http://topology.auburn.edu/tp/

TOPOLOGY PROCEEDINGS

Volume 56, 2020

Pages 195–218

http://topology.nipissingu.ca/tp/

Quantization for Uniform Distributions of Cantor Dusts on $\mathbb{R}^2$

by

Doğan Çömez and Mrinal Kanti Roychowdhury

Electronically published on November 16, 2019

Topology Proceedings

Web:	http://topology.auburn.edu/tp/
Mail:	Topology Proceedings Department of Mathematics & Statistics Auburn University, Alabama 36849, USA
E-mail:	topolog@auburn.edu
ISSN:	(Online) 2331-1290, (Print) 0146-4124

http://topology.auburn.edu/tp/
http://topology.nipissingu.ca/tp/

TOPOLOGY PROCEEDINGS
Volume 56 (2020)
Pages 195-218

E-Published on November 16, 2019

QUANTIZATION FOR UNIFORM DISTRIBUTIONS OF CANTOR DUSTS ON $\mathbb{R}^2$

DOĞAN ÇÖMEZ AND MRINAL KANTI ROYCHOWDHURY

Abstract. Let P be a Borel probability measure on $\mathbb{R}^2$ supported by the Cantor dusts generated by a set of 4^u, $u \geq 1$, contractive similarity mappings satisfying the strong separation condition. For this probability measure, we determine the optimal sets of n-means and the n^{th} quantization errors for all $n \geq 2$. In addition, it is shown that though the quantization dimension of the measure P is known, the quantization coefficient for P does not exist.

1. Introduction

Quantization for a probability distribution is the process of estimating it by a discrete probability that assumes only a finite number of levels in its support. For an in-depth analysis of quantization of probability measures, one may consult the excellent source by Sigfried Graf and Harald Luschgy [7] and the sources [1], [8], [9], [10], and [15], to name a few. In this paper, we are interested in the quantization of a particular type of continuous singular self-similar probability measures.

Let $\mathbb{R}^d$ denote the d-dimensional Euclidean space with the Euclidean norm $\| \ \|$. For any $d \geq 1$ and $n \in \mathbb{N}$, the n^{th} *quantization error* for a Borel probability measure P on $\mathbb{R}^d$ is defined by

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

2010 *Mathematics Subject Classification.* Primary 28A80; Secondary 94A34, 60Exx.

Key words and phrases. optimal quantizer, probability measure, quantization dimension, quantization error, Sierpiński carpet.

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

If $\int \|x\|^2 dP(x) < \infty$, then there is some set α for which the infimum is achieved ([6], [7], [8]). Such a set α for which the infimum occurs and contains no more than n points is called an *optimal set of n-means*. The elements in an optimal set of n-means are called *optimal 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 *distortion error* for α and is denoted by $V(P;\alpha)$. Then the n^{th} *quantization error for* P is defined by $V_n := V_n(P) = \inf\{V(P;\alpha) : \alpha \subset \mathbb{R}^d,\ \text{card}(\alpha) \leq n\}$. It is known that for a continuous probability measure P, an optimal set of n-means always has exactly n elements [7]. Naturally, $\lim_n V_n(P) = 0$; hence, one would like to know more about the rate of convergence. To know some recent results in the direction of optimal sets of n-means and the n^{th} quantization errors, one can see [2], [3], [12], [13], and [14]. The numbers

$$\underline{D}(P) := \liminf_{n\to\infty} \frac{2\log n}{-\log V_n(P)} \quad \text{and} \quad \overline{D}(P) := \limsup_{n\to\infty} \frac{2\log n}{-\log V_n(P)}$$

are called the *lower* and *upper quantization dimensions*, respectively, of the probability measure P. If $\underline{D}(P) = \overline{D}(P)$, the common value is called the *quantization dimension of* P and is denoted by $D(P)$. Quantization dimension measures the speed at which the specified measure of the error tends to zero as n approaches to infinity. Quantization dimension is also connected with other dimensions of dynamical systems; for details, one can see [7]. More accurate information about the asymptotics of the quantization error is provided by the quantization coefficients. For any $s \in (0, +\infty)$, the numbers $\liminf_n n^{\frac{2}{s}} V_n(P)$ and $\limsup_n n^{\frac{2}{s}} V_n(P)$ are called the s-dimensional *lower* and *upper quantization coefficients*, respectively, for P. If the s-dimensional lower and upper quantization coefficients for P are finite and positive, then s coincides with the quantization dimension of P.

Given a finite set $\alpha \subset \mathbb{R}^d$, 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 nearest to a. The family of sets $\{M(a|\alpha) : a \in \alpha\}$ is called the *Voronoi diagram* or *Voronoi tessellation* of $\mathbb{R}^d$ with respect to α. The generator $a \in \alpha$ is called the *centroid* of its own Voronoi region with respect to the probability distribution P, if

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

A Voronoi tessellation is called a *centroidal Voronoi tessellation* (CVT) if the generators are the centroids of their own Voronoi regions with respect to the measure P. The following proposition provides further information on the Voronoi regions generated by an optimal set of n-means.

Proposition 1.1 ([5], [7]). *Let α be an optimal set of n-means, $a \in \alpha$, and let $M(a|\alpha)$ 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. 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 [4], [5], [13]. Also, the optimal quantizers are the centroids of their own Voronoi regions [7].

Let C be the standard Cantor set generated by the similarity mappings $U_1(x) = \frac{1}{3}x$ and $U_2(x) = \frac{1}{3}x + \frac{2}{3}$, $x \in \mathbb{R}$, and let P_c be the associated self-similar Borel probability measure on $\mathbb{R}$. It is well known that $P_c = \frac{1}{2}P_c \circ U_1^{-1} + \frac{1}{2}P_c \circ U_2^{-1}$ with the Cantor set C as support [11]. For this probability measure, Graf and Luschgy in [6] determine the optimal sets of n-means and the n^{th} quantization errors for all $n \geq 2$. They also prove that the quantization dimension exists, but the quantization coefficient does not exist. Let S be a Cantor dust on $\mathbb{R}^2$ which is generated by a set of 4^u similarity mappings $\{S_j : j = 1, 2, 3, \cdots, 4^u\}$, where $u \in \mathbb{N}$. Let P be a Borel probability measure on $\mathbb{R}^2$ such that

$$P = \frac{1}{4^u} \sum_{j=1}^{4^u} P \circ S_j^{-1}.$$

Then P is unique and P has the Cantor dust S as support. It is extremely difficult to calculate the optimal sets of n-means and the n^{th} quantization errors for all $n \in \mathbb{N}$ and all $u \in \mathbb{N}$. Even they were not known for $u = 1$. In this paper, we investigate them for $u = 1$. The Cantor dust for $u = 1$ is called the *standard Cantor dust*, and it is generated by the maps $S_1(x_1, x_2) = \frac{1}{3}(x_1, x_2)$, $S_2(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (\frac{2}{3}, 0)$, $S_3(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (0, \frac{2}{3})$, and $S_4(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (\frac{2}{3}, \frac{2}{3})$, $(x_1, x_2) \in \mathbb{R}^2$. The associated self-similar Borel probability measure P on $\mathbb{R}^2$ is given by $P = \frac{1}{4}P \circ S_1^{-1} + \frac{1}{4}P \circ S_2^{-1} + \frac{1}{4}P \circ S_3^{-1} + \frac{1}{4}P \circ S_4^{-1}$. Then P has support S [11]. For this probability measure P, we have determined the optimal sets of n-means and the n^{th} quantization errors for all $n \geq 1$. The standard (as well as more general) Cantor dust S satisfies the strong separation condition (SSC); i.e., $S_1(S)$, $S_2(S)$, $S_3(S)$, and $S_4(S)$ are pairwise disjoint;

hence, the quantization dimension of the probability measure P exists and equals the Hausdorff dimension of S [6]. In this paper, we show that the quantization coefficient for P with support the standard Cantor dust does not exist.

For any n points, in the case of the (standard) Cantor distribution P_c, one can easily determine whether the n points form a CVT [6]; however, in the case of the probability distribution P supported by the Cantor dust, considered in this paper, it is quite difficult to determine whether the n points form a CVT. From the construction, the Cantor dust S above can be viewed as $C \times C$ and the probability distribution P can be considered as $P_c \times P_c$. Hence, it might be surmised that for any $n \geq 1$, the optimal sets of n-means can be deduced as products of those of C. This is far from the truth; as it will be clear from the arguments in the next sections that, for instance, while $\alpha = \{\frac{1}{18}, \frac{5}{18}, \frac{5}{6}\}$ is an optimal set of 3-means for P_c, the set $\alpha \times \alpha$ is not an optimal set of 9-means for P. By the same token, there is no direct connection between n^{th} quantization errors in the cases C and S.

The technique we utilize can be extended to determine the optimal sets and the corresponding quantization error for many other singular continuous probability measures on $\mathbb{R}^2$, such as probability measures on more general Cantor dusts generated by 4^u self-similar maps, $u \geq 2$.

2. Preliminaries

Let $I := \{1, 2, 3, 4\}$ be an alphabet and, for $k \geq 1$, call any finite sequence $\sigma := \sigma_1\sigma_2 \cdots \sigma_k$, where $\sigma_i \in I$, a *word* σ over I with length k. A word of length zero is called the *empty word* and is denoted by $\emptyset$. By I^*, we denote the set of all words over I of some finite length k, including the empty word $\emptyset$. For any two words $\sigma := \sigma_1\sigma_2 \cdots \sigma_k$ and $\tau := \tau_1\tau_2 \cdots \tau_\ell$ in I^*, by $\sigma\tau := \sigma_1 \cdots \sigma_k\tau_1 \cdots \tau_\ell$, we mean the word obtained from the concatenation of the two words σ and τ. For $\sigma, \tau \in I^*$, σ is called an *extension of* τ if $\sigma = \tau x$ for some word $x \in I^*$. The maps $S_i : \mathbb{R}^2 \to \mathbb{R}^2$, $1 \leq i \leq 4$, defined above, are the generating maps of the standard Cantor dust S. For $\sigma := \sigma_1\sigma_2 \cdots \sigma_k \in I^k$, set $S_\sigma := S_{\sigma_1} \circ \cdots \circ S_{\sigma_k}$ and $J_\sigma := S_\sigma([0,1] \times [0,1])$. For the empty word $\emptyset$, by $S_\emptyset$, it is meant the identity mapping on $\mathbb{R}^2$, and we will write $J := J_\emptyset = S_\emptyset([0,1] \times [0,1]) = [0,1] \times [0,1]$. Hence, $S := \cap_{k \in \mathbb{N}} \cup_{\sigma \in I^k} J_\sigma$, and the elements of the set $\{J_\sigma : \sigma \in I^k\}$ are the 4^k squares in the k^{th} level in the construction of S, which will be called *basic squares at the* k^{th} *level.* The squares $J_{\sigma 1}$, $J_{\sigma 2}$, $J_{\sigma 3}$, and $J_{\sigma 4}$ into which J_σ is split up at the $(k+1)^{\text{th}}$ level are called the *children of* J_σ. The point of intersection of the two diagonals of J_σ will be called the *center* of J_σ.

A few observations on the S and self-similar probability measure P are in order. Clearly, S has four axes of symmetry: the lines h $(x_2 = \frac{1}{2})$, v $(x_1 = \frac{1}{2})$, r $(x_1 = x_2)$, and l $(x_1 + x_2 = 1)$. Furthermore, by the following statement, the marginal distributions of P are P_c. Since the proof is straightforward, it will not be given.

Proposition 2.1. *Let P_1 and P_2 be the marginal distributions of P; i.e., $P_1(A) = P(A \times \mathbb{R})$ for all $A \in \mathfrak{B}$, and $P_2(B) = P(\mathbb{R} \times B)$ for all $B \in \mathfrak{B}$, where $\mathfrak{B}$ is the Borel σ-algebra on $\mathbb{R}$. Then $P_1 = P_2 = P_c$, where P_c is the Cantor distribution associated with standard Cantor set.*

Since $P = \frac{1}{4}P \circ S_1^{-1} + \frac{1}{4}P \circ S_2^{-1} + \frac{1}{4}P \circ S_3^{-1} + \frac{1}{4}P \circ S_4^{-1}$, we have, by induction, that $P = \sum_{\sigma \in I^k} \frac{1}{4^k} P \circ S_\sigma^{-1}$. Hence, Lemma 2.2 follows.

Lemma 2.2. *Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+ \times \mathbb{R}^+$ be Borel measurable and $k \in \mathbb{N}$. Then*

$$\int f \, dP = \frac{1}{4^k} \sum_{\sigma \in I^k} \int f \circ S_\sigma \, dP.$$

Remark 2.3. With respect to the horizontal and vertical lines h and v, and the diagonals r and l, the Cantor dust S has *maximum symmetry*, i.e., with respect to the four lines, the Cantor dust is geometrically symmetric as well as symmetric with respect to the probability distribution P (see Figure 1). By the symmetry with respect to P, it is meant that if the two basic squares of similar geometrical shape lie in the opposite sides and are equidistant with respect to any of the lines h, v, r, and l, then they have the same probability. Notice that the Cantor dust has four lines of maximum symmetry. Due to this fact, we conjecture that among all the pairs of points whose Voronoi regions have the boundaries as oblique lines passing through the point $(\frac{1}{2}, \frac{1}{2})$, the pair of points whose Voronoi regions have the boundary as any of the two diagonals will give the smallest distortion error.

Notation 2.4. As usual, let $X = (X_1, X_2)$ be a random vector with distribution P where $E(X)$ and $V(X)$ will denote the expected vector and the expected squared Euclidean distance of X, respectively. For words $\beta, \gamma, \cdots, \delta \in I^*$, by $a(\beta, \gamma, \cdots, \delta)$, we mean the conditional expectation of the random variable X given $J_\beta \cup J_\gamma \cup \cdots \cup J_\delta$; i.e.,

$$\begin{aligned} a(\beta, \gamma, \cdots, \delta) &= E(X | X \in J_\beta \cup J_\gamma \cup \cdots \cup J_\delta) \\ &= \frac{1}{P(J_\beta \cup \cdots \cup J_\delta)} \int_{J_\beta \cup \cdots \cup J_\delta} x dP, \end{aligned} \tag{1}$$

where $\int x dP = \int (x_1, x_2) dP$.

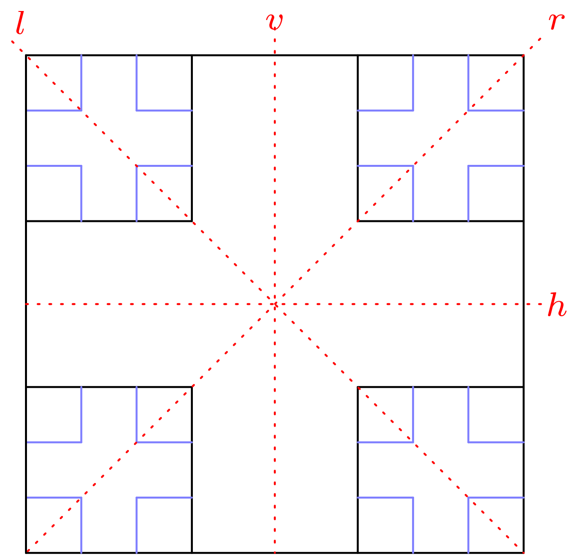


FIGURE 1. Symmetry axes of the Cantor dust.

Lemma 2.5. *Let $X := (X_1, X_2)$ be a random vector with distribution P. Then $E(X) = (E(X_1), E(X_2)) = (\frac{1}{2}, \frac{1}{2})$ and $V := V(X) = E\|X - (\frac{1}{2}, \frac{1}{2})\|^2 = \frac{1}{4}$.*

Proof. Since P_1 and P_2 are the marginal distributions of $X := (X_1, X_2)$, the random variables X_1 and X_2 have distributions P_1 and P_2, respectively. Since, by Proposition 2.1, $P_1 = P_2 = P_c$, both X_1 and X_2 are P_c-distributed random variables. Thus, by [6, Lemma 3.4], we obtain that $E(X_1) = E(X_2) = \frac{1}{2}$ and $V(X_1) = V(X_2) = \frac{1}{8}$. Consequently, it follows that $E\|X-(\frac{1}{2}, \frac{1}{2})\|^2 = E(X_1-\frac{1}{2})^2+E(X_2-\frac{1}{2})^2 = V(X_1)+V(X_2) = \frac{1}{4}$. $\square$

Remark 2.6. For any $(a, b) \in \mathbb{R}^2$,

$$\begin{aligned}
&E\|X - (a, b)\|^2 = \iint_{\mathbb{R}^2} [(x_1 - a)^2 + (x_2 - b)^2]\, dP(x_1, x_2)\\
&= \int_0^1 (x_1 - a)^2\, dP_1(x_1) + \int_0^1 (x_2 - b)^2\, dP_2(x_2)\\
&= E(X_1 - a)^2 + E(X_2 - b)^2 = V(X_1) + V(X_2) + (a - \frac{1}{2})^2 + (b - \frac{1}{2})^2\\
&= V + \|(a, b) - (\frac{1}{2}, \frac{1}{2})\|^2.
\end{aligned}$$

Hence, by Lemma 2.2, for any $\sigma \in I^k$, $k \geq 1$, we have

$$\int_{J_\sigma} \|x-(a,b)\|^2 dP = \frac{1}{4^k} \int \|x-(a,b)\|^2 dP \circ S_\sigma^{-1},$$

which implies

$$\int_{J_\sigma} \|x-(a,b)\|^2 dP = \frac{1}{4^k}\Big(\frac{1}{9^k} V + \|S_\sigma(\frac{1}{2}, \frac{1}{2}) - (a,b)\|^2\Big). \tag{2}$$

3. Optimal Sets of n-Means and the Quantization Errors for All $n \geq 2$

In this section, we determine the optimal sets of n-means, $n \geq 1$. First, we will determine optimal sets of n-means for $n = 1, 2, 3$, which will form the base in determining optimal sets of n-means for $n \geq 4$. To determine the quantization error, we will frequently use the relation (2). Below, an *oblique line* will mean any line which is neither horizontal nor vertical.

First, from Lemma 2.5, it follows that the optimal set of one-mean is $E(X) = \{(\frac{1}{2}, \frac{1}{2})\}$ and the corresponding quantization error is $V(X) = \frac{1}{4}$ of the random variable X. For any word $\sigma \in I^k$, $k \geq 1$, since $a(\sigma) = E(X|X \in J_\sigma)$, using Lemma 2.2, 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_i are similarity mappings, it follows that $E(S_j(X)) = S_j(E(X))$ for $j = 1, 2, 3, 4$, and so, by induction,

$$a(\sigma) = E(S_\sigma(X)) = S_\sigma(E(X)) = S_\sigma(\frac{1}{2}, \frac{1}{2}), \text{ for } \sigma \in I^k,\ k \geq 1.$$

From this observation, we surmise that the sets

$$\{(\frac{1}{6}, \frac{1}{2}), (\frac{5}{6}, \frac{1}{2})\} \text{ and } \{(\frac{1}{2}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$$

form two different optimal sets of two-means. In order to prove this fact, first we will show that an optimal set of two-means cannot lie on an oblique line.

Lemma 3.1. *The points in an optimal set of two-means cannot lie on a oblique line of S.*

Proof. The boundary of the Voronoi regions generated by the points in $\beta := \{(\frac{1}{6}, \frac{1}{2}), (\frac{5}{6}, \frac{1}{2})\}$ is the line v; i.e., $J_1 \cup J_3 \subset M((\frac{1}{6}, \frac{1}{2})|\beta)$ and $J_2 \cup J_4 \subset M((\frac{5}{6}, \frac{1}{2})|\beta)$. Let $V_{2,1}$ be the distortion error due to the set β. Then

$$V_{2,1} = \int\limits_{J_1 \cup J_3} \|x - (\frac{1}{6}, \frac{1}{2})\|^2 dP + \int\limits_{J_3 \cup J_4} \|x - (\frac{5}{6}, \frac{1}{2})\|^2 dP = \frac{5}{36}.$$

Among all the oblique lines through the point $(\frac{1}{2}, \frac{1}{2})$, as mentioned in Remark 2.3, the pair of points whose Voronoi regions have the boundary as any of the two diagonals will give the smallest distortion error.

We also know that the points which give the smallest distortion error are the centroids of their own Voronoi regions. Let (a_1, b_1) and (a_2, b_2) be the centroids of the left half and the right half, respectively, of the Cantor dust with respect to the diagonal passing through the origin. Thus, using (1), we have

$$\begin{aligned}
&(a_1, a_2) = E\left(X | X \in J_2 \cup (J_{12} \cup J_{42}) \cup (J_{112} \cup J_{142} \cup J_{412} \cup J_{442}) \cup \cdots\right) \\
&= 2\Big(\frac{1}{4} S_2(\frac{1}{2}, \frac{1}{2}) + \frac{1}{4^2}\Big(S_{12}(\frac{1}{2}, \frac{1}{2}) + S_{42}(\frac{1}{2}, \frac{1}{2})\Big) + \frac{1}{4^3}\Big(S_{112}(\frac{1}{2}, \frac{1}{2}) \\
&\qquad + S_{142}(\frac{1}{2}, \frac{1}{2}) + S_{412}(\frac{1}{2}, \frac{1}{2}) + S_{442}(\frac{1}{2}, \frac{1}{2})\Big) + \cdots\Big) \\
&= 2\Big((\frac{5}{24}, \frac{1}{24}) + (\frac{11}{144}, \frac{7}{144}) + (\frac{29}{864}, \frac{25}{864}) + (\frac{83}{5184}, \frac{79}{5184}) + \cdots\Big) \\
&= 2\Big(\frac{5}{24} + \frac{11}{144} + \frac{29}{864} + \frac{83}{5184} + \cdots, \frac{1}{24} + \frac{7}{144} + \frac{25}{864} + \frac{79}{5184} + \cdots\Big) \\
&= 2\Big(\frac{7}{20}, \frac{3}{20}\Big),
\end{aligned}$$

i.e,

$$(a_1, a_2) = \Big(\frac{7}{10}, \frac{3}{10}\Big). \tag{3}$$

Similarly, one can show that

$$(a_2, b_2) = E\left(X | X \in J_3 \cup (J_{13} \cup J_{43}) \cup (J_{113} \cup J_{143} \cup J_{413} \cup J_{443}) \cup \cdots\right),$$

implying

$$(a_2, b_2) = \Big(\frac{3}{10}, \frac{7}{10}\Big). \tag{4}$$

Hence, if $V_{2,2}$ is the distortion error due to the points $(\frac{7}{10}, \frac{3}{10})$ and $(\frac{3}{10}, \frac{7}{10})$, then, writing $A = J_{12} \cup J_{42}$, $B = J_{112} \cup J_{412} \cup J_{142} \cup J_{442}$, and $C = J_{1112} \cup J_{1412} \cup J_{1142} \cup J_{1442} \cup J_{4112} \cup J_{4412} \cup J_{4142} \cup J_{4442}$, we have

$$\begin{aligned}
&V_{2,2}\\
&= 2\Big(\text{distortion error due to } (\frac{7}{10}, \frac{3}{10})\Big) > 2\Big(\int_{J_2} \|(x_1, x_2) - (\frac{7}{10}, \frac{3}{10})\|^2 dP\\
&\quad + \int_A \|(x_1, x_2) - (\frac{7}{10}, \frac{3}{10})\|^2 dP + \int_B \|(x_1, x_2) - (\frac{7}{10}, \frac{3}{10})\|^2 dP\\
&\quad + \int_C \|(x_1, x_2) - (\frac{7}{10}, \frac{3}{10})\|^2 dP\Big)\\
&= 2(0.0747492) = 0.1494984 > \frac{5}{36} = V_{2,1}.
\end{aligned}$$

Therefore, the points in an optimal set of two-means cannot lie on an oblique line of the Cantor dust S. □

The following proposition gives all optimal sets of two-means.

Proposition 3.2. *The sets* $\{(\frac{1}{6}, \frac{1}{2}), (\frac{5}{6}, \frac{1}{2})\}$ *and* $\{(\frac{1}{2}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ *form two different optimal sets of two-means with quantization error* $V_2 = \frac{5}{36}$.

Proof. By Lemma 3.1, the points in an optimal set of two-means cannot lie on an oblique line of S; hence, the two optimal quantizers lie either on a horizontal line or on a vertical line. We will show that these lines are v or h.

First, assume that the optimal sets lie on a horizontal line. Let $\alpha := \{(a,p),(b,p)\}$ be an optimal set of two-means. Since the optimal quantizers are the centroids of their own Voronoi regions, it follows by the properties of centroids that

$$(a,p)P(M((a,p)|\alpha)) + (b,p)P(M((b,p)|\alpha)) = (\frac{1}{2}, \frac{1}{2}),$$

which implies that

$$\begin{aligned}
aP(M((a,p)|\alpha)) + bP(M((b,p)|\alpha)) &= pP(M((a,p)|\alpha)) + pP(M((b,p)|\alpha))\\
&= \frac{1}{2}.
\end{aligned}$$

Thus, $p = \frac{1}{2}$, and the two optimal quantizers $(a, \frac{1}{2})$ and $(b, \frac{1}{2})$ lie on the line h and are in opposite sides of the point $(\frac{1}{2}, \frac{1}{2})$. Again, since the optimal quantizers are the centroids of their own Voronoi regions, it follows that $0 \leq a < \frac{1}{2} < b \leq 1$. Thus,

$$\begin{aligned} V_2 &= \int \min_{c \in \alpha} \|x - c\|^2 dP = \int \min_{c \in \{a,b\}} \|x - (c, \frac{1}{2})\|^2 dP \\ &= \int\limits_{[0, \frac{a+b}{2}] \times [0,1]} \|x - (a, \frac{1}{2})\|^2 dP \ + \int\limits_{[\frac{a+b}{2}, 1] \times [0,1]} \|x - (b, \frac{1}{2})\|^2 dP \\ &= \int\limits_{[0, \frac{a+b}{2}]} (x_1 - a)^2 dP_c \ + \int\limits_{[\frac{a+b}{2}, 1]} (x_1 - b)^2 dP_c \\ &+ \Big(P_c([0, \frac{a+b}{2}]) \ + P_c([\frac{a+b}{2}, 1])\Big) \int (x_2 - \frac{1}{2})^2 dP_c \\ &= \int \min_{c \in \{a,b\}} (x_1 - c)^2 dP_c \ + \int (x_2 - \frac{1}{2})^2 dP_c. \end{aligned}$$

Now, $V_2(P_c) = \int \min_{c \in \{a,b\}} (x - c)^2 dP_c = \frac{1}{72}$ (see [6, Proposition 4.6]), and it occurs when $a = \frac{1}{6}$ and $b = \frac{5}{6}$. Since $\int (x_2 - \frac{1}{2})^2 dP_c = \frac{1}{8}$, we deduce that

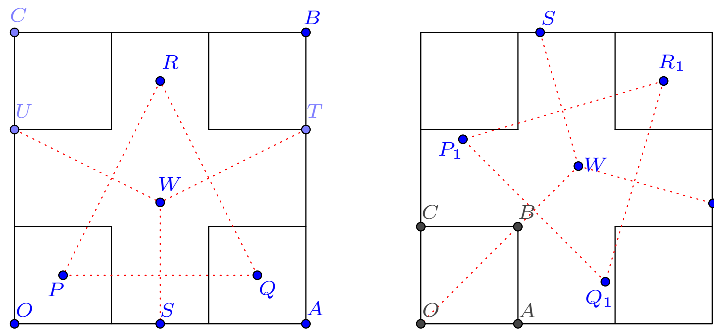
$$V_2 = \int \min_{c \in \{a,b\}} (x_1 - c)^2 dP_c + \int (x_2 - \frac{1}{2})^2 dP_c = \frac{1}{72} + \frac{1}{8} = \frac{5}{36},$$

and $\{(\frac{1}{6}, \frac{1}{2}), (\frac{5}{6}, \frac{1}{2})\}$ is an optimal set of two-means. Due to symmetry, the set $\{(\frac{1}{2}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ also forms an optimal set of two-means. □

We now state and prove Lemma 3.3 and Lemma 3.5. These two lemmas are useful to determine the optimal sets of three means.

Lemma 3.3. *The set* $\alpha_3 = \{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ *forms a CVT with three-means of distortion error* $\frac{1}{12}$.

Proof. Recall that the boundaries of the Voronoi regions lie along the perpendicular bisectors of the line segments joining their centers. The perpendicular bisectors of the line segments joining each pair of points from the set α_3 are SW, TW, and UW with equations $x_1 = \frac{1}{2}$, $x_2 = \frac{1}{2}x_1 + \frac{1}{6}$, and $x_2 = -\frac{1}{2}x_1 + \frac{2}{3}$, respectively, and they concur at the point $W(\frac{1}{2}, \frac{5}{12})$ as shown in Figure 2(a).



(a) (b)

FIGURE 2. (a) CVT with three-means P, Q, and R with one on the vertical line of the symmetry; (b) CVT with three-means P_1, Q_1, and R_1 with one on a diagonal of the square.

If (p_1, p_2), (q_1, q_2), and (r_1, r_2) are the centroids of these regions associated with P, respectively, we have

$$
\begin{aligned}
(p_1, p_2) &= \frac{1}{P(M_1)} \int_{M_1} x dP = \frac{1}{P(J_1)} \int_{J_1} x dP = \int x d(P \circ S_1^{-1}) \\
&= S_1(\frac{1}{2}, \frac{1}{2}) = (\frac{1}{6}, \frac{1}{6}), \\
(q_1, q_2) &= \frac{1}{P(M_2)} \int_{M_2} x dP = \frac{1}{P(J_2)} \int_{J_2} x dP = \int x d(P \circ S_2^{-1}) \\
&= S_2(\frac{1}{2}, \frac{1}{2}) = (\frac{5}{6}, \frac{1}{6}), \text{ and} \\
(r_1, r_2) &= \frac{1}{P(M_3)} \int_{M_3} x dP = \frac{1}{P(J_3 \cup J_4)} \int_{J_3 \cup J_4} x dP \\
&= \frac{1}{P(J_3 \cup J_4)} \Big(\int_{J_3} x dP + \int_{J_4} x dP \Big) \\
&= \frac{1}{P(J_3 \cup J_4)} \Big(P(J_3) \int x d(P \circ S_3^{-1}) + P(J_4) \int x d(P \circ S_4^{-1}) \Big) \\
&= 2 \Big(\frac{1}{4} S_3(\frac{1}{2}, \frac{1}{2}) + \frac{1}{4} S_4(\frac{1}{2}, \frac{1}{2}) \Big) = (\frac{1}{2}, \frac{5}{6}).
\end{aligned}
$$

Thus, we see that the given set α_3 forms a CVT with three-means. By (2), the corresponding distortion error is

$$\begin{aligned}
&\int \min_{a\in\alpha_3} \|x-a\|^2 dP\\
&=\int_{J_1} \|x-(\frac{1}{6},\frac{1}{6})\|^2 dP+\int_{J_2} \|x-(\frac{5}{6},\frac{1}{6})\|^2 dP+\int_{J_3\cup J_4} \|x-(\frac{1}{2},\frac{5}{6})\|^2 dP\\
&=\frac{1}{36}V+\frac{1}{36}V+\frac{1}{36}\Big[2V+\|S_3(\frac{1}{2},\frac{1}{2})-(\frac{1}{2},\frac{5}{6})\|^2+\|S_4(\frac{1}{2},\frac{1}{2})-(\frac{1}{2},\frac{5}{6})\|^2\Big]\\
&=\frac{1}{12},
\end{aligned}$$

proving the assertion. □

Remark 3.4. The points in the set $\alpha_3=\{(\frac{1}{6},\frac{1}{6}),(\frac{5}{6},\frac{1}{6}),(\frac{1}{2},\frac{5}{6})\}$ given by Lemma 3.3 form vertices of an isosceles triangle. Due to rotational symmetry and uniformity of P, there are four such sets giving the same distortion error $\frac{1}{12}$.

Figure 2(b) suggests that the set of points $\{(\frac{5}{6},\frac{5}{6}),(\frac{13}{90},\frac{19}{30}),(\frac{19}{30},\frac{13}{90})\}$ also forms a CVT with three-means. The following statement confirms this prediction; however, the associated distortion error is larger than $\frac{1}{12}$.

Lemma 3.5. *The set $\beta_3=\{(\frac{5}{6},\frac{5}{6}),(\frac{13}{90},\frac{19}{30}),(\frac{19}{30},\frac{13}{90})\}$ forms a CVT with three-means and the corresponding distortion error is larger than $\frac{1}{12}$.*

Proof. The perpendicular bisectors of the line segments joining each pair of points from the set of points $\{(\frac{5}{6},\frac{5}{6}),(\frac{13}{90},\frac{19}{30}),(\frac{19}{30},\frac{13}{90})\}$ are SW, OW, and TW with equations $x_2=\frac{979}{405}-\frac{31x_1}{9}$, $x_2=x_1$, and $x_2=\frac{979}{1395}-\frac{9x_1}{31}$, respectively, which meet at the point $W(\frac{979}{1800},\frac{979}{1800})$. Let (p_1,p_2), (q_1,q_2), and (r_1,r_2) be the centroids of the three Voronoi regions with centers $P_1(\frac{13}{90},\frac{19}{30})$, $Q_1(\frac{19}{30},\frac{13}{90})$, and $R_1(\frac{5}{6},\frac{5}{6})$, respectively. Since the similarity mappings preserve the ratio of the distances of a point from any other two points, by (3) and (4), with respect to the probability measure P, the centroids of the triangles OBC and OAB are obtained as $S_1(\frac{3}{10},\frac{7}{10})=(\frac{3}{30},\frac{7}{30})$ and $S_1(\frac{7}{10},\frac{3}{10})=(\frac{7}{30},\frac{3}{30})$, respectively. Therefore, using the definition of centroids, we have

$$\begin{aligned}(p_1,p_2)&=\frac{1}{P(J_3)+P(\triangle OBC)}\Big(P(J_3)\int_{J_3}xdP+P(\triangle OBC)\int_{\triangle OBC}xdP\Big)\\&=\frac{1}{\frac14+\frac18}\Big(\frac14(\frac16,\frac56)+\frac18(\frac{3}{30},\frac{7}{30})\Big)=\Big(\frac{13}{90},\frac{19}{30}\Big),\\(q_1,q_2)&=\frac{1}{P(J_2)+P(\triangle OAB)}\Big(P(J_2)\int_{J_2}xdP+P(\triangle OAB)\int_{\triangle OAB}xdP\Big)\\&=\frac{1}{\frac14+\frac18}\Big(\frac14(\frac16,\frac56)+\frac18(\frac{7}{30},\frac{3}{30})\Big)=\Big(\frac{19}{30},\frac{13}{90}\Big),\text{ and}\\(r_1,r_2)&=S_4(\frac12,\frac12)=\Big(\frac56,\frac56\Big).\end{aligned}$$

Thus, the set β_3 forms a CVT with three-means. Now, using (2), the corresponding distortion error is

$$\begin{aligned}&\int\min_{a\in\beta_3}\|x-a\|^2dP\\&=\Big(\text{distortion error due to}(\frac56,\frac56)\Big)+2\Big(\text{distortion error due to}(\frac{13}{90},\frac{19}{30})\Big)\\&>\frac{1}{36}V+2\Big(\int_{J_3}\|x-(\frac{13}{90},\frac{19}{30})\|^2dP+\int_{J_{13}}\|x-(\frac{13}{90},\frac{19}{30})\|^2dP\\&+\int_{J_{113}\cup J_{143}}\|x-(\frac{13}{90},\frac{19}{30})\|^2dP\\&+\int_{J_{1113}\cup J_{1143}\cup J_{1413}\cup J_{1443}}\|x-(\frac{13}{90},\frac{19}{30})\|^2dP\\&+\int_{J_{11113}\cup J_{11143}\cup J_{11413}\cup J_{11443}\cup J_{14113}\cup J_{14143}\cup J_{14413}\cup J_{14443}}\|x-(\frac{13}{90},\frac{19}{30})\|^2dP\\&+\int_{J_{111113}\cup J_{111143}\cup J_{111413}\cup J_{111443}\cup J_{114113}\cup J_{114143}}\|x-(\frac{13}{90},\frac{19}{30})\|^2dP\Big)\\&=\frac{1247143}{14929920}>\frac{1}{12},\end{aligned}$$

completing the proof of the lemma. □

Remark 3.6. One of the points of the CVT β_3 in Lemma 3.5 is the centroid of the child J_4 and the other two points are equidistant from the diagonal passing through the centroid. Due to rotational symmetry of S, there are four such CVTs with three-means in which one point is the centroid of one of the children J_1, J_2, J_3, or J_4 and the other two points are equidistant from the diagonal passing through the centroid; all have the same distortion error larger than $\frac{1}{12}$.

Proposition 3.7. *Let α_3 be the set given by Lemma 3.3. Then α_3 forms an optimal set of three-means with quantization error $\frac{1}{12}$. The number of optimal sets of three-means is four.*

Proof. As mentioned in Remark 2.3, the Cantor dust has four lines of maximum symmetry: h, v, r, and l (see Figure 1). Due to this, we conjecture that one of the points in an optimal set of three-means must lie on any of the above four lines and other two will be equidistant from the line. Thus, comparing the distortion errors for the CVTs given by Lemma 3.3 and Lemma 3.5, we deduce that the CVT α_3, given by Lemma 3.3, forms an optimal set of three-means with quantization error $\frac{1}{12}$ (see Figure 3). Remark 3.4 implies that the number of such sets is four. □

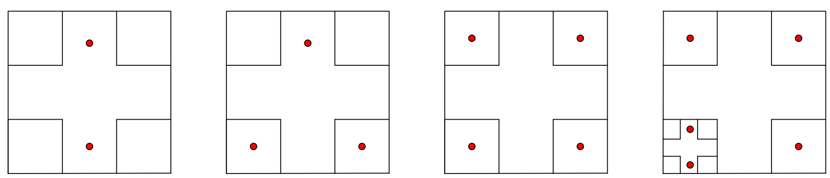


FIGURE 3. Optimal configuration of n points for $2 \leq n \leq 5$.

Remark 3.8. Lemma 3.3 and Lemma 3.5 together show that under squared error distortion measure, the centroid condition is not sufficient for optimal quantization for singular continuous probability measures on $\mathbb{R}^2$, which is already known for absolutely continuous probability measures on $\mathbb{R}^2$ [4] and for singular continuous probability measure on $\mathbb{R}$ [13].

Lemma 3.9. *Let $n \geq 4$ and let α_n be an optimal set of n-means such that $\alpha_n \cap J_i \neq \emptyset$ for $1 \leq i \leq 4$ and α_n does not contain any point from $J \setminus \cup_{i=1}^4 J_i$. If $\beta_i := \alpha_n \cap J_i$ with $n_i :=$ card(β_i), then $S_i^{-1}(\beta_i)$ is an optimal set of n_i-means. Moreover,*

$$V_n = \frac{1}{36}\left(V_{n_1} + V_{n_2} + V_{n_3} + V_{n_4}\right).$$

Proof. By the hypothesis, $\beta_i \neq \emptyset$ for all $1 \leq i \leq 4$, and α_n does not contain any point from $J \setminus \cup_{i=1}^4 J_i$; hence, $\alpha_n = \cup_{i=1}^4 \beta_i$. Since α_n is an optimal set of n-means,

$$V_n = \sum_{i=1}^4 \int_{J_i} \min_{a \in \alpha_n} \|x-a\|^2 dP = \sum_{i=1}^4 \int_{J_i} \min_{a \in \beta_i} \|x-a\|^2 dP.$$

Now, using Lemma 2.2, we have
(5)
$$V_n = \frac{1}{36}\sum_{i=1}^{4}\int \min_{a\in\beta_i}\|x-S_i^{-1}(a)\|^2dP = \frac{1}{36}\sum_{i=1}^{4}\int \min_{a\in S_i^{-1}(\beta_i)}\|x-a\|^2dP.$$
If $S_1^{-1}(\beta_1)$ is not an optimal set of n_1-means, then we can find a set $\gamma_1 \subset \mathbb{R}^2$ with $\text{card}(\gamma_1) = n_1$ such that
$$\int \min_{a\in\gamma_1}\|x-a\|^2dP < \int \min_{a\in S_1^{-1}(\beta_1)}\|x-a\|^2dP.$$
But then $S_1(\gamma_1)\cup\beta_2\cup\beta_3\cup\beta_4$ will be a set of cardinality n, and
$$\begin{aligned}
&\int \min\{\|x-a\|^2 : a\in S_1(\gamma_1)\cup\beta_2\cup\beta_3\cup\beta_4\}dP\\
&=\int_{J_1} \min_{a\in S_1(\gamma_1)}\|x-a\|^2dP + \frac{1}{36}\sum_{i=2}^{4}\int \min_{a\in S_i^{-1}(\beta_i)}\|x-a\|^2dP\\
&=\frac{1}{36}\int \min_{a\in S_1(\gamma_1)}\|x-S_1^{-1}(a)\|^2dP + \frac{1}{36}\sum_{i=2}^{4}\int \min_{a\in S_i^{-1}(\beta_i)}\|x-a\|^2dP\\
&=\frac{1}{36}\int \min_{a\in\gamma_1}\|x-a\|^2dP + \frac{1}{36}\sum_{i=2}^{4}\int \min_{a\in S_i^{-1}(\beta_i)}\|x-a\|^2dP\\
&<\frac{1}{36}\int \min_{a\in S_1^{-1}(\beta_1)}\|x-a\|^2dP + \frac{1}{36}\sum_{i=2}^{4}\int \min_{a\in S_i^{-1}(\beta_i)}\|x-a\|^2dP.
\end{aligned}$$
Thus, by (5), we have $\int \min\{\|x-a\|^2 : a\in S_1(\gamma_1)\cup\beta_2\cup\beta_3\cup\beta_4\}dP < V_n$, which contradicts the fact that α_n is an optimal set of n-means, and so $S_1^{-1}(\beta_1)$ is an optimal set of n_1-means. Similarly, one can show that $S_i^{-1}(\beta_i)$ are optimal sets of n_i-means for all $2\le i\le 4$. Thus, (5) implies $V_n = \frac{1}{36}\left(V_{n_1}+V_{n_2}+V_{n_3}+V_{n_4}\right)$. This completes the proof of the lemma. □

The similarity mappings S_i preserve the ratio of the distances of a point from any other two points; hence, the following proposition is an immediate corollary of Lemma 3.9.

Proposition 3.10. *Let $n\ge 4$ and α_n be an optimal set of n-means such that $\alpha_n\cap J_\sigma\neq\emptyset$ for $\sigma\in I^{\ell(n)}$ for some $\ell(n)\in\mathbb{N}$, and α_n does not contain any point from $J\setminus\cup_{\sigma\in I^{\ell(n)}}J_\sigma$. If $\beta_\sigma := \alpha_n\cap J_\sigma$ with $n_\sigma := \text{card}(\beta_\sigma)$, then $S_\sigma^{-1}(\beta_\sigma)$ is an optimal set of n_σ-means and*
$$V_n = \frac{1}{36^{\ell(n)}}\sum_{\sigma\in I^{\ell(n)}} V_{n_\sigma}.$$

Proposition 3.10 provides a major step in describing the optimal sets of n-means and the associated quantization errors, subject to the conditions that $\alpha_n \cap J_\sigma \neq \emptyset$ and $\alpha_n \cap (J \setminus \cup_{\sigma \in I^{\ell(n)}} J_\sigma) = \emptyset$. Thus, it remains to prove that these conditions are satisfied, which will be the focus of the next three statements.

Lemma 3.11. *If α_n is an optimal set of n-means for $n \geq 4$, then $\alpha_n \cap J_i \neq \emptyset$ for $1 \leq i \leq 4$.*

Proof. Recall that the Cantor dust has four lines of symmetry: h, v, r, and l. First, we will prove that, for $n \geq 4$, an optimal set of n-means α_n meets each of the quadrants determined by the lines h and v. Consider the set $\beta := \{S_i(\frac{1}{2}, \frac{1}{2}) : 1 \leq i \leq 4\}$. Then the distortion error due to the set β is given by

$$\int \min_{a \in \beta} \|x - a\|^2 dP = \sum_{i=1}^{4} \int_{J_i} \|x - S_i(\frac{1}{2}, \frac{1}{2})\|^2 dP = \frac{1}{9} V = \frac{1}{36}.$$

Since V_n is the quantization error for n-means for $n \geq 4$, we have $V_n \leq V_4 \leq \frac{1}{36}$. If all the elements of α_n lie on the line h, then, for any $x \in \cup_{i=1}^4 J_{ii}$, $\min_{a \in \alpha_n} \|x - a\|^2 \geq (\frac{1}{2} - \frac{1}{9})^2 = \frac{49}{324}$, and the distortion error is

$$\begin{aligned} &\int \min_{a \in \alpha_n} \|x - a\|^2 dP \\ &> \sum_{i=1}^{4} \int_{J_{ii}} \min_{a \in \alpha_n} \|x - a\|^2 dP \geq \sum_{i=1}^{4} \frac{49}{324} P(J_{ii}) = \frac{49}{1296} > \frac{1}{36} \geq V_n, \end{aligned}$$

which contradicts the optimality of α. Therefore, $\alpha_n \not\subseteq h$. It follows by the same argument that $\alpha_n \not\subseteq v$. If $\alpha_n \subset r$, then, for any $x \in J_{22} \cup J_{33}$, $\min_{a \in \alpha_n} \|x - a\|^2 \geq \|(\frac{1}{9}, \frac{8}{9}) - (\frac{1}{2}, \frac{1}{2})\|^2 = \frac{49}{162}$. Hence,

$$\begin{aligned} \int \min_{a \in \alpha_n} \|x - a\|^2 dP &> \int_{J_{22} \cup J_{33}} \min_{a \in \alpha_n} \|x - a\|^2 dP \geq 2 \cdot \frac{49}{162} \cdot \frac{1}{16} \\ &= \frac{49}{1296} > V_n, \end{aligned}$$

which, again, contradicts optimality. Similarly, $\alpha_n \not\subseteq l$, either. Now, by the properties of centroids, we have

$$\sum_{i=1}^{4} (a_i, b_i) P(M((a_i, b_i)|\alpha_n)) = (\frac{1}{2}, \frac{1}{2}),$$

yielding that

$$\sum_{i=1}^{4} a_i P(M((a_i, b_i)|\alpha_n)) = \frac{1}{2} \text{ and } \sum_{i=1}^{4} b_i P(M((a_i, b_i)|\alpha_n)) = \frac{1}{2}.$$

This implies that all the elements of α_n cannot lie only on one side of the line h nor only on one side of the line v.

Next, we claim that at least a pair of points of α_n lie in the region above h and at least another pair lie in the region below of h. Suppose that there is only one point of α_n that lies above the line h. Due to symmetry, we can assume that the point lies on the line v. Then, for $x \in J_{33} \cup J_{44}$, $\min_{a\in\alpha_n} \|x-a\|^2 \geq (\frac{1}{2}-\frac{1}{9})^2 = \frac{49}{324}$, and for $x \in A = J_{31} \cup J_{32} \cup J_{34} \cup J_{41} \cup J_{42} \cup J_{43}$, $\min_{a\in\alpha_n} \|x-a\|^2 \geq (\frac{1}{2}-\frac{1}{3})^2 = \frac{1}{36}$. Then

$$\begin{aligned}\int \min_{a\in\alpha_n} \|x-a\|^2 dP &> \int_{J_{33}\cup J_{44}} \min_{a\in\alpha_n} \|x-a\|^2 dP + \int_A \min_{a\in\alpha_n} \|x-a\|^2 dP \\ &\geq 2\cdot\frac{49}{324}\cdot\frac{1}{16} + 6\cdot\frac{1}{36}\cdot\frac{1}{16} = \frac{19}{648} > V_n,\end{aligned}$$

which is a contradiction, proving the claim. It follows in the same manner that at least two of the points of α_n lie on one side of the line v and at least two of the points of α_n lie on the other side the line v. Therefore, α_n contains points from each of the four quadrants. Since $supp(P) \subset J_1 \cup J_2 \cup J_3 \cup J_4$ and P is symmetrically distributed over J, we deduce that α_n contains points from each J_i, $1 \leq i \leq 4$; i.e., $\alpha_n \cap J_i \neq \emptyset$ for $1 \leq i \leq 4$. $\square$

Below, for any $1 \leq i, j \leq 4$, we will use the notation $ij^{(k)}$ for the concatenation of i with $j^{(k)}$, where $j^{(k)}$ denotes the k-times concatenation of j with itself.

Lemma 3.12. *If $n \geq 4$ and α_n is an optimal set of n-means, then $\alpha_n \cap (J \setminus \overset{4}{\underset{i=1}{\cup}} J_i) = \emptyset$.*

Proof. First, assume that $n = m4^{\ell(n)}$, where $m, \ell(n) \in \mathbb{N}$. Then, due to the maximum symmetry (i.e., symmetry with respect to h, l, r, v, and P) and Lemma 3.11, the optimal set α_n contains $m4^{\ell(n)-1}$ elements from each of J_i, $1 \leq i \leq 4$. By Proposition 1.1, the Voronoi region of each point in an optimal set of n-means has positive probability measure. Since support of P lies in $\cup_{i=1}^4 J_i$, it follows that $\alpha_n \cap (J \setminus \overset{4}{\underset{i=1}{\cup}} J_i) = \emptyset$. Thus, it remains to prove the assertion for $n = m4^{\ell(n)} + k$, $1 \leq k \leq 3$.

Suppose $\alpha_n \cap (J \setminus \overset{4}{\underset{i=1}{\cup}} J_i) \neq \emptyset$, $n = m4^{\ell(n)} + k$, $1 \leq k \leq 3$. Again, due to the maximum symmetry, without loss of generality, we can assume that

(i) when $k = 1$, α_n contains $m4^{\ell(n)-1}$ elements from each of J_i, and the remaining one element is the center $(\frac{1}{2}, \frac{1}{2})$;

(ii) when $k = 2$, α_n contains $m4^{\ell(n)-1}$ elements from each J_i, and the remaining two elements are the points $(\frac{1}{2}, \frac{1}{6})$ and $(\frac{1}{2}, \frac{5}{6})$; and

(iii) when $k = 3$, α_n contains $m4^{\ell(n)-1} + 1$ elements from each J_1 and J_2, and $m4^{\ell(n)-1}$ elements from each of J_3 and J_4, and the remaining one element is $(\frac{1}{2}, \frac{5}{6})$.

In all these cases, if the Voronoi regions $M((\centerdot, \frac{1}{2})|\alpha_n)$ of these points do not contain any point from $\underset{i=1}{\overset{4}{\cup}} J_i$, then $P(M((\centerdot, \frac{1}{2})|\alpha_n)) = 0$, and this contradicts Proposition 1.1.

In case (i), if $M((\frac{1}{2}, \frac{1}{2})|\alpha_n)$ contains points from J_i for all $1 \leq i \leq 4$, then there is a positive integer k such that $J_{14^{(k)}} \cup J_{23^{(k)}} \cup J_{32^{(k)}} \cup J_{41^{(k)}} \subset M((\frac{1}{2}, \frac{1}{2})|\alpha_n)$. Then the distortion error due to the set α_n is larger than the distortion error due to the set α_n when the extra one point is moved to any of the children J_i, $1 \leq i \leq 4$. This contradicts the fact that α_n is an optimal set of n-means. Thus, $\alpha_n \cap (J \setminus \underset{i=1}{\overset{4}{\cup}} J_i) = \emptyset$.

In case (ii), if $M((\frac{1}{2}, \frac{1}{6})|\alpha_n)$ and $M((\frac{1}{2}, \frac{5}{6})|\alpha_n)$ contain points from $\underset{i=1}{\overset{4}{\cup}} J_i$, then $M((\frac{1}{2}, \frac{1}{6})|\alpha_n)$ will contain points from both J_1 and J_2, and $M((\frac{1}{2}, \frac{5}{6})|\alpha_n)$ will contain points from both J_3 and J_4. But, in that situation, the distortion error due to the set α_n is larger than the distortion error due to the set α_n when the point $(\frac{1}{2}, \frac{1}{6})$ is moved to either J_1 or J_2, and the point $(\frac{1}{2}, \frac{5}{6})$ is moved to either J_3 or J_4, contradicting the optimality of α_n. Thus, $\alpha_n \cap (J \setminus \underset{i=1}{\overset{4}{\cup}} J_i) = \emptyset$.

In case (iii), if $M((\frac{1}{2}, \frac{5}{6})|\alpha_n)$ contains points from $\underset{i=1}{\overset{4}{\cup}} J_i$, then we see that $M((\frac{1}{2}, \frac{5}{6})|\alpha_n)$ will contain points from both J_3 and J_4. But then the distortion error due to the set α_n is larger than the distortion error due to the set α_n when the point $(\frac{1}{2}, \frac{5}{6})$ is moved to either J_3 or J_4, contradicting the optimality of α_n. Thus, $\alpha_n \cap (J \setminus \underset{i=1}{\overset{4}{\cup}} J_i) = \emptyset$. □

In the next proposition, σ^- will denote the word $\sigma_1\sigma_2 \cdots \sigma_{n-1}$, where $\sigma = \sigma_1\sigma_2 \cdots \sigma_n$.

Proposition 3.13. *Let α_n be an optimal set of n-means for $n \geq 4^{\ell(n)}$ for some $\ell(n) \in \mathbb{N}$. Then, for each $\sigma \in I^{\ell(n)}$, $\alpha_n \cap J_\sigma \neq \emptyset$ and $\alpha_n \cap (J_{\sigma^-} \setminus \underset{i=1}{\overset{4}{\cup}} J_{\sigma^- i}) = \emptyset$.*

Proof. First, we will consider the case $n = 4^{\ell(n)}$ for some $\ell(n) \in \mathbb{N}$. If $\ell(n) = 1$, $n = 4$, then the assertion follows from Lemma 3.11 and Lemma 3.12. So let $\ell(n) \geq 2$. Then, by Lemma 3.11, Lemma 3.12, and the symmetry of P, the set α_n contains $4^{\ell(n)-1}$ elements from each of J_{i_1}, where $1 \leq i_1 \leq 4$. Since the probability measure $P \circ S_{i_1}^{-1}$ on J_{i_1} is also symmetric as the probability measure P on J, by applying Lemma 3.11 and Lemma 3.12 on J_{i_1}, we see that α_n contains $4^{\ell(n)-2}$ elements from each of $J_{i_1 i_2}$, where $1 \leq i_2 \leq 4$. Proceeding in this way inductively, we see that α_n contains one element from each of J_σ, where $\sigma \in I^{\ell(n)}$; in other words, $\alpha_n \cap J_\sigma \neq \emptyset$ for $\sigma \in I^{\ell(n)}$. Since $\alpha_n \cap J_{\sigma^-} = \cup_{i=1}^4 \alpha_n \cap J_{\sigma^- i}$, it follows that $\alpha_n \cap (J_{\sigma^-} \setminus \bigcup\limits_{i=1}^{4} J_{\sigma^- i}) = \emptyset$.

Now let $n = 4^{\ell(n)} + 1$ for some $\ell(n) \in \mathbb{N}$. Again, by Lemma 3.11, Lemma 3.12, and the symmetry of P, there exists an element $i_1 \in I$ such that α_n contains $4^{\ell(n)-1} + 1$ elements from J_{i_1}. Applying Lemma 3.11 and Lemma 3.12 on J_{i_1} again, and proceeding in this way inductively, we can see that there exists a word $\sigma \in I^{\ell(n)}$, where σ is an extension of i_1, such that α_n contains two elements from J_σ, and α_n contains only one element from each J_τ for $\tau \in I^{\ell(n)}$ with $\tau \neq \sigma$. Thus, the proposition is true for $n = 4^{\ell(n)} + 1$.

When $n = 4^{\ell(n)} + 2$ for some $\ell(n) \in \mathbb{N}$, via Lemma 3.11, Lemma 3.12, and the symmetry of P, there exists two elements $i_1, i_2 \in I$, such that α_n contains $4^{\ell(n)-1}+1$ elements from each of J_{i_1} and J_{i_2} and $4^{\ell(n)-1}$ elements from each of J_j for $j \in I \setminus \{i_1, i_2\}$. Applying Lemma 3.11 and Lemma 3.12 on both J_{i_1} and J_{i_2} again, and proceeding in this way inductively, we can see that there exist words $\sigma, \tau \in I^{\ell(n)}$, where σ is an extension of i_1 and τ is an extension of i_2, such that α_n contains two elements from each of J_σ and J_τ, and α_n contains only one element from each J_δ for $\delta \in I^{\ell(n)} \setminus \{\sigma, \tau\}$. Thus, the proposition is true for $n = 4^{\ell(n)} + 2$.

If $n = 4^{\ell(n)} + 3$, the assertion is obtained similarly to the case $n = 4^{\ell(n)} + 1$. Thus, we deduce that the statement is valid for any $n \geq 4^{\ell(n)}$, $\ell(n) \in \mathbb{N}$. □

By Lemma 2.5, the set $\alpha_1 = \{(\frac{1}{2}, \frac{1}{2})\}$ is the only optimal set of one-mean with quantization error $V = \frac{1}{4}$. By Proposition 3.2 and Proposition 3.7, the sets $\alpha_2 = \{(\frac{1}{2}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ and $\alpha_3 = \{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ are optimal sets of two- and three-means with quantization error $\frac{5}{36}$ and $\frac{1}{12}$, respectively. The sets α_2 and α_3 are not the only optimal sets of two- and three-means; indeed, the total number of optimal sets of two-means is two and the total number of optimal sets of three-means is four. With this, the optimal sets of n-means for all $n \geq 4$, their numbers, and the associated quantization error are determined by the following theorem.

Theorem 3.14. *Let P be a Borel probability measure on $\mathbb{R}^2$ supported by the Cantor dust S. Let $n \geq 4$, $1 \leq m \leq 3$. Then*

(i) *if $n = m4^{\ell(n)}$ for some $\ell(n) \in \mathbb{Z}^+$, then $\alpha_n = \{S_\sigma(\alpha_m) : \sigma \in I^{\ell(n)}\}$ is an optimal set of n-means. The number of such sets is*

$(2^{m-1})^{4^{\ell(n)}}$ *and the corresponding quantization error is*

$$V_n = \sum_{\sigma \in I^{\ell(n)}} \int_{J_\sigma} \min_{a \in S_\sigma(\alpha_m)} \|x-a\|^2 dP.$$

(ii) *if* $n = m4^{\ell(n)} + k$, *where* $1 \leq k < 4^{\ell(n)}$ *for some* $\ell(n) \in \mathbb{Z}^+$, *and* $t \subset I^{\ell(n)}$ *with card*$(t) = k$, *then* $\alpha_n(t) = \{S_\sigma(\alpha_m) : \sigma \in I^{\ell(n)} \setminus t\} \cup \{S_\sigma(\alpha_{m+1}) : \sigma \in t\}$ *is an optimal set of* n*-means. The number of such sets is* $(2^{m-1})^{4^{\ell(n)}-k} \cdot {}^{4^{\ell(n)}}C_k \cdot 2^{mk}$ *if* $m = 1, 2$ *and* $(2^{m-1})^{4^{\ell(n)}-k} \cdot {}^{4^{\ell(n)}}C_k$ *if* $m = 3$*; the corresponding quantization error is*

$$V_n = \sum_{\sigma \in I^{\ell(n)} \setminus t} \int_{J_\sigma} \min_{a \in S_\sigma(\alpha_m)} \|x-a\|^2 dP$$
$$+ \sum_{\sigma \in t} \int_{J_\sigma} \min_{a \in S_\sigma(\alpha_{m+1})} \|x-a\|^2 dP,$$

where ${}^uC_v = \binom{u}{v}$, *the binomial coefficients.*

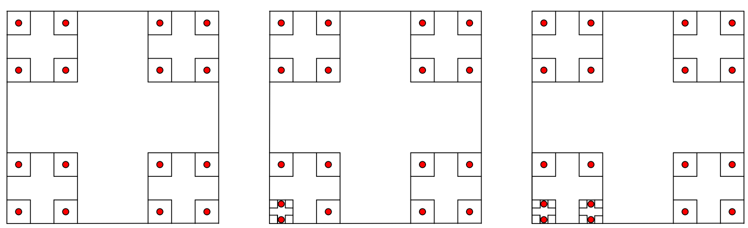


FIGURE 4. Optimal configuration of n points for $16 \leq n \leq 18$.

Proof. Let $m = 1, 2, 3$. Let $n = m4^{\ell(n)}$, $1 \leq m \leq 3$, for some $\ell(n) \in \mathbb{N}$. Then, by Proposition 3.13, it follows that α_n contains m elements from each J_σ for $\sigma \in I^{\ell(n)}$, and Proposition 3.10 implies that $S_\sigma^{-1}(\alpha_n \cap J_\sigma)$ is an optimal set of m-means (i.e., $\alpha_n \cap J_\sigma = S_\sigma(\alpha_m)$); thus,

$$\alpha_n = \bigcup_{\sigma \in I^{\ell(n)}} S_\sigma(\alpha_m) = \{S_\sigma(\alpha_m) : \sigma \in I^{\ell(n)}\}.$$

Since α_m can be chosen in 2^{m-1} different ways, the number of such sets is $(2^{m-1})^{4^{\ell(n)}}$, and the corresponding quantization error is given by

$$V_n = \int \min_{a\in\alpha_n} \|x-a\|^2 dP = \sum_{\sigma\in I^{\ell(n)}} \int_{J_\sigma} \min_{a\in\alpha_n} \|x-a\|^2 dP$$
$$= \sum_{\sigma\in I^{\ell(n)}} \int_{J_\sigma} \min_{a\in S_\sigma(\alpha_m)} \|x-a\|^2 dP.$$

To prove (ii), we proceed as follows: Let $t \subset I^{\ell(n)}$ with $\text{card}(t) = k$. Then, by Proposition 3.13, it follows that α_n contains m elements from J_σ for each $\sigma \in I^{\ell(n)} \setminus t$ and $(m+1)$ elements from J_σ for $\sigma \in t$. In other words, $\alpha_n \cap J_\sigma = S_\sigma(\alpha_m)$ for $\sigma \in I^{\ell(n)} \setminus t$ and $\alpha_n \cap J_\sigma = S_\sigma(\alpha_{m+1})$ for $\sigma \in t$. Thus,

$$\alpha_n(t) = \{S_\sigma(\alpha_m) : \sigma \in I^{\ell(n)} \setminus t\} \cup \{S_\sigma(\alpha_{m+1}) : \sigma \in t\}.$$

The corresponding quantization error is given by

$$V_n = \int \min_{a\in\alpha_n} \|x-a\|^2 dP$$
$$= \sum_{\sigma\in I^{\ell(n)}\setminus t} \int_{J_\sigma} \min_{a\in\alpha_n} \|x-a\|^2 dP + \sum_{\sigma\in t} \int_{J_\sigma} \min_{a\in\alpha_n} \|x-a\|^2 dP$$
$$= \sum_{\sigma\in I^{\ell(n)}\setminus t} \int_{J_\sigma} \min_{a\in S_\sigma(\alpha_m)} \|x-a\|^2 dP + \sum_{\sigma\in t} \int_{J_\sigma} \min_{a\in S_\sigma(\alpha_{m+1})} \|x-a\|^2 dP.$$

Recall that α_2 can be chosen in two different ways, α_3 can be chosen in three different ways, and α_4 can be chosen in only one way. Thus, if $m=1$, the number of α_n is ${}^{4^{\ell(n)}}C_k \cdot 2^k$; if $m=2$, the number of α_n is $2^{4^{\ell(n)}-k} \cdot {}^{4^{\ell(n)}}C_k \cdot 4^k$; and if $m=3$, the number of α_n is $4^{4^{\ell(n)}-k} \cdot {}^{4^{\ell(n)}}C_k$. Hence, the proof of the theorem is complete. $\square$

4. Quantization Dimension and Quantization Coefficient

Since the standard Cantor dust S under investigation satisfies the strong separation condition, with each S_i having contracting factor of $\frac{1}{3}$, its Hausdorff dimension is equal to the similarity dimension. Hence, from the equation $4(\frac{1}{3})^\beta = 1$, we have $\dim_{\mathrm{H}}(S) = \beta = \frac{\log 4}{\log 3}$. By [7, Theorem 14.17], the quantization dimension $D_2(P)$ exists and is equal to β.

Next, we proceed to show that β dimensional quantization coefficient for P does not exist. The proof of it follows the same general strategy as [6, Theorem 6.3]. We begin with the following simple observation. If $f : [1,2] \to \mathbb{R}$ is a function defined by $f(x) = \frac{1}{36}x^{\frac{2}{\beta}}(13-x)$, then it is strictly increasing on $[1,2]$; hence; $f([1,2]) = [\frac{1}{3}, \frac{11}{12}]$.

Theorem 4.1. *β-dimensional quantization coefficient for P does not exist.*

Proof. Let $(n_k)_{k\in\mathbb{N}}$ be a subsequence of the set of natural numbers such that $4^{\ell(n_k)} \leq n_k < 2\cdot 4^{\ell(n_k)}$. The assertion of the theorem will follow if we show that the set of accumulation points of the sequence $(n_k^{\frac{2}{\beta}} V_{n_k})_{k\geq 1}$ is $[\frac{1}{3}, \frac{11}{12}]$. Let $y \in [\frac{1}{3}, \frac{11}{12}]$, then $y = f(x)$ for some $x \in [1,2]$. Set $n_{k_\ell} = \lfloor x4^\ell \rfloor$, where $\lfloor x4^\ell \rfloor$ denotes the greatest integer less than or equal to $x4^\ell$. Then $n_{k_\ell} < n_{k_{\ell+1}}$ and $\ell(n_{k_\ell}) = \ell$, and there exists $x_{k_\ell} \in [1,2]$ such that $n_{k_\ell} = x_{k_\ell}4^\ell$. Recall that by $\ell(n_{k_\ell}) = \ell$, it is meant that $4^\ell \leq n_{k_\ell} < 4^{\ell+1}$. Notice that if $4^{\ell(n)} \leq n \leq 4^{\ell(n)+1}$, then, by Theorem 3.14, we have

$$\begin{aligned} V_n &= (2\cdot 4^{\ell(n)} - n)\frac{1}{36^{\ell(n)}}\frac{1}{4} + (n - 4^{\ell(n)})\frac{1}{36^{\ell(n)}}\frac{5}{36} \\ &= \frac{1}{36^{\ell(n)+1}}(13\cdot 4^{\ell(n)} - 4n). \end{aligned}$$

Thus, putting the values of n_{k_ℓ} and $V_{n_{k_\ell}}$, we obtain

$$\begin{aligned} n_{k_\ell}^{\frac{2}{\beta}} V_{n_{k_\ell}} &= n_{k_\ell}^{\frac{2}{\beta}} \frac{1}{36^{\ell+1}}(13\cdot 4^\ell - 4n_k) = x_{k_\ell}^{\frac{2}{\beta}} 4^{\frac{2}{\beta}} \frac{1}{36^{\ell+1}}(13\cdot 4^\ell - 4x_{k_\ell}4^\ell) \\ &= x_{k_\ell}^{\frac{2}{\beta}} 9^\ell \frac{1}{36^{\ell+1}}(13\cdot 4^\ell - 4x_{k_\ell}4^\ell), \end{aligned}$$

which yields

$$n_{k_\ell}^{\frac{2}{\beta}} V_{n_{k_\ell}} = \frac{1}{36} x_{k_\ell}^{\frac{2}{\beta}}(13 - 4x_{k_\ell}) = f(x_{k_\ell}). \tag{6}$$

Again, $x_{k_\ell}4^\ell \leq x4^\ell < x_{k_\ell}4^\ell + 1$, which implies $x - \frac{1}{4^\ell} < x_{k_\ell} \leq x$, and so $\lim\limits_{\ell\to\infty} x_{k_\ell} = x$. Since f is continuous, we have

$$\lim_{\ell\to\infty} n_{k_\ell}^{\frac{2}{\beta}} V_{n_{k_\ell}} = f(x) = y,$$

which yields the fact that y is an accumulation point of the subsequence $(n_k^{\frac{2}{\beta}} V_{n_k})_{k\geq 1}$ whenever $y \in [\frac{1}{3}, \frac{11}{12}]$. To prove the converse, let y be an accumulation point of the sequence $(n_k^{\frac{2}{\beta}} V_{n_k})_{k\geq 1}$. Then there exists a subsequence $(n_{k_i}^{\frac{2}{\beta}} V_{n_{k_i}})_{i\geq 1}$ of $(n_k^{\frac{2}{\beta}} V_{n_k})_{k\geq 1}$ such that $\lim\limits_{i\to\infty} n_{k_i}^{\frac{2}{\beta}} V_{n_{k_i}} = y$. Set $\ell_{k_i} = \ell(n_{k_i})$ and $x_{k_i} = \frac{n_{k_i}}{4^{\ell_{k_i}}}$. Then $x_{k_i} \in [1,2]$ and, as shown in (6), we have

$$n_{k_i}^{\frac{2}{\beta}} V_{n_{k_i}} = f(x_{k_i}).$$

Let $(x_{k_{i_j}})_{j\geq 1}$ be a convergent subsequence of $(x_{k_i})_{i\geq 1}$, then we obtain

$$y = \lim_{i\to\infty} n_{k_i}^{\frac{2}{\beta}} V_{n_{k_i}} = \lim_{j\to\infty} n_{k_{i_j}}^{\frac{2}{\beta}} V_{n_{k_{i_j}}} = \lim_{j\to\infty} f(x_{k_{i_j}}) \in [\frac{1}{3}, \frac{11}{12}].$$

Thus, the set of accumulation points of the sequence $(n_k^{\frac{2}{\beta}} V_{n_k})_{k\geq 1}$ is $[\frac{1}{3}, \frac{11}{12}]$, which yields the proof of the theorem. $\square$

References

[1] Efren F. Abaya and Gary L. Wise, *Some remarks on the existence of optimal quantizers*, Statist. Probab. Lett. **2** (1984), no. 6, 349–351.

[2] Doğan Çömez and Mrinal Kanti Roychowdhury, *Quantization for uniform distributions on stretched Sierpiński triangles*, Monatsh. Math. **190** (2019), no. 1, 79–100.

[3] Carl P. Dettmann and Mrinal Kanti Roychowdhury, *Quantization for uniform distributions on equilateral triangles*, Real Anal. Exchange **42** (2017), no. 1, 149–166.

[4] Qiang Du, Vance Faber, and Max Gunzburger, *Centroidal Voronoi tessellations: Applications and algorithms*, SIAM Rev. **41** (1999), no. 4, 637–676.

[5] Allen Gersho and Robert M. Gray, *Vector Quantization and Signal Compression.* New York: Kluwer Academic Publishers, 1992.

[6] S. Graf and H. Luschgy, *The quantization of the Cantor distribution*, Math. Nachr. **183** (1997), 113–133.

[7] Siegfried Graf and Harald Luschgy, *Foundations of Quantization for Probability Distributions.* Lecture Notes in Mathematics, 1730. Berlin: Springer-Verlag, 2000.

[8] R. M. Gray, J. C. Kieffer, and Y. Linde, *Locally optimal block quantizer design*, Inform. and Control **45** (1980), no. 2, 178–198.

[9] Robert M. Gray and David L. Neuhoff, *Quantization*, IEEE Trans. Inform. Theory **44** (1998), no. 6, 2325–2383.

[10] András György and Tamás Linder, *On the structure of optimal entropy-constrained scalar quantizers*, IEEE Trans. Inform. Theory **48** (2002), no. 2, 416–427.

[11] John E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30** (1981), no. 5, 713–747.

[12] Joseph Rosenblatt and Mrinal Kanti Roychowdhury, *Optimal quantization for piecewise uniform distributions*, Unif. Distrib. Theory **13** (2018), no. 2, 23–55.

[13] Mrinal Kanti Roychowdhury, *Quantization and centroidal Voronoi tessellations for probability measures on dyadic Cantor sets*, J. Fractal Geom. **4** (2017), no. 2, 127–146.

[14] Mrinal Kanti Roychowdhury, *Optimal quantizers for some absolutely continuous probability measures*, Real Anal. Exchange **43** (2018), no. 1, 105–136.

[15] Ram Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation and Multiuser Information Theory.* Cambridge: Cambridge Universitgy Press, 2014.

(Çömez) Department of Mathematics; 408E24 Minard Hall; North Dakota State University; Fargo, ND 58108-6050 USA
Email address: Dogan.Comez@ndsu.edu

(Roychowdhury) School of Mathematical and Statistical Sciences; University of Texas Rio Grande Valley; 1201 West University Drive; Edinburg, TX 78539-2999 USA
Email address: mrinal.roychowdhury@utrgv.edu