r}(\mu))_{n\in\mathbb{N}}$ tends to zero (cf. [5], Lemma 6.1), considering Lemma 2 we obtain

$$\underline{D}_r(\mu) \leq \liminf_{k\to\infty} \frac{kr\log(N)}{-\log(V_{N^k,r}(\mu))}$$

$$\leq \liminf_{k\to\infty} \frac{kr\log(N)}{-\log(M_2\cdot(\pi_k)^r)} = \liminf_{n\to\infty} \frac{\log(N)}{-\frac{1}{n}\sum_{i=1}^{n}\log(c_i)}.$$

By the same arguments the lower inequality can be shown. The identity for $\overline{D}_r(\mu)$ follows similarly.

Subsequently to the characterization of the quantization dimension we can ask whether the quantization coefficient $Q_r(\mu)$ exists. The following Proposition shows that under weak assumptions the quantization coefficient does not exist.

Proposition 5 *Let $r > 1$ and assume that (SQ) and (MQED) does hold. Further assume that* $\left(\frac{\log(N)}{-\frac{1}{n}\sum_{i=1}^{n}\log(c_i)}\right)_n$ *converges in $\mathbb{R}$. Then D_r exists and*

(i) $\limsup_{n\to\infty} n^{\frac{r}{D_r}} V_{n,r}(\mu) < \infty$, *if and only if* $\limsup_{k\to\infty} N^{\frac{k}{D_r}}\pi_k < \infty$,
(ii) $\liminf_{n\to\infty} n^{\frac{r}{D_r}} V_{n,r}(\mu) > 0$, *if and only if* $\liminf_{k\to\infty} N^{\frac{k}{D_r}}\pi_k > 0$.
(iii) Let $\sup_{k\in\mathbb{N}} c_k = \overline{c} < \frac{1}{N}$. If $0 < \liminf_{k\to\infty} N^{\frac{k}{D_r}}\pi_k$ and $\limsup_{k\to\infty} N^{\frac{k}{D_r}}\pi_k < \infty$, then

$$0 < \liminf_{n\to\infty} n^{\frac{r}{D_r}} V_{n,r}(\mu) < \limsup_{n\to\infty} n^{\frac{r}{D_r}} V_{n,r}(\mu) < \infty. \tag{37}$$

Proof Proposition 4 ensures the existence of the quantization dimension D_r. The assertions (i) and (ii) are an immediate consequence of Lemma 2.
(iii) Let $x \in]1,2[$ and consider the mapping

$$x \to s(x) = \frac{\log(x)}{\log(\frac{1}{2-x})}. \tag{38}$$

Clearly, $s(x) < 1$. Using de l'Hospital we obtain $\lim_{x\to 1} s(x) = 1$. Thus, we can choose an $\alpha \in]1,2[$ such that $\frac{-\log(N)}{\log(\overline{c})} < s(\alpha) < 1$. As a consequence, we get

$$D_r = \liminf_{n\to\infty} \frac{\log(N)}{-\frac{1}{n}\sum_{i=1}^{n}\log(c_i)} \leq \frac{-\log(N)}{\log(\overline{c})} < s(\alpha) < 1.$$

Now, let $n_k = N^k$ and $m_k = \alpha_k N^k$ with $\alpha_k = [\alpha N^k]/N^k$ for every $k \in \mathbb{N}$. Here $[\alpha N^k]$ denotes the integer part of αN^k. Using (32), we deduce $n_k^{\frac{r}{D_r}} V_{n_k,r}(\mu) = (N^k)^{\frac{r}{D_r}}(\pi_k)^r V_{1,r}(\mu^{(k+1)})$, respectively

$$m_k^{\frac{r}{D_r}} V_{m_k,r}(\mu)$$
$$= \alpha_k^{\frac{r}{D_r}}(N^k)^{\frac{r}{D_r}}\frac{(\pi_k)^r}{N^k}\left((2N^k - \alpha_k N^k)V_{1,r}(\mu^{(k+1)}) + (\alpha_k N^k - N^k)V_{2,r}(\mu^{(k+1)})\right)$$
$$\geq (2 - \alpha_k)\alpha_k^{\frac{r}{D_r}}(N^k)^{\frac{r}{D_r}}(\pi_k)^r V_{1,r}(\mu^{(k+1)}).$$

Clearly, $\alpha_k \xrightarrow{k \to \infty} \alpha$. Together with (38), we obtain

$$\frac{m_k^{\frac{r}{D_r}} V_{m_k,r}(\mu)}{n_k^{\frac{r}{D_r}} V_{n_k,r}(\mu)} = \alpha_k^{\frac{r}{D_r}} (2 - \alpha_k) \xrightarrow{k \to \infty} \alpha^{\frac{r}{D_r}} (2 - \alpha) > \alpha^{\frac{1}{s(\alpha)}} (2 - \alpha) = 1.$$

Combining this with (i) and (ii), we get

$$0 < \liminf_{n \to \infty} n^{\frac{r}{D_r}} V_{n,r}(\mu) \leq \liminf_{k \to \infty} n_k^{\frac{r}{D_r}} V_{n_k,r}(\mu)$$

$$< \limsup_{k \to \infty} m_k^{\frac{r}{D_r}} V_{m_k,r}(\mu) \leq \limsup_{n \to \infty} n^{\frac{r}{D_r}} V_{n,r}(\mu) < \infty.$$

Remark 12 It remains an open question whether the existence characterization (36) of the quantization dimension does also hold, if only condition (SQ) is required. The techniques developed by Zhu [24] need the boundary condition $\inf_{i \in \mathbb{N}} c_i > 0$ and therefore cannot be used here. Moreover, it remains unanswered whether it is possible to find conditions, which are equivalent to (36).

Remark 13 It is also still an open question whether (37) remains valid, if we drop some or all conditions in Proposition 5 (iii).

3.3 Criteria for the imposed conditions

In order to make our results applicable, we need a criterion ensuring that condition (SQ), respectively (MQED) is satisfied. Such a criterion may only depend on V, ω and $(c_k)_{k \in \mathbb{N}}$. Note that for every integer m and $i \in \{1, .., N\}$ the similitude $S_{m,i}(\cdot)$ can be written as

$$S_{m,i}(\cdot) = c_m O_{m,i}(\cdot - x_i) + x_i \tag{39}$$

with an orthonormal map $O_{m,i}$. (cf. [14], Proposition 2.3 (1)). Let us first note a simple but essential fact.

Lemma 3 *It is $\omega \subset \overline{V}$.*

Proof Let $m \in \mathbb{N}$ and $i \in \{1, .., N\}$. Because $S_{m,i}(.)$ is a similitude, we get from relation (24) that

$$S_{m,i}(\overline{V}) \subset \overline{V}. \tag{40}$$

Now, we proceed indirectly. Assume that $x_i \notin \overline{V}$. Thus, we get

$$\sup\{\delta > 0 : B(x_i, \delta) \cap \overline{V} = \emptyset\} =: \delta_i > 0.$$

Moreover, a $w \in \overline{V}$ exists such that $\| x_i - w \| = \delta_i$. Using (39), we obtain

$$S_{m,i}(w) \in S_{m,i}(B(x_i, \delta_i)) = B(S_{m,i}(x_i), c_m \delta_i) = B(x_i, c_m \delta_i).$$

Because $c_m < 1$, we have $B(x_i, c_m \delta_i) \cap \overline{V} = \emptyset$. Hence, $S_{m,i}(w) \notin \overline{V}$, which contradicts relation (40).

In the sequel, we will need the constant $d_0 = \frac{d_{\min}(\omega)}{10 \operatorname{diam}(\overline{V})}$.

Proposition 6 *If* $\bar{c} = \sup_{i \in \mathbb{N}} c_i < d_0$ *and* $r > 1$, *then (SQ) holds with* $\kappa = \omega$ *and* $l_m = \bar{c} \cdot \text{diam}(\overline{V})$ *for every* $m \in \mathbb{N}$.

Proof Let $m \in \mathbb{N}$ and $l_m = \bar{c} \cdot \text{diam}(\overline{V})$. The construction of $E^m(\overline{V})$ and relation (24) implies $E^m(\overline{V}) \subset \overline{V}$. Using (39) and Lemma 3, we obtain

$$S_{m,i}(E^m(\overline{V})) \subset S_{m,i}(\overline{V}) \subset B(x_i, \bar{c} \cdot \text{diam}(\overline{V})), \tag{41}$$

which proves part (a1) of condition (SQ) with $\kappa = \omega$.

Due to $c_m < \frac{d_{\min}(\omega)}{2 \text{diam}(\overline{V})}$, the balls $B(x, l_m), x \in \omega$ are pairwise disjoint. Thus, the construction of $\mu^{(m)}$ and (41) yields $\mu^{(m)}(B(x, l_m)) = \frac{1}{N}$ for every $x \in \omega$. Hence, $\mu^{(m)}$ is an (l_m, ω)−separated probability measure. Because of $l_m < d_0 \cdot \text{diam}(\overline{V}) = \frac{d_{\min}(\omega)}{10}$, we can apply Proposition 1 (b) and deduce that $\mu^{(m)}$ is separately quantizing. Thus, (a2) of (SQ) is also proved.

Now, we need to develop a criterion for condition (MQED). To this end, we define for any Borel probability distribution ν, $r \geq 1$ and $N > 2$ the value

$$\begin{aligned}
&D_{\min}(r, N, \nu) \\
&= \min\{V_{n,r}(\nu) - V_{n+1,r}(\nu) - (V_{n+1,r}(\nu) - V_{n+2,r}(\nu)) : n \in \{1, .., N-2\}\} \\
&= \min\{V_{n,r}(\nu) + V_{n+2,r}(\nu) - 2V_{n+1,r}(\nu) : n \in \{1, .., N-2\}\}.
\end{aligned}$$

Proposition 7 *Let* $m \in \mathbb{N}$ *and* $N > 2$. *Let* $D_{\min}(r, N, U_\omega) > 0$ *and*

$$d_1 = \begin{cases} \min\left(d_0, \dfrac{D_{\min}(r, N, U_\omega)}{4r\left(\frac{11}{10} \text{diam}(\omega)\right)^{r-1} \text{diam}(\overline{V})}\right) & , \text{if } N > 2 \\ d_0 & , \text{if } N = 2 \end{cases}$$

If $\bar{c} < d_1$ *and* $r > 1$, *then condition (MQED) and (SQ) does hold.*

Proof We divide the proof into two steps.

First, we will show that $D_{\min}(r, N, \mu^{(m)}) > 0$.

Due to $\bar{c} < \frac{d_{\min}(\omega)}{2 \text{diam}(\overline{V})}$, we know that $\mu^{(m)}$ is an (l, ω)−separated probability measure with $l = \bar{c} \cdot \text{diam}(\overline{V})$. Applying Proposition 2, we obtain

$$\begin{aligned}
D_{\min}(r, N, \mu^{(m)}) &> D_{\min}(r, N, U_\omega) - 4lr(l + \text{diam}(\omega))^{r-1} \\
&\geq D_{\min}(r, N, U_\omega) - 4\bar{c} \cdot \text{diam}(\overline{V})r\left(\frac{d_{\min}(\omega)}{10} + \text{diam}(\omega)\right)^{r-1} \\
&\geq D_{\min}(r, N, U_\omega) - \bar{c} \cdot \text{diam}(\overline{V})4r\left(\frac{11}{10} \text{diam}(\omega)\right)^{r-1} > 0.
\end{aligned}$$

Secondly, we will show that $V_{N-1,r}(\mu^{(m)}) - V_{N,r}(\mu^{(m)}) > V_{N,r}(\mu^{(m)}) - V_{N+1,r}(\mu^{(m)})$. Applying Proposition 1 (b), we deduce

$$\begin{aligned}
&V_{N-1,r}(\mu^{(m)}) - V_{N,r}(\mu^{(m)}) - (V_{N,r}(\mu^{(m)}) - V_{N+1,r}(\mu^{(m)})) \\
&> \left(1 + \frac{1}{N}\right)V_{N,r}(\mu^{(m)}) - 2V_{N,r}(\mu^{(m)}) + V_{N+1,r}(\mu^{(m)}).
\end{aligned}$$

With Proposition 3 and (30), the right hand side turns into

$$\frac{1}{N}V_{N,r}(\mu^{(m)}) - \frac{c_m^r}{N}\left(V_{1,r}(\mu^{(m+1)}) - V_{2,r}(\mu^{(m+1)})\right)$$
$$= \frac{c_m^r}{N}V_{2,r}(\mu^{(m+1)}) \geq 0.$$

Combining the two steps, we recognize that (MQED) does hold.

Remark 14 If we replace $D_{\min}(r,N,U_\omega)$ by $|D_{\min}(r,N,U_\omega)|$ in the definition of d_1, it is easy to see that $D_{\min}(r,N,U_\omega) < 0$ implies $D_{\min}(r,N,\mu) < 0$, provided $\bar{c} < d_1$.

Remark 15 According to [5], Lemma 3.2 (a), the optimal quantization error scales with s^r under a similitude with scaling number $s > 0$. Thus, any criterion ensuring that (SQ) respectively (MQED) holds, should be invariant under such a similarity transformation of the whole model. Indeed, each of our criteria in this section satisfies this request. Note also that d_0, respectively d_1 in the criteria above do not depend on the cardinality of ω. Our criteria, ensuring that (SQ) does hold, is working with $\kappa = \omega$ and l_m independent of m. Nevertheless (SQ) is not restricted to this special setting.

4 The self-similar case

If $c = c_m$ is independent of $m \in \mathbb{N}$, we have $\mu = \mu^{(m)}$ for every $m \in \mathbb{N}$. Thus, μ becomes self-similar. In this case, we have $D = D_r(\mu) = \frac{\log(N)}{-\log(c)}$ for the quantization dimension (cf. [6]) because the similitudes satisfy the open-set condition. In this self-similar case, we can determine all accumulation points of $\left(n^{\frac{r}{D}}V_{n,r}(\mu)\right)_{n\in\mathbb{N}}$. To this end, let $f_c : [1,N] \to \mathbb{R}$ with

$$f_c(x) = x^{\frac{r}{D}}\left((i+1)V_{i,r}(\mu) - iV_{i+1,r}(\mu) - x(V_{i,r}(\mu) - V_{i+1,r}(\mu))\right),$$

if $x \in [i, i+1[$, $i \in \{1,..,N-1\}$ and $f_c(x) = V_{1,r}(\mu)$, if $x = N$. By adapting arguments from [4] and [17] in addition to Theorem 1 we get the following result.

Proposition 8 *Let $r > 1$ and $n \geq N$. Let $k \geq 1$ and $i \in \{1,..,N-1\}$ such that $iN^k \leq n < (i+1)N^k$. Assume that condition (SQ) and (MQED) holds. Then*

$$V_{n,r}(\mu) = \frac{c^{kr}}{N^k}\left(((i+1)N^k - n)\cdot V_{i,r}(\mu) + (n - iN^k)\cdot V_{i+1,r}(\mu)\right). \tag{42}$$

If, furthermore, $c \leq \frac{1}{N}$, then the set of all accumulation points of the sequence $\left(n^{\frac{r}{D}}V_{n,r}(\mu)\right)_{n\in\mathbb{N}}$ equals the interval $f_c([1,N])$, which contains more than one point.

Proof Equation (42) follows immediately from Theorem 1.
Due to $\lim_{x\to i-0} f_c(x) = f_c(i)$ for every $i \in \{2,..,N\}$, the function f_c is continuous and $f_c([1,N])$ is an interval. Moreover, we see for every $i \in \{1,..,N-1\}$, that f_c is differentiable on $]i, i+1[$. Using $r > 1$ and $c \leq \frac{1}{N}$, we deduce $\frac{r}{D} = \frac{r\log(c)}{-\log(N)} > 1$. Thus, for all $x \in]i, i+1[$ we get

$$f_c'(x) = \frac{r}{D}x^{\frac{r}{D}-1}\left((i+1)V_i - iV_{i+1}\right) + x^{\frac{r}{D}}\left(1 + \frac{r}{D}\right)(V_{i+1} - V_i).$$

Hence, f_c' is continuous on $]i, i+1[$ and has at most one zero in $]i, i+1[$. Thus, f_c is not constant, yielding that $f_c([1,N])$ contains more than one point. It remains to prove that $f_c([1,N])$ equals the set of all accumulation points of the sequence $\left(n^{\frac{r}{D}}V_{n,r}(\mu)\right)_{n\in\mathbb{N}}$. Let $y \in f_c([1,N])$.

Then an $i_0 \in \{1,..,N-1\}$ and $x \in [1,N]$ exist, with $x \in [i_0, i_0+1]$ and $y = f_c(x)$. Let $k \in \mathbb{N}$ and define $n_k := [xN^k]$ as the greatest integer, smaller or equal to xN^k. Obviously, an $x_k \in [1,N]$ exists, with $n_k = x_k N^k$. The sequence $(x_k)_{k\in\mathbb{N}}$ converges to x. Due to $k(n_k) = k$ and $i(n_k) = i_0$ we get from (42)

$$n_k^{\frac{r}{D}} V_{n_k,r}(\mu)$$
$$= (x_k N^k)^{\frac{r}{D}} \frac{c^{kr}}{N^k} \left(V_{i_0,r}(\mu)((i_0+1)N^k - x_k N^k) + V_{i_0+1,r}(\mu)(x_k N^k - i_0 N^k) \right).$$

Because of $c^D = \frac{1}{N}$, we deduce

$$n_k^{\frac{r}{D}} V_{n_k,r}(\mu)$$
$$= x_k^{\frac{r}{D}} \left((i_0+1)V_{i_0,r}(\mu) - i_0 V_{i_0+1,r}(\mu) - x_k(V_{i_0,r}(\mu) - V_{i_0+1,r}(\mu)) \right)$$
$$= f_c(x_k).$$

The continuity of f_c implies

$$y = f_c(x) = \lim_{k\to\infty} f_c(x_k).$$

Thus, y is an accumulation point of $(n^{\frac{r}{D}} V_{n,r}(\mu))_{n\in\mathbb{N}}$.

Now, let y be an accumulation point of $(n^{\frac{r}{D}} V_{n,r}(\mu))_{n\in\mathbb{N}}$. Then a subsequence $(n_l^{\frac{r}{D}} V_{n_l,r}(\mu))_{l\in\mathbb{N}}$ exists, with

$$y = \lim_{l\to\infty} n_l^{\frac{r}{D}} V_{n_l,r}(\mu).$$

W.l.o.g. we can assume that $k(n_l) < k(n_l + 1)$ for all $l \in \mathbb{N}$. Otherwise, we take a suitable subsequence. Let $x_l := \frac{n_l}{N^{k(n_l)}}$. Then $x_l \in [i(n_l), i(n_l)+1]$ and (42) imply

$$f_c(x_l) = n_l^{\frac{r}{D}} V_{n_l,r}(\mu).$$

Because $(x_l)_{l\in\mathbb{N}}$ is bounded, a subsequence $(x_{l_q})_{q\in\mathbb{N}}$ exists, converging against a $x \in [1,N]$. According to the continuity of f_c, we obtain

$$f_c(x) = \lim_{q\to\infty} f_c(x_{l_q}) = \lim_{q\to\infty} n_{l_q}^{\frac{r}{D}} V_{n_{l_q},r}(\mu) = y,$$

yielding $y \in f_c([1,N])$.

Remark 16 Equation (42) implies

$$V_{n,r}(\mu) - V_{n+1,r}(\mu) \geq V_{n+1,r}(\mu) - V_{n+2,r}(\mu) \tag{43}$$

for every $n \geq N$. (43) turns into an equation, if

$$i(n)N^{k(n)} \leq n < (i(n)+1)N^{k(n)} - 1.$$

Hence, the strict monotonicity for $n = 1,..,N-1$ (cf. condition (MQED)) is not carried over to all $n \in \mathbb{N}$.

Remark 17 If condition (SQ) and (MQED) are satisfied and $c < \frac{1}{N}$, then the non-existence of the quantization coefficient follows from Proposition 5, but is also an immediate consequence of Proposition 8.

Remark 18 Let $n \in \mathbb{N}$. An $n-$tuple $s = (s_1,..,s_n)$ with $s_i > 0$ is called arithmetic, if a $t > 0$ exists such that $s_i/t \in \mathbb{N}$. For such an arithmetic s, we call

$$\vartheta := \max\{t > 0 : s_i/t \in \mathbb{N} \text{ for every } i = 1,..,n\}$$

the period of s. Hence, the $N-$tuple $(-\log(c),..,-\log(c))$ with c as contracting parameter is arithmetic with period $-\log(c)$. In this arithmetic case, Pötzelberger ([21], Theorem 1) has shown that a periodic function $h : \mathbb{R}_+ \to \mathbb{R}_+$ with period $-\log(c)$ exists such that

$$\lim_{n \to \infty} (n^{\frac{r}{b}} V_{n,r}(\mu) - h(\log(n))) = 0.$$

In contrast to our special setting, this function h in [21] is not calculated explicitly.

Remark 19 The construction of the Cantor-like set E starts with the closed set $\overline{V}$. Hutchinson [14] has shown for the self-similar case that a replacement of $\overline{V}$ by any non-empty compact set $K \subset \mathbb{R}^d$ with $\bigcup_{i=1}^N S_i(K) \subset K$ leads always to the same Cantor-like set E.

5 Modified versions of classical fractals

5.1 One-dimensional Cantor distribution

Let $d = 1$ and $\omega = \{0,1\}$. Let $r > 1$ and (c_k) be a sequence of contracting factors, satisfying $c_k \in]0,\frac{1}{2}]$ for every $k \in \mathbb{N}$. We consider the similitudes $S_{k,1}(x) = c_k x$, respectively $S_{k,2}(x) = c_k x + 1 - c_k$. We choose $V =]0,1[$. Obviously, the open set condition is satisfied. The uniform distribution μ on $E = E(\overline{V})$ is called one-dimensional Cantor distribution. If $c_k = c$ independent of k, the distribution μ becomes self-similar. In the (sub-)case of $c = \frac{1}{3}$, the measure μ is the classical Cantor distribution.

If $\overline{c} = \sup_{i \in \mathbb{N}} c_i < \frac{1}{10}$, Proposition 7 ensures that the conditions (SQ) and (MQED) are satisfied. Applying Theorem 1, we can solve the optimal quantization problem in this case. Using more refined methods, Kesseböhmer and Zhu (cf. [15]) proved the result for the optimal codebooks in Theorem 1 under the relaxed condition $\overline{c} \leq \frac{1}{4}$. In another work, Kreitmeier [17] has shown that all assertions of Theorem 1 remain true, if $\overline{c} \leq \frac{1}{3}$. Note that the solution of the optimal quantization problem in the classical case ($c_k = \frac{1}{3}$ for every k) goes back to Graf and Luschgy (cf. [4]). Thus, our results in this paper for the one-dimensional Cantor distribution are already incorporated in other works. In this context, it is interesting to see that Theorem 1 becomes wrong, if $\overline{c} > \frac{1}{3}$. To show this, let $(a_l)_{l \in \mathbb{N}}$ be a zero sequence with $\sup_{l \in \mathbb{N}} a_l \leq \frac{1}{3}$. For $k,l \in \mathbb{N}$ and $c \in]0,\frac{1}{2}[$ let

$$c_{k,l} := \begin{cases} c, & \text{if } k \neq 2 \\ a_l, & \text{if } k = 2 \end{cases}$$

Denote μ_l as the uniform distribution on E defined by $(c_{k,l})_{k \in \mathbb{N}}$. Due to $c_{1,l} = c$ and because $(c_{2,l})_{l \in \mathbb{N}}$ is a zero sequence, the series $(\mu_l)_{l \in \mathbb{N}}$ converges weak against $\nu_c = \frac{1}{4}(\delta_0 + \delta_c + \delta_{1-c} + \delta_1)$. Moreover,

$$\int \| x \|^r \, d\mu_l(x) \to \int \| x \|^r \, d\nu_c(x),$$

because the support of μ_l and ν_c is contained in $[0,1]$ for every l. Using well-known stability and consistency results (cf. [5], p. 57 ff), one gets

$$V_{n,r}(\mu_l) \to V_{n,r}(\nu_c). \tag{44}$$

Moreover, a sequence $(\alpha_l)_{l \in \mathbb{N}}$ of codebooks exists, with $\alpha_l \in C_{n,r}(\mu_l)$ for every $l \in \mathbb{N}$ and

$$d_H(\alpha_l, C_{n,r}(\nu_c)) \overset{l \to \infty}{\longrightarrow} 0 \tag{45}$$

Now assume that Theorem 1 is valid. Hence,

$$C_{3,r}(\mu_l) = \{\{\frac{c}{2}, 1 - c + \frac{ca_l}{2}, 1 - \frac{ca_l}{2}\}, \{\frac{ca_l}{2}, c - \frac{ca_l}{2}, 1 - \frac{c}{2}\}\}.$$

In case of $c \leq \frac{1}{3}$ it is easy to see that

$$C_{3,r}(\nu_c) = \{\{\frac{c}{2}, 1 - c, 1\}, \{0, c, 1 - \frac{c}{2}\}\}.$$

but

$$C_{3,r}(\nu_c) = \{\{0, \frac{1}{2}, 1\}\},$$

if $c > \frac{1}{3}$. In the last case, the statement about optimal codebooks in Theorem 1 gets wrong because otherwise we would derive with

$$\inf_{l \in \mathbb{N}} \inf\{d_H(\alpha, \beta) : \alpha \in C_{3,r}(\mu_l), \beta \in C_{3,r}(\nu_c)\} \geq \frac{1}{2} - c > 0,$$

a contradiction to (45). In the same way, the assumed validity of equation (32) in case of $c > \frac{1}{3}$ would lead to

$$\begin{aligned}
\lim_{l \to \infty} V_{3,r}(\mu_l) &= \lim_{l \to \infty} \frac{c_1^r}{2}\left((2^1 - 1)V_{1,r}\left(\mu_l^{(2)}\right)\right) \\
&= \frac{c^r}{2}V_{1,r}\left(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1\right) \\
&= \frac{1}{2}\left(\frac{c}{2}\right)^r > \frac{1}{2}\left(\frac{1}{2} - c\right)^r = V_{3,r}(\nu_c),
\end{aligned}$$

which contradicts (44). Hence, all statements of Theorem 1 will get wrong in general, if $\bar{c} > \frac{1}{3}$. As a consequence, the condition (SQ) cannot be replaced by the weaker condition of strong separation (cf. (25)). Even in the self-similar case, it can be shown (cf. [17], Example 6.3) that Theorem 1 becomes invalid, if $c_k = c > \frac{5 - \sqrt{17}}{2}$.

5.2 Uniform distribution on modified Sierpinski Gasket

Now, we intend to apply our results to a set ω consisting of three different points. In view of our criteria (cf. section 3.3) for condition (SQ) and (MQED) it is essential to know, under which conditions $D_{\min}(r, 3, U_\omega) > 0$ does hold. At first, we will answer this question in Proposition 9. To this end, we will use properties of the Fermat point of a triangle. After that, we introduce the uniform distribution on a modified version of the Sierpinski Gasket and discuss the optimal quantization of it.

Lemma 4 *Let $\omega \subset \mathbb{R}^d$ be finite and non-empty, $r = 2$ and $s(\omega) = \frac{1}{N}\sum_{x \in \omega} x$. Then*

$$V_{1,r}(U_\omega) = \int \| x - s(\omega) \|^r \, dU_\omega(x) = \frac{1}{N}\sum_{x \in \omega} \| x - s(\omega) \|^r. \tag{46}$$

Equation (46) also holds in case of $r > 1$, if $\| x - s(\omega) \|$ does not depend on x.

Proof The proof of Theorem 2.4 in [5] shows the strict convexity of

$$\mathbb{R}^d \ni z \xrightarrow{F_r} \int \| x - z \|^r \, dU_\omega(x) \in \mathbb{R}_0^+ .$$

If F_r has only one critical value, it is the global minimum of F_r. With ∇ as gradient, we calculate

$$\nabla F_r(z) = \frac{r}{N} \sum_{x \in \omega} \| x - z \|^{r-2} (z - x).$$

If $r = 2$ or if $\| x - s(\omega) \|$ is independent of $x \in \omega$, we get $\nabla F_r(s(\omega)) = 0$.

Proposition 9 *Let $d \geq 2$, card$(\omega) = 3$ and $r > 1$. Then*

$$D_{\min}(r, 3, U_\omega) \geq \left(\frac{d_{\min}(\omega)}{\sqrt{3}} \right)^r - \frac{4}{3} \left(\frac{d_{\min}(\omega)}{2} \right)^r . \tag{47}$$

In general $D_{\min}(r, 3, U_\omega) > 0$ if $r > 2$. If and only if $d_{\min}(\omega) = \mathrm{diam}(\omega)$, (47) turns into an equation.

Proof According to [5], Lemma 3.2 (a), the optimal quantization error remains unchanged under isometrical mappings. Thus, we can assume w.l.o.g. that $d = 2$. Let $\omega = \{A, B, C\}$. Let

$$u = \inf\{\| A - P \| + \| B - P \| + \| C - P \| : P \in \mathbb{R}^2\}.$$

A point $F \in \mathbb{R}^2$ is 1-optimal for Q_ω of order 1, this means, $\{F\} \in C_{1,1}(Q_\omega)$, if and only if

$$u = \| A - F \| + \| B - F \| + \| C - F \| . \tag{48}$$

The minimization problem defined by (48) has a unique solution (cf. [10]): exactly one point $F \in \overline{\mathrm{conv}(\omega)}$ satisfies equation (48).

(a) If the triangle, spanned by ω, has an angle equal or greater than $\frac{2}{3}\pi$, then F is the vertex of this angle.

(b) Otherwise, F is the unique point, at which each of the three angles BFC, CFA and AFB - containing F as vertex - does have the value $\frac{2}{3}\pi$.

The point F is called Fermat point. Due to [5], Remark 4.6 (a), this minimization problem does not have another solution on $\mathbb{R}^2 \setminus \overline{\mathrm{conv}(\omega)}$. Hence, the Fermat point is unique in $\mathbb{R}^2$ and we have $C_{1,1}(U_\omega) = \{\{F\}\}$.

If we construct above each side of the triangle ABC an equilateral triangle (cf. figure 1), we get the triangles $BA'C$, ACB' and BAC'. The line segments AA', BB', and CC' are all the same length u and their common point of intersection is the Fermat point (cf. [10]). Obviously, $V_{3,r}(U_\omega) = 0$, respectively $V_{2,r}(U_\omega) = \frac{2}{3}(\frac{d_{\min}(\omega)}{2})^r$. We proceed separately for each of the two cases (a) and (b).
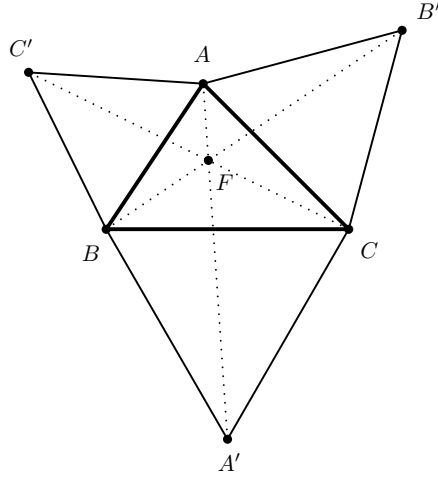
Case (a)

We obtain $u \geq 2 d_{\min}(\omega)$. Thus, the convexity of $t \to t^r$ yields

$$V_{1,r}(U_\omega) \geq \left(\frac{u}{3} \right)^r \geq \left(\frac{2}{3} d_{\min}(\omega) \right)^r > \left(\frac{d_{\min}(\omega)}{\sqrt{3}} \right)^r . \tag{49}$$

Case (b)

First we calculate u. To this end, we use the standard notation a, b and c for the sides of the

Fig. 1 Fermat point

triangle ABC. The angles α, β and γ belong to the vertices A, B and C. Let Δ denote the area of the triangle ABC. Note that $A'B = BC = a$. Using the law of cosine, we obtain

$$
\begin{aligned}
u^2 = (AA')^2 &= c^2 + a^2 - 2ac \cdot \cos(\beta + \frac{\pi}{3}) \\
&= c^2 + a^2 - 2ac \left(\frac{1}{2} \cos(\beta) - \frac{\sqrt{3}}{2} \sin(\beta) \right) \\
&= c^2 + a^2 - \frac{c^2 + a^2 - b^2}{2} + 2ac\sqrt{3} \frac{\Delta}{ac} \\
&= \frac{a^2 + b^2 + c^2}{2} + 2\sqrt{3}\Delta.
\end{aligned}
\tag{50}
$$

As before, the convexity of $t \to t^r$ yields

$$
V_{1,r}(U_\omega) \geq (\frac{u}{3})^r.
\tag{51}
$$

Note that

$$
\Delta \geq \frac{1}{2}(d_{\min}(\omega))^2 \cdot \max\left(\sin(\alpha), \sin(\beta), \sin(\gamma)\right).
$$

At least one angle of the triangle ABC lies in $[\frac{\pi}{3}, \frac{2\pi}{3}]$. Therefore, we get

$$
\Delta \geq \frac{1}{2}(d_{\min}(\omega))^2 \frac{1}{2}\sqrt{3}.
\tag{52}
$$

The combination of (50), (51) and (52) yields

$$V_{1,r}(U_\omega) \geq \left(\frac{\sqrt{\frac{a^2+b^2+c^2}{2} + 2\sqrt{3}\Delta}}{3} \right)^r \tag{53}$$

$$\geq \left(\frac{1}{3} \sqrt{\frac{3(d_{\min}(\omega))^2}{2} + 2\sqrt{3}\frac{1}{2}(d_{\min}(\omega))^2\frac{1}{2}\sqrt{3}} \right)^r$$

$$= \left(\frac{(d_{\min}(\omega))}{\sqrt{3}} \right)^r.$$

Thus, in both cases we get from (49) and (53),

$$D_{\min}(r,3,U_\omega) = V_{1,r}(U_\omega) - 2V_{2,r}(U_\omega)$$

$$\geq \left(\frac{d_{\min}(\omega)}{\sqrt{3}} \right)^r - \frac{4}{3}\left(\frac{d_{\min}(\omega)}{2} \right)^r,$$

which is strictly positive if $\frac{3}{4} > \left(\frac{\sqrt{3}}{2} \right)^r$. This is exactly the case if $r > 2$.

Now assume that $\mathrm{diam}(\omega) = d_{\min}(\omega)$.
From Lemma 4 we get $V_{1,r}(U_\omega) = (\frac{1}{\sqrt{3}})^r \mathrm{diam}^r(\omega)$, which yields equality in (47). If, on the other hand, (47) turns into an equation due to (49), only case (b) can happen. Thus (53), turns into an equation. As an immediate consequence, we obtain $a = b = c$, which implies $\mathrm{diam}(\omega) = d_{\min}(\omega)$.

Now, let $\omega = \{x_1, x_2, x_3\} \subset \mathbb{R}^2$ with $\mathrm{card}(\omega) = 3$. Assume that the points x_1, x_2, x_3 are not collinear. Let (c_k) be a sequence of contracting factors with $c_k \in]0, \frac{1}{2}]$ for every $k \in \mathbb{N}$. For $x \in \mathbb{R}^2$, the integers $k \in \mathbb{N}$ and $i \in \{1,2,3\}$ define the similitudes $S_{k,i}(x) = c_k x + (1 - c_k)x_i$. Let V be the interior of the convex hull of ω. Let μ be the uniform distribution on $E = E(\overline{V})$. We call E a modified Sierpinski gasket. Due to $\bar{c} \leq \frac{1}{2}$, the open set condition is satisfied. If $c_k = \frac{1}{2}$ for every k and ω consists of the vertices of an equilateral triangle, we call E the classical Sierpinski gasket. In the non-equilateral case, i.e. if $\mathrm{diam}(\omega) \neq d_{\min}(\omega)$, using Proposition 9 and Proposition 7 we can calculate a $d_1 > 0$, ensuring the applicability of Theorem 1, if $\bar{c} < d_1$. In the equilateral case we can determine such a $d_1 > 0$, too, but here we are restricted to $r > 2$. Note that in the self-similar case, provided $r = 2$ and $c < c_0$ for a small enough $c_0 > 0$, the optimal quantization errors $V_{n,2}(\mu)$ can be calculated explicitly for $n \in \{1,2,3\}$ by Theorem 5.4 in [18]. By means of Theorem 1, we thus can determine the optimal quantization errors $V_{n,2}(\mu)$ for every n in the non-equilateral case, provided $c < \min(c_0, d_1)$. Finally, we illustrate with this example that condition (MQED) cannot be dropped completely. To this end, let ω represent an equilateral triangle and let μ be self-similar with contracting factor c and $r \in]1, 2[$. Checking the proof of Proposition 9 it is easy to see that $D_{\min}(r, 3, U_\omega) < 0$. According to Remark 14, one gets $D_{\min}(r, 3, \mu) < 0$, provided c is smaller than d_1, which we will assume. Note that condition (SQ) is satisfied according to Proposition 6. Now, let $\{a\} \in C_{1,r}(\mu)$ and $\{b_1, b_2\} \in C_{2,r}(\mu)$. Let

$$\alpha' = \{S_1(a), S_2(a), S_{(3,1)}(a), S_{(3,2)}(a), S_{(3,3)}(a)\}$$

and

$$\alpha = \{S_1(a), S_2(\{b_1, b_2\}), S_3(\{b_1, b_2\})\}.$$

Because μ satisfies condition (SQ), Proposition 3 implies

$$\Psi_{\alpha',r}(\mu) - \Psi_{\alpha,r}(\mu) = c^r(V_{1,r}(\mu) + V_{3,r}(\mu) - 2V_{2,r}(\mu)) < 0.$$

Hence, $\alpha \notin C_{5,r}(\mu)$ and equation (42), respectively (32) gets wrong. Thus, the condition (MQED) cannot be dropped completely.

5.3 Further examples

Let μ be a uniform distribution on a Cantor-like set E, defined as in 3 by a set of similitudes with contracting factors c_k, $k \in \mathbb{N}$ and ω as the set of fixpoints with $\mathrm{card}(\omega) = N$.

According to Proposition 7, we can calculate a constant $d_1 > 0$ and apply Theorem 1, if $D_{\min}(r, N, U_\omega) > 0$ and $\bar{c} < d_1$. Additionally, as already mentioned in Example 5.2, in the self-similar case and $r = 2$ by combination of [18], Theorem 5.4 and Theorem 1 we can fix a $c_0 > 0$ and calculate all optimal quantization errors explicitly, if $\bar{c} < \min(c_0, d_1)$. Note that [18], Theorem 5.4 is only applicable, if the similitudes have no rotation part, i.e. the orthonormal transformation in (39) has to be the identity.

Thus, it remains to derive a criteria for ω, ensuring that $D_{\min}(r, N, U_\omega) > 0$. If $r = 2$, this question can always be answered by direct calculations, using the centroidal decomposition of $V_{n,2}(U_\omega)$ for $n = 1, .., N - 1$ (cf. [5], Example 3.5). If $r > 1$ in general, the situation becomes more difficult. In [16], Proposition 6.24 (ii), some special criteria are ensuring that $D_{\min}(r, N, U_\omega) > 0$. It needs further investigations to find general conditions in higher dimensions, ensuring that $D_{\min}(r, N, U_\omega) > 0$.

References

1. Bucklew, J.A., Wise, G.L.: Multidimensional asymptotic quantization theory with r-th power distortion measures. IEEE Trans. Inform. Theory **28**, 239-247 (1982)
2. Falconer, K. J.: Techniques in Fractal Geometry. Wiley, Chichester (1997)
3. Gersho, A., Gray, R.M.: Vector Quantization and Signal Compression. Kluwer, Boston (1992)
4. Graf, S., Luschgy, H.: The quantization of the Cantor Distribution. Math. Nachr. **183**, 113-133 (1997)
5. Graf, S., Luschgy, H.: Foundations of Quantization for Probability Distributions. Lecture Notes in Math. 1730, Springer (2000)
6. Graf, S., Luschgy, H.: The quantization dimension of self-similar probabilities. Math. Nachr. **241**, 103-109 (2002)
7. Graf, S., Luschgy, H.: Quantization for probability measures with respect to the geometric mean error. Math. Proc. Camb. Phil. Soc. **136**, 687-717 (2004)
8. Graf, S., Luschgy, H.: The point density measure in the quantization of self-similar probabilites. Math. Proc. Camb. Phil. Soc. **138**, 513-531 (2005)
9. Gray, R., Neuhoff, D.: Quantization. IEEE Trans. Inform. Theory **44**, 2325-2383 (1998)
10. Honsberger, R.: Mathematical gems I. From elementary combinatorics, number theory, and geometry. The Dolciani Mathematical Expositions (1973)
11. Hua, S.: On the dimension of generalized self-similar sets. Acta Math. Appl. Sinica **17**, 551-558 (1994)
12. Hua, S., Li, W.: Packing dimension of generalized Moran sets. Progress in Natural Science **6**, 148-152 (1996)
13. Hua, S., Wen, Z., Rao, H., Wu, J.: On the structures and dimensions of Moran sets. Sci. China Ser. A **43**, 836-852 (2000)
14. Hutchinson, J.E.: Fractals and self-similarity. Indiana Univ. Math. J. **30**, 713-747 (1981)
15. Kesseböhmer, M., Zhu, S.: Stability of quantization dimension and quantization for homogeneous Cantor measure. Math. Nachr. **280**, 866 - 881 (2007)

16. Kreitmeier, W.: Optimale Quantisierung verallgemeinerter Cantor-Verteilungen. Doctoral thesis. Univ. of Passau, Germany (2007)
    URN: http://nbn-resolving.de/urn:nbn:de:bvb:739-opus-913
17. Kreitmeier, W.: Optimal Quantization for Dyadic Homogeneous Cantor Distributions. Math. Nachr. **281**, in press (2008)
18. Kreitmeier, W.: Optimal quantization of probabilities concentrated on small balls. to appear in Forum Mathematicum.
19. Lindsay, L.J.: Quantization dimension for Probability Distributions. Doctoral thesis. Univ. of North Texas, Denton (2001)
20. Marion, J.: Mesure de Hausdorff d'un fractal à similitude interne. Ann. Sci. Math. Quebec **10**, 51-84 (1986)
21. Pötzelberger, K.: The quantization error of self-similar distributions. Math. Proc. Camb. Phil. Soc. **137**, 725-740 (2004)
22. Zador, P.L.: Development and evaluation of procedures for quantizing multivariate distributions. Doctoral thesis. Stanford Univ. (1963)
23. Zador, P.L.: Asymptotic quantization error of continuous signals and the quantization dimension. IEEE Trans. Inform. Theory **28**, 139-149 (1982)
24. Zhu, S.: Quantization dimension of probability measures supported on Cantor-like sets. J. Math. Anal. Appl. **338**, 742-750 (2008)