

Quantization and centroidal Voronoi tessellations for probability measures on dyadic Cantor sets

Mrinal Kanti Roychowdhury¹

Abstract. Quantization of a probability distribution is the process of estimating a given probability by a discrete probability that assumes only a finite number of levels in its support. Centroidal Voronoi tessellations (CVT) are Voronoi tessellations of a region such that the generating points of the tessellations are also the centroids of the corresponding Voronoi regions. In this paper, we investigate the optimal quantization and the centroidal Voronoi tessellations with n generators for a Borel probability measure P on \mathbb{R} supported by a dyadic Cantor set generated by two self-similar mappings with similarity ratios r , where $0 < r \leq \frac{5-\sqrt{17}}{2}$.

Mathematics Subject Classification (2010). 60Exx, 28A80, 94A34.

Keywords. Probability measure, Cantor set, quantization, centroidal Voronoi tessellation.

1. Introduction

In the context of communication theory, quantization is the process by which data is reduced to a simpler, more coarse representation which is more compatible with digital processing. Loosely speaking, quantization is the heart of analog to digital conversion. It is an area which has increased in importance in the last few decades due to the burgeoning advances in digital technology. Quantization for probability distributions refers to the idea of estimating a given probability measure by a discrete probability measure with finite support. We refer to [4, 9, 11] for surveys on the subject and comprehensive lists of references to the literature, see also [1, 3, 5, 6]. For mathematical treatment of quantization one is referred to Graf-Luschgy's book (see [7]). Let \mathbb{R}^d denote the d -dimensional Euclidean space, $\|\cdot\|$ denote the Euclidean norm on \mathbb{R}^d for any $d \geq 1$, and $n \in \mathbb{N}$. Then,

¹ The research of the author was supported by U.S. National Security Agency (NSA) Grant H98230-14-1-0320.

the n th *quantization error* for a Borel probability measure P on \mathbb{R}^d is defined by

$$V_n := V_n(P) = \inf \left\{ \int \min_{a \in \alpha} \|x - a\|^2 dP(x) : \alpha \subset \mathbb{R}^d, \text{card}(\alpha) \leq n \right\}, \quad (1)$$

where the infimum is taken over all subsets α of \mathbb{R}^d with $\text{card}(\alpha) \leq n$. If $\int \|x\|^2 dP(x) < \infty$, then there is some set α for which the infimum is achieved (see [5, 6, 7]). Such a set α for which the infimum occurs and contains no more than n points is called an *optimal set of n -means*, or *optimal set of n -quantizers*. If α is a finite set, in general, the error $\int \min_{a \in \alpha} \|x - a\|^2 dP(x)$ is often referred to as the *cost* or *distortion error* for α , and is denoted by $V(P; \alpha)$. Thus, $V_n := V_n(P) = \inf \{V(P; \alpha) : \alpha \subset \mathbb{R}^d, \text{card}(\alpha) \leq n\}$. It is known that for a continuous probability measure P an optimal set of n -means always has exactly n elements (see [7]). Given a finite subset $\alpha \subset \mathbb{R}^d$, the Voronoi region generated by $a \in \alpha$ is defined by

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

i.e., the Voronoi region generated by $a \in \alpha$ is the set of all elements in \mathbb{R}^d which are closer to a than to any other element in α , and the set $\{M(a|\alpha) : a \in \alpha\}$ is called the *Voronoi diagram* or *Voronoi tessellation* of \mathbb{R}^d with respect to α . A Borel measurable partition $\{A_a : a \in \alpha\}$ of \mathbb{R}^d is called a *Voronoi partition* of \mathbb{R}^d with respect to α (and P) if P -almost surely $A_a \subset M(a|\alpha)$ for every $a \in \alpha$. Notice that if $\alpha = \{a_1, a_2, \dots, a_n\}$ is an optimal set of n -means for P and $\{A_1, A_2, \dots, A_n\}$ is a Voronoi partition with respect to α , then $V_n = \sum_{i=1}^n \int_{A_i} \|x - a_i\|^2 dP(x)$. *Centroidal Voronoi tessellations* (CVTs) are Voronoi tessellations of a region such that the generating points of the tessellations are also the centroids of the corresponding Voronoi regions. A CVT with n generators, also called a CVT with n -means, associated with a probability measure P is called an *optimal centroidal Voronoi tessellation* (OCVT) if the n generators form an optimal set of n -means for P . Let us now state the following proposition (see [4, 7]).

Proposition 1.1. *Let α be an optimal set of n -means, $a \in \alpha$, and $M(a|\alpha)$ be the Voronoi region generated by $a \in \alpha$, i.e., $M(a|\alpha) = \{x \in \mathbb{R}^d : \|x - a\| = \min_{b \in \alpha} \|x - b\|\}$. Then, for every $a \in \alpha$,*

- (i) $P(M(a|\alpha)) > 0$,
- (ii) $P(\partial M(a|\alpha)) = 0$,
- (iii) $a = E(X : X \in M(a|\alpha))$, and
- (iv) P -almost surely the set $\{M(a|\alpha) : a \in \alpha\}$ forms a Voronoi partition of \mathbb{R}^d .

Remark 1.2. Let α be an optimal set of n -means and $a \in \alpha$, then by Proposition 1.1, we have

$$a = \frac{1}{P(M(a|\alpha))} \int_{M(a|\alpha)} x dP = \frac{\int_{M(a|\alpha)} x dP}{\int_{M(a|\alpha)} dP},$$

which implies that a is the centroid of the Voronoi region $M(a|\alpha)$ associated with the probability measure P (see also [2]). Thus, we can say that for a Borel probability measure P on \mathbb{R}^d , an optimal set of n -means forms a centroidal Voronoi tessellation of \mathbb{R}^d ; however, the converse is not true in general (see [2, 4]).

Let $S_1, S_2: \mathbb{R} \rightarrow \mathbb{R}$ be two contractive similarity mappings such that $S_1(x) = rx$ and $S_2(x) = rx + (1-r)$ for $0 < r < \frac{1}{2}$. Then, there exists a unique Borel probability measure P on \mathbb{R} such that $P = \frac{1}{2}P \circ S_1^{-1} + \frac{1}{2}P \circ S_2^{-1}$, where $P \circ S_i^{-1}$ denotes the image measure of P with respect to S_i for $i = 1, 2$ (see [10]). Such a P has support the Cantor set generated by the two mappings S_1 and S_2 . In this paper, in Section 3, we have given a centroidal Voronoi tessellation (CVT) for the probability measure P supported by the Cantor set generated by $S_1(x) = \frac{4}{9}x$ and $S_2(x) = \frac{4}{9}x + \frac{5}{9}$. The formula in this paper can be used to obtain a CVT for P on any Cantor set generated by $S_1(x) = rx$ and $S_2(x) = rx + (1-r)$, where $0.4364590141 \leq r \leq 0.4512271429$ (written up to ten decimal places). For the classical Cantor set C , i.e., when $r = \frac{1}{3}$, in the paper [8], Graf and Luschgy determined the optimal sets of n -means for the probability measure P for all $n \geq 2$. For a long time it was believed that using the same formula given in [8], one could determine the optimal sets of n -means for all $n \geq 2$ for the probability measure P supported by any Cantor set generated by the two mappings $S_1(x) = rx$ and $S_2(x) = rx + (1-r)$ for $0 < r \leq \frac{5-\sqrt{17}}{2}$, i.e., if $0 < r \leq 0.4384471872$. In Proposition 4.3 we have shown that it is not always true, by showing that if $0.4371985206 < r \leq 0.4384471872$ and n is not of the form $2^{\ell(n)}$ for any positive integer $\ell(n)$, then the distortion error of the CVT obtained using the formula in this paper is smaller than the distortion error of the CVT obtained using the formula given by Graf-Luschgy in [8]. In fact, in Section 5, we have further improved this bound which is given in Remark 5.3. In addition, the work in this paper shows that under squared error distortion measure, the centroid condition is not sufficient for optimal quantization for a singular continuous probability measure. Recall that the centroid condition is not sufficient for optimal quantization for an absolutely continuous probability measure is already known (see [2] and [4, Chapter 6]).

2. Preliminaries

By a *string* or a *word* σ over an alphabet $\{1, 2\}$, we mean a finite sequence $\sigma := \sigma_1\sigma_2\dots\sigma_k$ of symbols from the alphabet, where $k \geq 1$, and k is called the length of the word σ . A word of length zero is called the *empty word*, and is denoted by \emptyset . By $\{1, 2\}^*$, we denote the set of all words over the alphabet $\{1, 2\}$ of some finite length k including the empty word \emptyset . By $|\sigma|$, we denote the length of a word $\sigma \in \{1, 2\}^*$. For any two words $\sigma := \sigma_1\sigma_2\dots\sigma_k$ and $\tau := \tau_1\tau_2\dots\tau_\ell$ in $\{1, 2\}^*$, by $\sigma\tau := \sigma_1\dots\sigma_k\tau_1\dots\tau_\ell$, we mean the word obtained from the concatenation of the two words σ and τ . Let S_1 and S_2 be two contractive similarity mappings on \mathbb{R} given by $S_1(x) = \frac{4}{9}x$ and $S_2(x) = \frac{4}{9}x + \frac{5}{9}$. For $\sigma := \sigma_1\sigma_2\dots\sigma_k \in \{1, 2\}^k$, set $S_\sigma := S_{\sigma_1} \circ \dots \circ S_{\sigma_k}$ and $J_\sigma := S_\sigma([0, 1])$. For the empty word \emptyset , by S_\emptyset , it is meant the identity mapping on \mathbb{R} , and write $J := J_\emptyset = S_\emptyset([0, 1]) = [0, 1]$. Then, the set $C := \bigcap_{k \in \mathbb{N}} \bigcup_{\sigma \in \{1, 2\}^k} J_\sigma$ is known as the *Cantor set* generated by the two mappings S_1 and S_2 , and equals the support of the probability measure P given by $P = \frac{1}{2}P \circ S_1^{-1} + \frac{1}{2}P \circ S_2^{-1}$. For any $\sigma \in \{1, 2\}^*$, the intervals $J_{\sigma 1}$ and $J_{\sigma 2}$ into which J_σ is split up are called the basic intervals of J_σ . For $\sigma = \sigma_1\sigma_2\dots\sigma_k \in \{1, 2\}^*$, $k \geq 0$, write $p_\sigma := \frac{1}{2^k}$ and $s_\sigma := (\frac{4}{9})^k$.

Let X be a random variable with probability distribution P . By $E(X)$ and $V := V(X)$, we mean the expectation and the variance of the random variable X . For words $\beta, \gamma, \dots, \delta$ in $\{1, 2\}^*$, by $a(\beta, \gamma, \dots, \delta)$, we mean the conditional expectation of the random variable X given $J_\beta \cup J_\gamma \cup \dots \cup J_\delta$, i.e.,

$$\begin{aligned} a(\beta, \gamma, \dots, \delta) &= E(X \mid X \in J_\beta \cup J_\gamma \cup \dots \cup J_\delta) \\ &= \frac{1}{P(J_\beta \cup \dots \cup J_\delta)} \int_{J_\beta \cup \dots \cup J_\delta} x dP. \end{aligned} \quad (2)$$

Let us now give the following lemmas.

Lemma 2.1. *Let $f: \mathbb{R} \rightarrow \mathbb{R}^+$ be Borel measurable and $k \in \mathbb{N}$. Then*

$$\int f dP = \sum_{\sigma \in \{1, 2\}^k} \frac{1}{2^k} \int f \circ S_\sigma dP.$$

Proof. We know $P = \frac{1}{2}P \circ S_1^{-1} + \frac{1}{2}P \circ S_2^{-1}$, and so by induction

$$P = \sum_{\sigma \in \{1, 2\}^k} \frac{1}{2^k} P \circ S_\sigma^{-1},$$

and thus the lemma is yielded. \square

Lemma 2.2. $E(X) = \frac{1}{2}$ and $V := V(X) = \frac{5}{52}$, and for any $x_0 \in \mathbb{R}$,

$$\int (x - x_0)^2 dP(x) = V(X) + \left(x_0 - \frac{1}{2}\right)^2.$$

Proof. We have

$$\begin{aligned} E(X) &= \int x dP(x) \\ &= \frac{1}{2} \int \frac{4}{9} x dP(x) + \frac{1}{2} \int \left(\frac{4}{9} x + \frac{5}{9}\right) dP(x) \\ &= \frac{4}{18} E(X) + \frac{4}{18} E(X) + \frac{5}{18} \\ &= \frac{4}{9} E(X) + \frac{5}{18}, \end{aligned}$$

which implies $E(X) = \frac{1}{2}$;

$$\begin{aligned} E(X^2) &= \int x^2 dP(x) \\ &= \frac{1}{2} \int x^2 dP \circ S_1^{-1}(x) + \frac{1}{2} \int x^2 dP \circ S_2^{-1}(x) \\ &= \frac{1}{2} \int \frac{16}{81} x^2 dP(x) + \frac{1}{2} \int \left(\frac{4}{9} x + \frac{5}{9}\right)^2 dP(x) \\ &= \frac{16}{81} E(X^2) + \frac{20}{81} E(X) + \frac{25}{162} \\ &= \frac{16}{81} E(X^2) + \frac{20}{162} + \frac{25}{162}, \end{aligned}$$

which implies $E(X^2) = \frac{9}{26}$, and hence

$$V(X) = E(X - E(X))^2 = E(X^2) - (E(X))^2 = \frac{9}{26} - \left(\frac{1}{2}\right)^2 = \frac{5}{52}.$$

Then, following the standard theory of probability, we have

$$\int (x - x_0)^2 dP(x) = V(X) + (x_0 - E(X))^2,$$

and thus the lemma is yielded. \square

Corollary 2.3. *Let $\sigma \in \{1, 2\}^*$. Then, for any $x_0 \in \mathbb{R}$,*

$$\int_{J_\sigma} (x - x_0)^2 dP(x) = p_\sigma \left(s_\sigma^2 V + \left(S_\sigma \left(\frac{1}{2} \right) - x_0 \right)^2 \right). \quad (3)$$

Note 2.4. Notice that from the above lemma it follows that the optimal set of one-mean is the expected value and the corresponding quantization error is the variance V of the random variable X . For $\sigma \in \{1, 2\}^k$, $k \geq 1$, since $a(\sigma) = E(X: X \in J_\sigma)$, using Lemma 2.1, we have

$$\begin{aligned} a(\sigma) &= \frac{1}{P(J_\sigma)} \int_{J_\sigma} x dP(x) \\ &= \int_{J_\sigma} x dP \circ S_\sigma^{-1}(x) \\ &= \int S_\sigma(x) dP(x) \\ &= E(S_\sigma(X)). \end{aligned}$$

Since S_1 and S_2 are similarity mappings, it is easy to see that $E(S_j(X)) = S_j(E(X))$ for $j = 1, 2$ and so by induction,

$$a(\sigma) = E(S_\sigma(X)) = S_\sigma(E(X)) = S_\sigma \left(\frac{1}{2} \right)$$

for $\sigma \in \{1, 2\}^k$, $k \geq 1$.

In the next section, Proposition 3.3, Proposition 3.11 and Proposition 3.12 determine the centroidal Voronoi tessellations with n generators for the probability measure P for all $n \geq 2$.

3. Centroidal Voronoi tessellations for all $n \geq 2$

In this section, we determine the CVTs with n -means for each $n \geq 2$ of the Cantor set C generated by the two mappings S_1 and S_2 defined by $S_1(x) = \frac{4}{9}x$ and $S_2(x) = \frac{4}{9}x + \frac{5}{9}$ for $x \in \mathbb{R}$. As the probability distribution P has support the Cantor set C and $C \subset J$, a CVT of J with respect to the probability distribution is also a CVT of C and vice versa. Once we know a CVT, using the formula (3), the corresponding distortion error can easily be obtained. Write

$$A_\sigma := \{a(\sigma 11, \sigma 121, \sigma 1221), a(\sigma 1222, \sigma 21), a(\sigma 22)\},$$

or

$$A_\sigma := \{a(\sigma 11), a(\sigma 12, \sigma 2111), a(\sigma 2112, \sigma 212, \sigma 22)\}.$$

If σ is the empty word \emptyset , then we have

$$A := A_\emptyset = \{a(11, 121, 1221), a(1222, 21), a(22)\},$$

or

$$A := A_\emptyset = \{a(11), a(12, 2111), a(2112, 212, 22)\}.$$

Let us now prove the following lemma.

Lemma 3.1. *Similarity mappings preserve the ratio of the distances of a point from any other two points.*

Proof. Let $a, b, c \in \mathbb{R}$. For any $\sigma \in \{1, 2\}^k$ it is enough to prove that

$$\frac{|b - a|}{|c - b|} = \frac{|S_\sigma(b) - S_\sigma(a)|}{|S_\sigma(c) - S_\sigma(b)|},$$

which is clearly true, since

$$\frac{|S_\sigma(b) - S_\sigma(a)|}{|S_\sigma(c) - S_\sigma(b)|} = \frac{s_\sigma |b - a|}{s_\sigma |c - b|} = \frac{|b - a|}{|c - b|}. \quad \square$$

Remark 3.2. The two mappings S_1 and S_2 defined in this paper are increasing mappings, i.e., for any $x, y \in \mathbb{R}$, $x < y$ implies $S_i(x) < S_i(y)$ for $i = 1, 2$, and so by Lemma 3.1 if $a < b < c$, we have

$$(c - b)(S_\sigma(b) - S_\sigma(a)) = (b - a)(S_\sigma(c) - S_\sigma(b)).$$

Proposition 3.3. *Let $n \in \mathbb{N}$ and $n = 2^k$ for some $k \in \mathbb{N}$. Then, $\alpha_n = \{S_\sigma(\frac{1}{2}) : \sigma \in \{1, 2\}^k\}$ forms a unique optimal CVT with n -means with distortion error $V_n = (\frac{4}{9})^{2k} V$.*

Proof. By Remark 3.2, for any $\sigma \in \{1, 2\}^k$ we have

$$\left(1 - \frac{1}{2}\right) \left(S_\sigma\left(\frac{1}{2}\right) - S_\sigma(0)\right) = \left(\frac{1}{2} - 0\right) \left(S_\sigma(1) - S_\sigma\left(\frac{1}{2}\right)\right),$$

which implies $S_\sigma(\frac{1}{2}) = \frac{1}{2}(S_\sigma(0) + S_\sigma(1))$, i.e., $S_\sigma(\frac{1}{2})$ are the midpoints of the basic intervals J_σ for all $\sigma \in \{1, 2\}^k$. In addition, by Remark 1.2 and Note 2.4, $S_\sigma(\frac{1}{2})$ is the centroid of J_σ . Thus, the set $\{S_\sigma(\frac{1}{2}) : \sigma \in \{1, 2\}^k\}$ forms a CVT of the Cantor set. Moreover, by Corollary 2.3, for $a \in \mathbb{R}$, $\int_{J_\sigma} (x - a)^2 dP$ is minimum

when $a = S_\sigma(\frac{1}{2})$. Hence, the set α_n forms a unique optimal CVT of the Cantor set, and then

$$\begin{aligned} V_n &= \int \min_{a \in \alpha_n} \|x - a\|^2 dP \\ &= \sum_{\sigma \in \{1,2\}^k} \int_{J_\sigma} \min_{a \in \alpha} (x - a)^2 dP \\ &= \sum_{\sigma \in \{1,2\}^k} p_\sigma s_\sigma^2 V \\ &= \left(\frac{4}{9}\right)^{2k} V. \end{aligned}$$

This completes the proof of the proposition. □

Lemma 3.4. *Let*

$$A = \{a(11, 121, 1221), a(1222, 21), a(22)\}$$

or

$$A = \{a(11), a(12, 2111), a(2112, 212, 22)\}.$$

Then, A forms a CVT with three-means of the Cantor set C .

Proof. We have

$$\begin{aligned} S_{1221}(1) &= 0.395671 < \frac{1}{2}(a(11, 121, 1221) + a(1222, 21)) \\ &= 0.400854 \\ &< S_{1222}(0) \\ &= 0.405426 \end{aligned}$$

and

$$\begin{aligned} S_{21}(1) &= 0.753086 \\ &< \frac{1}{2}(a(1222, 21) + a(22)) \\ &= 0.754839 \\ &< S_{22}(0) \\ &= 0.802469. \end{aligned}$$

Thus, $A = \{a(11, 121, 1221), a(1222, 21), a(22)\}$ forms a CVT of C . Due to symmetry $A = \{a(11), a(12, 2111), a(2112, 212, 22)\}$ also forms a CVT of C . □

Lemma 3.5. *The set*

$$\{a(111, 1121, 11221), a(11222, 121), a(122), a(21), a(22)\}$$

forms a CVT with five-means.

Proof. We have

$$\begin{aligned} S_1^{-1}(\{a(111, 1121, 11221), a(11222, 121), a(122)\}) \\ = \{a(11, 121, 1221), a(1222, 21), a(22)\}. \end{aligned}$$

So, by Lemma 3.4, the set

$$S_1^{-1}(\{a(111, 1121, 11221), a(11222, 121), a(122)\})$$

forms a CVT with three-means. Similarly, by Proposition 3.3, the set

$$S_2^{-1}(\{a(21), a(22)\}) = \{a(1), a(2)\}$$

forms a CVT with two-means. By Lemma 3.1, we know that the similarity mappings preserve the ratio of the distances of a point from any other two points, and so the set $\{a(111, 1121, 11221), a(11222, 121), a(122)\}$ forms a CVT with three-means of J_1 and the set $\{a(21), a(22)\}$ forms a CVT with two-means of J_2 . Thus, the union of the CVTs of J_1 and J_2 will form a CVT of the Cantor set C if we can prove that

$$S_{122}(1) \leq \frac{1}{2}(a(122) + a(21)) \leq S_{21}(0),$$

which is clearly true since

$$S_{122}(1) = 0.444444 < \frac{1}{2}(a(122) + a(21)) = 0.527435 < S_{21}(0) = 0.555556.$$

Hence the given set forms a CVT with five-means. \square

Remark 3.6. Similarly, we can prove that the sets

$$\{a(111), a(112, 12111), a(12112, 1212, 122), a(21), a(22)\},$$

$$\{a(11), a(12), a(211, 2121, 21221), a(21222, 221), a(222)\},$$

$$\{a(11), a(12), a(211), a(212, 22111), a(22112, 2212, 222)\}$$

also form CVTs with five-means, i.e., the number of CVTs with five-means is four.

Lemma 3.7. *The set*

$$\{a(111, 1121, 11221), a(11222, 121), a(122), \\ a(211, 2121, 21221), a(21222, 221), a(222)\}$$

forms a CVT with six-means.

Proof. The set

$$\{a(11, 121, 1221), a(1222, 21), a(22)\}$$

forms a CVT of J with three-means. So, by Lemma 3.1, the sets

$$S_1(\{a(11, 121, 1221), a(1222, 21), a(22)\})$$

and

$$S_2(\{a(11, 121, 1221), a(1222, 21), a(22)\})$$

form CVTs of J_1 and J_2 , respectively. Thus, the given set will form a CVT with six-means if we can prove that

$$S_{122}(1) \leq \frac{1}{2}(a(122) + a(211, 2121, 21221)) \leq S_{211}(0),$$

which is clearly true since

$$\begin{aligned} S_{122}(1) &= 0.444444 \\ &< \frac{1}{2}(a(122) + a(211, 2121, 21221)) \\ &= 0.521 \\ &< S_{211}(0) \\ &= 0.555556. \end{aligned}$$

Thus, the lemma is yielded. \square

Remark 3.8. By Lemma 3.4, since there are two different CVTs of J with three-means, one can say that each of the basic intervals J_1 and J_2 has two different CVTs, and thus using all possible combinations one can see that the total number of CVTs with six-means is four.

Lemma 3.9. *The set*

$$\{a(111, 1121, 11221), a(11222, 121), a(122), a(211), a(212), a(221), a(222)\}$$

forms a CVT with seven-means.

Proof. The set

$$\{a(11, 121, 1221), a(1222, 21), a(22)\}$$

forms a CVT of J with three-means. So, by Lemma 3.1, the set

$$\{a(111, 1121, 11221), a(11222, 121), a(122)\}$$

forms a CVT of J_1 with three-means. Again by Proposition 3.3, the set

$$\{a(211), a(212), a(221), a(222)\}$$

forms a CVT of J_2 with four-means. Hence, the union of the two CVTs will form a CVT with seven-means if we can prove that

$$S_{122}(1) \leq \frac{1}{2}(a(122) + a(211)) \leq S_{211}(0),$$

which is clearly true since

$$S_{122}(1) = 0.444444 < \frac{1}{2}(a(122) + a(211)) = 0.5 < S_{211}(0) = 0.555556.$$

Thus, the lemma is obtained. \square

Remark 3.10. Each J_i for $i = 1, 2$, has two different CVTs, and so using all possible combinations we see that the total number of CVTs with seven-means is four.

Let us now prove the following two propositions.

Proposition 3.11. *Let $n \in \mathbb{N}$ be such that $n = 2^{\ell(n)} + k$, where $1 \leq k \leq 2^{\ell(n)-1}$. Let $I \subseteq \{1, 2\}^{\ell(n)-1}$ with $\text{card}(I) = k = n - 2^{\ell(n)}$. Then, the set $\alpha_n := \alpha_n(I)$, where*

$$\alpha_n(I) = \begin{cases} \bigcup_{\sigma \in I} A_\sigma \cup \left\{ S_{\sigma 1} \left(\frac{1}{2} \right), S_{\sigma 2} \left(\frac{1}{2} \right) : \sigma \in \{1, 2\}^{\ell(n)-1} \setminus I \right\} & \text{if } 1 \leq k < 2^{\ell(n)-1}, \\ \bigcup_{\sigma \in I} A_\sigma & \text{if } k = 2^{\ell(n)-1} \neq 1, \\ A_\emptyset & \text{if } k = 2^{\ell(n)-1} = 1, \text{ i.e., when } n = 3, \end{cases}$$

forms a CVT with n -means. The number of CVTs for $1 \leq k < 2^{\ell(n)-1}$ is $2^{n-2^{\ell(n)}} \times 2^{\ell(n)-1} C_{n-2^{\ell(n)}}$, and the number of CVTs for $k = 2^{\ell(n)-1}$ is $2^{2^{\ell(n)-1}}$.

Proof. Let us first assume that $n = 2^{\ell(n)} + k$, where $1 \leq k < 2^{\ell(n)-1}$. Let $I \subseteq \{1, 2\}^{\ell(n)-1}$ with $\text{card}(I) = n - 2^{\ell(n)}$. By Lemma 3.4, for each $\sigma \in I$, the set $S_{\sigma}^{-1}(A_{\sigma})$ forms a CVT of J with three-means, and so by Lemma 3.1 the set A_{σ} forms a CVT of J_{σ} with three-means. Now, proceeding in the similar way as Lemma 3.7, we can say that the set $\bigcup_{\sigma \in I} A_{\sigma}$ forms a CVT of $\bigcup_{\sigma \in I} J_{\sigma}$. Again, for each $\sigma \in \{1, 2\}^{\ell(n)-1} \setminus I$ the set $S_{\sigma}^{-1}(\{S_{\sigma 1}(\frac{1}{2}), S_{\sigma 2}(\frac{1}{2})\})$ forms a CVT of J with two-means, and so by Lemma 3.1 the set $\{S_{\sigma 1}(\frac{1}{2}), S_{\sigma 2}(\frac{1}{2})\}$ forms a CVT of J_{σ} with two-means. Now, proceeding in the similar way as Lemma 3.5, we can say that the set $\bigcup_{\sigma \in I} A_{\sigma} \cup \{S_{\sigma 1}(\frac{1}{2}), S_{\sigma 2}(\frac{1}{2}) : \sigma \in \{1, 2\}^{\ell(n)-1} \setminus I\}$ forms a CVT with n -means. Notice that $\text{card}(I) = n - 2^{\ell(n)}$ and I can be chosen in $2^{\ell(n)-1} C_{n-2^{\ell(n)}}$ ways. For each $\sigma \in I$ there are two different choices for A_{σ} . Hence, the number of CVTs in this case is $2^{n-2^{\ell(n)}} \times 2^{\ell(n)-1} C_{n-2^{\ell(n)}}$. Similarly, by Lemma 3.1, Lemma 3.4 and proceeding in the similar way as Lemma 3.7, we can prove that if $n = 2^{\ell(n)} + k$, where $k = 2^{\ell(n)-1} \neq 1$ or $k = 2^{\ell(n)-1} = 1$, the set $\bigcup_{\sigma \in I} A_{\sigma}$ forms a CVT of C , and the number of CVTs in either case is given by $2^{2^{\ell(n)-1}}$. Hence, the proposition is yielded. \square

Proposition 3.12. *Let $n \in \mathbb{N}$ be such that $n = 3 \cdot 2^{\ell(n)-1} + k$, where $1 \leq k \leq 2^{\ell(n)-1} - 1$. Let $I \subseteq \{1, 2\}^{\ell(n)-1}$ with $\text{card}(I) = n - 3 \cdot 2^{\ell(n)-1}$. Then, the set $\alpha_n := \alpha_n(I)$, where*

$$\alpha_n(I) = \left(\bigcup_{\sigma \in \{1, 2\}^{\ell(n)-1} \setminus I} A_{\sigma} \right) \cup \left\{ S_{\sigma \tau} \left(\frac{1}{2} \right) : \sigma \in I \text{ and } \tau \in \{1, 2\}^2 \right\},$$

forms a CVT with n -means. The number of such sets is

$$2^{2^{\ell(n)+1-n}} \times 2^{\ell(n)-1} C_{n-3 \cdot 2^{\ell(n)-1}}.$$

Proof. Let $n = 3 \cdot 2^{\ell(n)-1} + k$, where $1 \leq k \leq 2^{\ell(n)-1} - 1$. Let $I \subseteq \{1, 2\}^{\ell(n)-1}$ with $\text{card}(I) = n - 3 \cdot 2^{\ell(n)-1}$. For each $\sigma \in \{1, 2\}^{\ell(n)-1} \setminus I$ we have $S_{\sigma}^{-1}(A_{\sigma}) = A$, which is a CVT of C with three-means, and so the set A_{σ} forms a CVT of J_{σ} for each $\sigma \in \{1, 2\}^{\ell(n)-1} \setminus I$. Again, the set $\{J_{\sigma \tau} : \tau \in \{1, 2\}^2\}$ forms a CVT with four-means of J_{σ} for each $\sigma \in I$. Thus, proceeding in the similar way as Lemma 3.9, we can prove that the set

$$\alpha_n(I) = \left(\bigcup_{\sigma \in \{1, 2\}^{\ell(n)-1} \setminus I} A_{\sigma} \right) \cup \left\{ S_{\sigma \tau} \left(\frac{1}{2} \right) : \sigma \in I \text{ and } \tau \in \{1, 2\}^2 \right\}$$

forms a CVT of C with n -means. Notice that I can be chosen in $2^{\ell(n)-1} C_{n-3 \cdot 2^{\ell(n)-1}}$ ways and $\text{card}(\{1, 2\}^{\ell(n)-1} \setminus I) = 2^{\ell(n)-1} - (n - 3 \cdot 2^{\ell(n)-1}) = 2^{\ell(n)+1} - n$.

For each $\sigma \in \{1, 2\}^{\ell(n)-1} \setminus I$ there are two different choices for A_σ . Hence, the number of CVTs in this case is $2^{2^{\ell(n)+1}-n} \times 2^{\ell(n)-1} C_{n-3, 2^{\ell(n)-1}}$. Thus, the proposition is yielded. \square

Let us now prove the following lemma.

Lemma 3.13. *Let $P = \frac{1}{2}P \circ S_1^{-1} + \frac{1}{2}P \circ S_2^{-1}$ be the probability measure supported by the Cantor set generated by $S_1(x) = rx$ and $S_2(x) = rx + (1-r)$. Let $n \in \mathbb{N}$, $n \geq 2$ and n is not of the form $2^{\ell(n)}$ for any $\ell(n) \in \mathbb{N}$. Then, $\alpha_n(I)$, given by Proposition 3.11 or Proposition 3.12, for each $n \geq 2$ forms a CVT if*

$$0.4364590141 \leq r \leq 0.4512271429$$

(written up to ten decimal places).

Proof. Let

$$\alpha_3(I) = \{a(11, 121, 1221), a(1222, 21), a(22)\}.$$

It forms a CVT if

$$S_{1221}(1) \leq \frac{1}{2}(a(11, 121, 1221) + a(1222, 21)) \leq S_{1222}(0)$$

and

$$S_{21}(1) \leq \frac{1}{2}(a(1222, 21) + a(22)) \leq S_{22}(0),$$

i.e., if

$$0.4364590141 \leq r \leq 0.4521904271$$

and

$$0.2076973455 \leq r \leq 0.4512271429,$$

which yields

$$0.4364590141 \leq r \leq 0.4512271429.$$

Similarly, if

$$\alpha_3(I) = \{a(11), a(12, 2111), a(2112, 212, 22)\},$$

it will form a CVT if

$$0.4364590141 \leq r \leq 0.4512271429.$$

By Lemma 3.1, we can say that the set $\alpha_n(I)$ for each $n \geq 2$ also forms a CVT if $0.4364590141 \leq r \leq 0.4512271429$, and thus the lemma is yielded. \square

Remark 3.14. Proposition 3.3, Proposition 3.11 and Proposition 3.12 give the CVTs with n -means for the probability distribution P supported by the Cantor set generated by the mappings $S_1(x) = \frac{4}{9}x$ and $S_2(x) = \frac{4}{9}x + \frac{5}{9}$ for any positive integer $n \geq 2$. Lemma 3.13 says that using the formula given in this paper, if n is not of the form $2^{\ell(n)}$ for any $\ell(n) \in \mathbb{N}$, one can determine the CVTs with n -means for any $n \geq 2$, and hence the corresponding distortion error, for the probability measure P supported by any Cantor set generated by $S_1(x) = rx$ and $S_2(x) = rx + (1-r)$, where $0.4364590141 \leq r \leq 0.4512271429$.

4. Distortion errors for two different CVTs

In this section we compare the distortion errors for two different CVTs with n -means: one is obtained using the formula given by Proposition 3.3, Proposition 3.11, or Proposition 3.12 in this paper, and one is obtained using the formula given in [8]. Let $P = \frac{1}{2}P \circ S_1^{-1} + \frac{1}{2}P \circ S_2^{-1}$ be the probability measure supported by the Cantor set generated by $S_1(x) = rx$ and $S_2(x) = rx + (1-r)$. Then, it can be shown that if V is the variance of a random variable with distribution P in this case, then $V = \frac{1-r}{4(r+1)}$.

Definition 4.1. For $n \in \mathbb{N}$ with $n \geq 2$ let $\ell(n)$ be the unique natural number with $2^{\ell(n)} \leq n < 2^{\ell(n)+1}$. For $I \subset \{1, 2\}^{\ell(n)}$ with $\text{card}(I) = n - 2^{\ell(n)}$ let $\beta_n(I)$ be the set consisting of all midpoints a_σ of intervals J_σ with $\sigma \in \{1, 2\}^{\ell(n)} \setminus I$ and all midpoints $a_{\sigma_1}, a_{\sigma_2}$ of the basic intervals of J_σ with $\sigma \in I$. Formally,

$$\beta_n(I) = \{a_\sigma : \sigma \in \{1, 2\}^{\ell(n)} \setminus I\} \cup \{a_{\sigma_1} : \sigma \in I\} \cup \{a_{\sigma_2} : \sigma \in I\}.$$

In [8], it was shown that $\beta_n(I)$ forms an optimal set of n -means for $r = \frac{1}{3}$. Let us now prove the following lemma.

Lemma 4.2. *Let $\beta_n(I)$ be the set given by Definition 4.1. Then, $\beta_n(I)$ forms a CVT with n -means for each $n \geq 2$ if $0 < r \leq \frac{5-\sqrt{17}}{2}$, i.e., if $0 < r \leq 0.4384471872$ (written up to ten decimal places).*

Proof. Let a_{11} be the midpoint of J_{11} , a_{12} be the midpoint of J_{12} , and a_2 be the midpoint of J_2 . Then, $\beta_3(\{1\}) = \{a_{11}, a_{12}, a_2\}$, and it will form a CVT if $S_{12}(1) \leq \frac{1}{2}(a_{12} + a_2) \leq S_2(0)$, which implies $r \leq \frac{1}{2}(-r^2 + r + 2) \leq 1 - r$, which after simplification yields $0 < r \leq \frac{5-\sqrt{17}}{2}$, i.e., $0 < r \leq 0.4384471872$. Thus, by Lemma 3.1, it can be seen that if $0 < r \leq 0.4384471872$, $\beta_n(I)$ for each $n \geq 2$ also forms a CVT, and thus the lemma is yielded. \square

Let us now prove the following proposition.

Proposition 4.3. *Let $\alpha_n(I)$ be the set as defined by Proposition 3.3, Proposition 3.11, or Proposition 3.12, and $\beta_n(I)$ be the set given by Definition 4.1. Suppose n is not of the form $2^{\ell(n)}$ for any positive integer $\ell(n)$. Then, $V(P, \alpha_n(I)) < V(P, \beta_n(I))$ if $0.4371985206 < r \leq 0.4384471872$, where $V(P, \alpha_n(I))$ and $V(P, \beta_n(I))$ respectively denote the distortion errors for the sets $\alpha_n(I)$ and $\beta_n(I)$.*

Proof. If n is of the form $2^{\ell(n)}$ for some positive integer $\ell(n)$, then it is easy to see that $V(P, \alpha_n(I)) = V(P, \beta_n(I))$. Let us assume that n is not of the form $2^{\ell(n)}$ for any positive integer $\ell(n)$. To prove $V(P, \alpha_n(I)) < V(P, \beta_n(I))$, due to Lemma 3.1, it is enough to prove the inequality for $n = 3$, then it will be satisfied for all other values $n = 5, 6, 7, 9, 10$, etc., which are not of the form $2^{\ell(n)}$. Notice that in $\alpha_3(I)$ as defined in Proposition 3.11, the set I is an empty set. By Lemma 3.4, take $\alpha_3(I) = \{a(11, 121, 1221), a(1222, 21), a(22)\}$. Then, using (1) and (3), we have

$$\begin{aligned}
 V(P, \alpha_3(I)) &= V(P, \{a(11, 121, 1221), a(1222, 21), a(22)\}) \\
 &= \int_{J_{11} \cup J_{121} \cup J_{1221}} (x - a(11, 121, 1221))^2 dP(x) \\
 &\quad + \int_{J_{1222} \cup J_{21}} (x - a(1222, 21))^2 dP(x) \\
 &\quad + \int_{J_{22}} (x - a(22))^2 dP(x) \\
 &= \int_{J_{11}} (x - a(11, 121, 1221))^2 dP(x) \\
 &\quad + \int_{J_{121}} (x - a(11, 121, 1221))^2 dP(x) \\
 &\quad + \int_{J_{1221}} (x - a(11, 121, 1221))^2 dP(x) \\
 &\quad + \int_{J_{1222}} (x - a(1222, 21))^2 dP(x) \\
 &\quad + \int_{J_{21}} (x - a(1222, 21))^2 dP(x) \\
 &\quad + \int_{J_{22}} (x - a(22))^2 dP(x),
 \end{aligned}$$

which implies

$$\begin{aligned}
 V(P, \alpha_3(I)) &= \frac{1}{2^2}(r^4V + (a(11) - a(11, 121, 1221))^2) \\
 &\quad + \frac{1}{2^3}(r^6V + (a(121) - a(11, 121, 1221))^2) \\
 &\quad + \frac{1}{2^4}(r^8V + (a(1221) - a(11, 121, 1221))^2) \\
 &\quad + \frac{1}{2^4}(r^8V + (a(1222) - a(1222, 21))^2) \\
 &\quad + \frac{1}{2^2}(r^4V + (a(21) - a(1222, 21))^2) \\
 &\quad + \frac{1}{2^2}r^4V.
 \end{aligned}$$

Now, use (2), and simplify to obtain

$$V(P, \alpha_3(I)) = \frac{1}{560(r+1)}(-3r^9 - 3r^8 + 14r^7 - 22r^6 - 71r^5 + 49r^4 + 4r^3 + 88r^2 - 84r + 28). \quad (4)$$

To calculate $V(P, \beta_3(I))$ we take $I = \{1\}$, then $\beta_3(I) = \{a(11), a(12), a(2)\}$. Thus,

$$\begin{aligned}
 &V(P, \beta_3(I)) \\
 &= \int_{J_{11}} (x - a(11))^2 dP + \int_{J_{12}} (x - a(12))^2 dP + \int_{J_2} (x - a(2))^2 dP \\
 &= \frac{1}{2}r^4V + \frac{1}{2}r^2V.
 \end{aligned}$$

We see that $V(P, \alpha_3(I)) < V(P, \beta_3(I))$ if $0.4371985206 < r$. Combining this with the values of r in Lemma 4.2, we see that $V(P, \alpha_3(I)) < V(P, \beta_3(I))$ if $0.4371985206 < r \leq 0.4384471872$, which yields the proposition. \square

Remark 4.4. Proposition 4.3 says that if $0.4371985206 < r \leq 0.4384471872$ and n is not of the form $2^{\ell(n)}$ for any positive integer $\ell(n)$, then the distortion error for the CVT $\alpha_n(I)$ obtained in this paper is less than the distortion error for the CVT obtained using the formula in [8]. But, until now it is not known whether this $\alpha_n(I)$ forms an optimal CVT with n -means for $0.4371985206 < r \leq 0.4384471872$. In the following section, in Theorem 5.2, we give an answer of it.

5. The CVT $\alpha_n(I)$ does not form an optimal CVT

In this section, we show that if n is not of the form $2^{\ell(n)}$ for any positive integer $\ell(n)$, then the CVT $\alpha_n(I)$ does not form an optimal CVT for

$$0.4371985206 < r \leq \frac{5 - \sqrt{17}}{2} \approx 0.4384471872.$$

Write

$$C_\sigma := \{a(\sigma 11, \sigma 1211, \sigma 12121), a(\sigma 12122, \sigma 122, \sigma 211), a(\sigma 212, \sigma 22)\},$$

or

$$C_\sigma := \{a(\sigma 11, \sigma 121), a(\sigma 122, \sigma 211, \sigma 21211), a(\sigma 21212, \sigma 2122, \sigma 22)\}.$$

If σ is the empty word \emptyset , then we have

$$C := C_\emptyset = \{a(11, 1211, 12121), a(12122, 122, 211), a(212, 22)\},$$

or

$$C := C_\emptyset = \{a(11, 121), a(122, 211, 21211), a(21212, 2122, 22)\}.$$

Let $n \in \mathbb{N}$. If $n = 2^{\ell(n)} + k$, where $1 \leq k \leq 2^{\ell(n)-1}$, then write

$$\delta_n(I) = \begin{cases} \bigcup_{\sigma \in I} C_\sigma \cup \left\{ S_{\sigma 1} \left(\frac{1}{2} \right), S_{\sigma 2} \left(\frac{1}{2} \right) : \sigma \in \{1, 2\}^{\ell(n)-1} \setminus I \right\} & \text{if } 1 \leq k < 2^{\ell(n)-1}, \\ \bigcup_{\sigma \in I} C_\sigma & \text{if } k = 2^{\ell(n)-1} \neq 1, \\ C_\emptyset \text{ if } k = 2^{\ell(n)-1} = 1, & \text{i.e., when } n = 3, \end{cases}$$

where $I \subseteq \{1, 2\}^{\ell(n)-1}$ with $\text{card}(I) = k = n - 2^{\ell(n)}$. If $n = 3 \cdot 2^{\ell(n)-1} + k$, where $1 \leq k \leq 2^{\ell(n)-1} - 1$, then write

$$\delta_n(I) = \left(\bigcup_{\sigma \in \{1, 2\}^{\ell(n)-1} \setminus I} C_\sigma \right) \cup \left\{ S_{\sigma \tau} \left(\frac{1}{2} \right) : \sigma \in I \text{ and } \tau \in \{1, 2\}^2 \right\},$$

where $I \subseteq \{1, 2\}^{\ell(n)-1}$ with $\text{card}(I) = k = n - 3 \cdot 2^{\ell(n)-1}$.

We now prove the following proposition.

Proposition 5.1. *Let $0.4364590141 \leq r \leq 0.4486234903$. Then, both $\delta_n(I)$ and $\alpha_n(I)$ form CVTs. Moreover, if n is not of the form $2^{\ell(n)}$ for any positive integer $\ell(n)$, then $V(P, \delta_n(I)) < V(P, \alpha_n(I))$ if $0.4364590141 \leq r \leq 0.4486234903$, where $V(P, \delta_n(I))$ and $V(P, \alpha_n(I))$ respectively denote the distortion errors for the CVTs $\delta_n(I)$ and $\alpha_n(I)$.*

Proof. Let us first find the values of r for which $\delta_n(I)$ forms a CVT. Proceeding in the similar way as Lemma 3.13, we see that $\delta_n(I)$ forms a CVT if $0.4332840530 \leq r \leq 0.4486234903$. Moreover, by Lemma 3.13, $\alpha_n(I)$ forms a CVT if $0.4364590141 \leq r \leq 0.4512271429$. Thus, both $\delta_n(I)$ and $\alpha_n(I)$ forms a CVT if $0.4364590141 \leq r \leq 0.4486234903$. Now, to find the values of r for which $V(P, \delta_n(I)) < V(P, \alpha_n(I))$, we proceed in the similar way as the proof of Proposition 4.3. Take $\delta_3(I) = \{a(11, 1211, 12121), a(12122, 122, 211), a(212, 22)\}$. Then, using (1) and (3), we have

$$\begin{aligned} V(P, \delta_3(I)) &= \frac{1}{2^2}(r^4V + (a(11) - a(11, 1211, 12121))^2) \\ &\quad + \frac{1}{2^4}(r^8V + (a(1211) - a(11, 1211, 12121))^2) \\ &\quad + \frac{1}{2^5}(r^{10}V + (a(12121) - a(11, 1211, 12121))^2) \\ &\quad + \frac{1}{2^5}(r^{10}V + (a(12122) - a(12122, 122, 211))^2) \\ &\quad + \frac{1}{2^3}(r^6V + (a(122) - a(12122, 122, 211))^2) \\ &\quad + \frac{1}{2^3}(r^6V + (a(211) - a(12122, 122, 211))^2) \\ &\quad + \frac{1}{2^3}(r^6V + (a(212) - a(212, 22))^2) \\ &\quad + \frac{1}{2^2}(r^4V + (a(22) - a(212, 22))^2). \end{aligned}$$

Then, using (2),

$$\begin{aligned} V(P, \delta_3(I)) &= -\frac{1}{3168(r+1)}(5r^{11} + 5r^{10} - 2r^9 + 18r^8 + 89r^7 + 21r^6 \\ &\quad + 180r^5 - 48r^4 - 140r^3 - 568r^2 + 660r - 220). \end{aligned}$$

Equation (4) gives $V_3(P, \alpha_3(I))$. Thus, we see that $V(P, \delta_3(I)) < V(P, \alpha_3(I))$ if $0.4307442489 < r$. Combining this with the values of r for which both $\delta_3(I)$ and $\alpha_3(I)$ simultaneously form a CVT, we see that $V(P, \gamma_3(I)) < V(P, \alpha_3(I))$ if $0.4364590141 \leq r \leq 0.4486234903$, which yields the proposition. \square

Let us now give the following theorem.

Theorem 5.2. *Let $n \in \mathbb{N}$ be such that n is not of the form $2^{\ell(n)}$ for any $\ell(n) \in \mathbb{N}$. Let $\alpha_n(I)$ be the set as defined in Section 3. Then, $\alpha_n(I)$ does not form an optimal CVT for $0.4371985206 < r \leq 0.4384471872$.*

Proof. By Proposition 5.1 we see that for $0.4364590141 \leq r \leq 0.4486234903$, both $\delta_n(I)$ and $\alpha_n(I)$ form CVTs, and $V(P, \delta_n(I)) < V(P, \alpha_n(I))$. Thus, $\alpha_n(I)$ does not form an optimal CVT for $0.4371985206 < r \leq 0.4384471872$, which is the theorem. \square

Remark 5.3. Comparing Proposition 4.3 and Proposition 5.1, if n is not of the form $2^{\ell(n)}$ for any positive integer $\ell(n)$, we can say that if $0.4364590141 \leq r \leq \frac{5-\sqrt{17}}{2} \approx 0.4384471872$, then the CVT $\beta_n(I)$, which is obtained using the formula given in [8] does not form an optimal CVT. The least upper bound of r for which $\beta_n(I)$ forms an optimal CVT is still unknown. The investigation of it will appear elsewhere.

References

- [1] E. F. Abaya and G. L. Wise, Some remarks on the existence of optimal quantizers. *Stat. Probab. Lett.* **2** (1984), 349–351. [Zbl 0557.62005](#)
- [2] Q. Du, V. Faber, and M. Gunzburger, Centroidal Voronoi tessellations: applications and algorithms. *SIAM Rev.* **41** (1999), no. 4, 637–676. [MR 1722997](#) [Zbl 0983.65021](#)
- [3] C. P. Dettmann and M. K. Roychowdhury, Quantization for uniform distributions on equilateral triangles. *Real Anal. Exchange* **42** (2017), no. 1, 149–166.
- [4] A. Gersho and R. M. Gray, *Vector quantization and signal compression*. Kluwer Academy publishers, Boston, MA, 1992. [Zbl 0782.94001](#)
- [5] R. M. Gray, J. C. Kieffer, and Y. Linde, Locally optimal block quantizer design. *Inform. and Control* **45** (1980), no. 2, 178–198. [MR 0584832](#) [Zbl 0452.94011](#)
- [6] A. György and T. Linder, On the structure of optimal entropy-constrained scalar quantizers. *IEEE Trans. Inform. Theory* **48** (2002), no. 2, 416–427. [MR 1891255](#) [Zbl 1071.94526](#)
- [7] S. Graf and H. Luschgy, *Foundations of quantization for probability distributions*. Lecture Notes in Mathematics, 1730. Springer-Verlag, Berlin, 2000. [MR 1764176](#) [Zbl 0951.60003](#)
- [8] S. Graf and H. Luschgy, The Quantization of the Cantor Distribution. *Math. Nachr.* **183** (1997), 113–133. [MR 1434978](#) [Zbl 0874.28013](#)
- [9] R. Gray and D. Neuhoff, Quantization. *IEEE Trans. Inform. Theory* **44** (1998), no. 6, 2325–2383. [MR 1658787](#) [Zbl 1016.94016](#)
- [10] J. Hutchinson, Fractals and self-similarity. *Indiana Univ. Math. J.* **30** (1981), no. 5, 713–747. [MR 0625600](#) [Zbl 0598.28011](#)
- [11] R. Zam, *Lattice coding for signals and networks*. A structured coding approach to quantization, modulation, and multiuser information theory. Cambridge University Press, Cambridge, 2014.

Received September 22, 2015

Mrinal Kanti Roychowdhury, School of Mathematical and Statistical Sciences,
University of Texas Rio Grande Valley, 1201 West University Drive, Edinburg,
TX 78539-2999, USA

e-mail: mrinal.roychowdhury@utrgv.edu