

University of Texas Rio Grande Valley

ScholarWorks @ UTRGV

Mathematical and Statistical Sciences Faculty
Publications and Presentations

College of Sciences

4-25-2022

Quantization for Infinite Affine Transformations

Dogan Comez

Mrinal Kanti Roychowdhury

Follow this and additional works at: https://scholarworks.utrgv.edu/mss_fac



Part of the [Mathematics Commons](#)



Article

Quantization for Infinite Affine Transformations

Doğan Çömez¹ and Mrinal Kanti Roychowdhury^{2,*}

¹ Department of Mathematics, 408E24 Minard Hall, North Dakota State University, Fargo, ND 58108-6050, USA; dogan.comez@ndsu.edu

² School of Mathematical and Statistical Sciences, University of Texas Rio Grande Valley, 1201 West University Drive, Edinburg, TX 78539-2999, USA

* Correspondence: mrinal.roychowdhury@utrgv.edu

Abstract: Quantization for a probability distribution refers to the idea of estimating a given probability by a discrete probability supported by a finite set. In this article, we consider a probability distribution generated by an infinite system of affine transformations $\{S_{ij}\}$ on \mathbb{R}^2 with associated probabilities $\{p_{ij}\}$ such that $p_{ij} > 0$ for all $i, j \in \mathbb{N}$ and $\sum_{i,j=1}^{\infty} p_{ij} = 1$. For such a probability measure P , the optimal sets of n -means and the n th quantization error are calculated for every natural number n . It is shown that the distribution of such a probability measure is the same as that of the direct product of the Cantor distribution. In addition, it is proved that the quantization dimension $D(P)$ exists and is finite; whereas, the $D(P)$ -dimensional quantization coefficient does not exist, and the $D(P)$ -dimensional lower and the upper quantization coefficients lie in the closed interval $[\frac{1}{12}, \frac{5}{4}]$.

Keywords: affine transformations; affine set; affine measure; optimal quantizers; quantization error

MSC: 60Exx; 28A80; 94A34



Citation: Çömez, D.; Roychowdhury, M.K. Quantization for Infinite Affine Transformations. *Fractal Fract.* **2022**, *6*, 239. <https://doi.org/10.3390/fractalfract6050239>

Academic Editor: Carlo Cattani

Received: 4 April 2022

Accepted: 24 April 2022

Published: 25 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The quantization problem for probability measures is concerned with approximating a given measure by discrete measures of finite support in L_r -metrics. This problem has roots in information theory and engineering technology, in particular in signal processing and pattern recognition [1,2]. For a Borel probability measure P on \mathbb{R}^d , a quantizer is a function q mapping d -dimensional vectors in the domain $\Omega \subset \mathbb{R}^d$ into a finite set of vectors $\alpha \subset \mathbb{R}^d$. In this case, the error $\int \min_{a \in \alpha} \|x - a\|^2 dP(x)$, where $\|\cdot\|$ is the Euclidean norm \mathbb{R}^d , is often referred to as the variance, cost, or distortion error for α with respect to the measure P , and is denoted by $V(\alpha) := V(P; \alpha)$. The value $\inf\{V(P; \alpha) : \alpha \subset \mathbb{R}^d, \text{card}(\alpha) \leq n\}$ is called the n th quantization error for the P , and is denoted by $V_n := V_n(P)$. A set α on which this infimum is attained and contains no more than n points is called an optimal set of n -means. The elements of an optimal set are called optimal quantizers. It is known that for a Borel probability measure P if its support contains infinitely many elements and $\int \|x\|^2 dP(x)$ is finite, then an optimal set of n -means always has exactly n -elements [3–6]. The number $\lim_{n \rightarrow \infty} \frac{2 \log n}{-\log V_n(P)}$, if exists, is called the *quantization dimension* of the measure P , and is denoted by $D(P)$; likewise, for any $s \in (0, +\infty)$, the number $\lim_{n \rightarrow \infty} n^{\frac{2}{s}} V_n(P)$, if it exists, is called the s -dimensional quantization coefficient for P .

For a finite set $\alpha \subset \mathbb{R}^d$, the Voronoi region generated by $a \in \alpha$, denoted by $M(a|\alpha)$, is the set of all points in \mathbb{R}^d which are closer to $a \in \alpha$ than to all other elements in α . For a probability distribution P on \mathbb{R}^d the centroids of the regions $M(a|\alpha)$ are given by $a^* = \frac{1}{P(M(a|\alpha))} \int_{M(a|\alpha)} x dP$. A Voronoi tessellation is called a centroidal Voronoi tessellation (CVT) if $a^* = a$, i.e., if the generators are also the centroids of their own Voronoi regions. For a Borel probability measure P on \mathbb{R}^d , an optimal set of n -means forms a CVT; however, the converse is not true in general [7,8]. The following fact is known [6,9]:

Proposition 1. Let α be an optimal set of n -means and $a \in \alpha$. Then,

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

Let $X = \mathbb{R}$ and consider the probability distribution $P_c := \frac{1}{2}P_c \circ U_1^{-1} + \frac{1}{2}P_c \circ U_2^{-1}$, where $U_1(x) = \frac{1}{3}x$ and $U_2(x) = \frac{1}{3}x + \frac{2}{3}$, for all $x \in \mathbb{R}$. Because its support is the standard Cantor set generated by U_1 and U_2 , P_c is called the Cantor distribution. S. Graf and H. Luschgy determined the optimal sets of n -means and the n th quantization errors for the Cantor distribution, for all $n \geq 1$, completing its quantization program [10]. This result has been extended to the setting of a nonuniform Cantor distribution by L. Roychowdhury [11]. Analogously, the Cantor dust is generated by the contractive mappings $\{S_i\}_{i=1}^4$ on \mathbb{R}^2 , where $S_1(x_1, x_2) = \frac{1}{3}(x_1, x_2)$, $S_2(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (\frac{2}{3}, 0)$, $S_3(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (0, \frac{2}{3})$, and $S_4(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (\frac{2}{3}, \frac{2}{3})$. If P is a Borel probability measure on \mathbb{R}^2 such that $P = \frac{1}{4}P \circ S_1^{-1} + \frac{1}{4}P \circ S_2^{-1} + \frac{1}{4}P \circ S_3^{-1} + \frac{1}{4}P \circ S_4^{-1}$, then P has support the Cantor dust. For this measure, D. Çömez and M.K. Roychowdhury determined the optimal sets of n -means and the n th quantization errors [12]. Let P be a probability measure on \mathbb{R} generated by an infinite collection of similitudes $\{S_j\}_{j=1}^\infty$, where $S_j(x) = \frac{1}{3^j}x + 1 - \frac{1}{3^{j-1}}$ for all $x \in \mathbb{R}$ and P is given by $P = \sum_{j=1}^\infty \frac{1}{2^j}P \circ S_j^{-1}$. For this measure, M.K. Roychowdhury determined the optimal sets of n -means and the n th quantization errors [13], which is an infinite extension of the result of S. Graf and H. Luschgy in [10]. The quantization dimension for probability distributions generated by an infinite collection of similitudes was determined by E. Mihailescu and M.K. Roychowdhury in [14], which is an infinite extension of the result of S. Graf and H. Luschgy in [15]. In this article, we study extension of the result of D. Çömez and M.K. Roychowdhury in [12] to the setting of countably infinite affine maps on \mathbb{R}^2 , which will also complete the program initiated in [14].

Let $\{S_{(i,j)} : i, j \in \mathbb{N}\}$ be a collection of countably infinite affine transformations on \mathbb{R}^2 , where $S_{(i,j)}(x_1, x_2) = (r^i x_1 + 1 - r^{i-1}, r^j x_2 + 1 - r^{j-1})$, where $0 < r \leq \frac{1}{3}$. Clearly, these affine transformations are all contractive but are not similarity mappings. Associate the mappings $S_{(i,j)}$ with the probabilities $p_{(i,j)}$ such that $p_{(i,j)} = \frac{1}{2^{i+j}}$ for all $i, j \in \mathbb{N}$, where $\mathbb{N} := \{1, 2, 3, \dots\}$. Then, there exists a unique Borel probability measure P on \mathbb{R}^2 ([16–18], etc.) such that

$$P = \sum_{i,j=1}^{\infty} p_{(i,j)} P \circ S_{(i,j)}^{-1}.$$

The support of such a probability measure lies in the unit square $[0, 1]^2$. We call such a measure an affine measure on \mathbb{R}^2 , or more specifically, an *infinitely generated affine measure* on \mathbb{R}^2 . This article deals with the quantization of this measure P . The arrangement of the paper is as follows: in Section 2, we discuss the basic definitions and lemmas about the optimal sets of n -means and the n th quantization errors. The arguments in this section point out that determining optimal sets of n -means and the n th quantization errors for all $n \geq 3$ and for arbitrary $r \in (0, \frac{1}{3})$ require very intricate and complicated analysis; hence, for clarity purposes, in the remaining sections the focus will be on the case $r = \frac{1}{3}$. Section 3 is devoted to determining the optimal sets of n -means for $n = 2$ and $n = 3$. In Section 4, we first define a mapping F which enables us to convert the infinitely generated affine measure P to a finitely generated product measure $P_c \times P_c$, each P_c is the Cantor distribution. Having this connection between P and P_c ; together with the optimal sets of n -means for $n = 1, 2, 3$, in Section 5 we will utilize the dynamics of the affine maps to obtain the main results of the paper: closed formulas to determine the optimal sets of n -means and the corresponding quantization errors for all $n \geq 4$. For clarity of the exposition, we also provide some examples and figures to illustrate the constructions. Lastly, having closed form for the quantization errors for each n , we prove the existence of the quantization dimension $D(P)$ and show that the $D(P)$ -dimensional quantization coefficient for P does not exist (but are

finite) and the $D(P)$ -dimensional lower and the upper quantization coefficients lie in the closed interval $[\frac{1}{12}, \frac{5}{4}]$.

The results and the arguments in this article are not straightforward generalizations of those in [13]; in particular, this is the case for optimal sets. By the nature of the affine transformations considered in this paper, the optimal sets of order $n = k^2$, $k \geq 1$, are the same as the cross product of optimal sets of order k obtained in [13]; however, the same cannot be said for other $n \geq 3$. Clearly, for n a prime number, optimal sets of n -means cannot be obtained this way. Furthermore, as will be seen from the main theorem, even for $n = kl$, optimal sets of n -means are not the same as the cross product of optimal sets of k - and l -means in [13]. For example, optimal sets of 2- and 3-means in [13] are $\{\frac{1}{6}, \frac{5}{6}\}$ and $\{\frac{1}{6}, \frac{13}{18}, \frac{17}{18}\}$ (or $\{\frac{1}{18}, \frac{5}{18}, \frac{5}{6}\}$), respectively; hence, the cross product of these sets produce some of the optimal sets of 6-means. On the other hand, one of the optimal sets of 6-means is $\{(\frac{1}{18}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{13}{18}, \frac{1}{6}), (\frac{1}{18}, \frac{5}{6}), (\frac{13}{18}, \frac{5}{6}), (\frac{5}{6}, \frac{5}{6})\}$, which cannot be obtained as the cross product of optimal sets of 2- and 3-means in [13].

2. Preliminaries

Let P be the affine measure on \mathbb{R}^2 generated by the affine maps $\{S_{(i,j)} : i, j \in \mathbb{N}\}$ defined above. Consider the alphabet $\mathcal{I} = \mathbb{N}^2 = \{(i, j) : i, j \in \mathbb{N}\}$. By a "string" or a "word" ω over \mathcal{I} , it is meant a finite sequence $\omega := \omega_1\omega_2 \cdots \omega_k$ of symbols from the alphabet, $k \geq 1$, where k is called the length of the word ω . A word of length zero is called the "empty word", and is denoted by \emptyset . By \mathcal{I}^* we denote the set of all words over the alphabet \mathcal{I} of some finite length k , including the empty word \emptyset . By $|\omega|$, we denote the length of a word $\omega \in \mathcal{I}^*$. For any two words $\omega := \omega_1\omega_2 \cdots \omega_k$ and $\tau := \tau_1\tau_2 \cdots \tau_\ell$ in \mathcal{I}^* , by $\omega\tau := \omega_1 \cdots \omega_k\tau_1 \cdots \tau_\ell$ we mean the word obtained from the concatenation of ω and τ . For $n \geq 1$ and $\omega = \omega_1\omega_2 \cdots \omega_n \in \mathcal{I}^*$ we define $\omega^- := \omega_1\omega_2 \cdots \omega_{n-1}$. Note that ω^- is the empty word if the length of ω is one. Analogously, by \mathbb{N}^* we denote the set of all words over the alphabet \mathbb{N} , and for any $\tau \in \mathbb{N}^*$, $|\tau|$, τ^- , etc. are defined similarly. Let $\omega \in \mathcal{I}^k$, $k \geq 1$, be such that $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)$, then $\omega^{(1)}$ and $\omega^{(2)}$ will denote the "coordinate words"; i.e., $\omega^{(1)} := i_1i_2 \cdots i_k$ and $\omega^{(2)} := j_1j_2 \cdots j_k$. Thus, $\omega_{|\omega|}^{(1)} = i_k$ and $\omega_{|\omega|}^{(2)} = j_k$. These lead us to define the following notations: For $\omega \in \mathcal{I}^*$, by $\omega(\emptyset, \infty)$ it is meant the set of all words $\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)} + j)$ obtained by concatenating the word ω^- with the word $(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)} + j)$ for $j \in \mathbb{N}$, i.e.,

$$\omega(\emptyset, \infty) := \{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)} + j) : j \in \mathbb{N}\}.$$

Similarly, $\omega(\infty, \emptyset)$ and $\omega(\infty, \infty)$ represent the sets

$$\omega(\infty, \emptyset) := \{\omega^-(\omega_{|\omega|}^{(1)} + i, \omega_{|\omega|}^{(2)}) : i \in \mathbb{N}\} \text{ and } \omega(\infty, \infty) := \{\omega^-(\omega_{|\omega|}^{(1)} + i, \omega_{|\omega|}^{(2)} + j) : i, j \in \mathbb{N}\},$$

respectively. Analogously, for any $\tau \in \mathbb{N}^*$, by (τ, ∞) it is meant the set $(\tau, \infty) := \{\tau + i : i \in \mathbb{N}\}$, and (τ, \emptyset) represents the set $(\tau, \emptyset) := \{\tau\}$. Thus, if $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)(\infty, \emptyset)$, then we write $\omega^{(1)} := (i_1i_2 \cdots i_k, \infty)$ and $\omega^{(2)} := j_1j_2 \cdots j_k$; if $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)(\emptyset, \infty)$, then we write $\omega^{(1)} := i_1i_2 \cdots i_k$ and $\omega^{(2)} := (j_1j_2 \cdots j_k, \infty)$; and if $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)(\infty, \infty)$, then we write $\omega^{(1)} := (i_1i_2 \cdots i_k, \infty)$ and $\omega^{(2)} := (j_1j_2 \cdots j_k, \infty)$. For $\omega = \omega_1\omega_2 \cdots \omega_k \in \mathcal{I}^k$, $k \geq 1$, let us write

$$S_\omega := S_{\omega_1} \circ \cdots \circ S_{\omega_k}, p_\omega := p_{\omega_1}p_{\omega_2} \cdots p_{\omega_k} \text{ and } J_\omega := S_\omega([0, 1] \times [0, 1]).$$

In particular, $S_\emptyset = I$, the identity mapping on \mathbb{R}^2 , and $J := J_\emptyset = S_\emptyset([0, 1] \times [0, 1])$. Then, the probability measure P supports the closure of the limit set \mathcal{S} , where $\mathcal{S} = \bigcap_{k \in \mathbb{N}} \bigcup_{\omega \in \mathcal{I}^k} J_\omega$. The limit set \mathcal{S} is called the affine set or infinitely generated affine set. For $\omega \in \mathcal{I}^k$ and

$i, j \in \mathbb{N}$, the rectangles $J_{\omega(i,j)}$, into which J_{ω} is split up at the $(k + 1)$ th level are called the children or the basic rectangles of J_{ω} (see Figure 1). For $\omega \in \mathcal{I}^*$, we write

$$J_{\omega(\emptyset, \infty)} := \bigcup_{j=1}^{\infty} J_{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)}+j)}, \quad J_{\omega(\infty, \emptyset)} := \bigcup_{i=1}^{\infty} J_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)})}, \quad J_{\omega(\infty, \infty)} := \bigcup_{i,j=1}^{\infty} J_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}$$

$$p_{\omega(\emptyset, \infty)} := P(J_{\omega(\emptyset, \infty)}) = \sum_{j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)}+j)}, \quad p_{\omega(\infty, \emptyset)} := P(J_{\omega(\infty, \emptyset)}) = \sum_{i=1}^{\infty} p_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)})}, \quad \text{and}$$

$$p_{\omega(\infty, \infty)} := P(J_{\omega(\infty, \infty)}) = \sum_{i,j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}.$$

Notice that for any $\omega \in \mathcal{I}^*$, $p_{\omega(\emptyset, \infty)} = p_{\omega^-} \sum_{j=1}^{\infty} \frac{1}{2^{\omega_{|\omega|}^{(1)} + \omega_{|\omega|}^{(2)} + j}} = p_{\omega^-} p_{\omega_{|\omega|}} \sum_{j=1}^{\infty} \frac{1}{2^j} = p_{\omega^-} p_{\omega_{|\omega|}} = p_{\omega}$; and similarly, $p_{\omega(\infty, \emptyset)} = p_{\omega(\infty, \infty)} = p_{\omega}$.

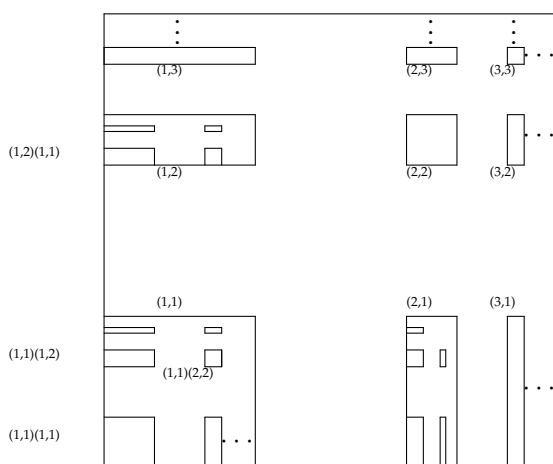


Figure 1. Basic rectangles of the infinite affine transformations.

Because $P = \sum_{i,j=1}^{\infty} p_{(i,j)} P \circ S_{(i,j)}^{-1}$, then, by induction, $P = \sum_{\omega \in \mathcal{I}^k} p_{\omega} P \circ S_{\omega}^{-1}$ for any $k \in \mathbb{N}$.

Hence, we have the following statement:

Lemma 1. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ be Borel measurable and $k \in \mathbb{N}$. Then,

$$\int f dP = \sum_{\omega \in \mathcal{I}^k} p_{\omega} \int f \circ S_{\omega} dP.$$

Let $S_{(i,j)}^{(1)}$ and $S_{(i,j)}^{(2)}$ be the horizontal and vertical components of the transformations $S_{(i,j)}$. Then, for all $(x_1, x_2) \in \mathbb{R}^2$, we have $S_{(i,j)}^{(1)}(x_1) = r^i x_1 + 1 - r^{i-1}$ and $S_{(i,j)}^{(2)}(x_2) = r^j x_2 + 1 - r^{j-1}$; hence, $S_{(i,j)}^{(1)}$ and $S_{(i,j)}^{(2)}$ are similarity mappings on \mathbb{R} with similarity ratios $s_{(i,j)}^{(1)} := r^i$ and $s_{(i,j)}^{(2)} := r^j$, respectively. Similarly, for $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k) \in \mathcal{I}^k, k \geq 1$, let $S_{\omega}^{(1)}$ and $S_{\omega}^{(2)}$ represent the horizontal and vertical components of the transformation S_{ω} on \mathbb{R}^2 . Then, $S_{\omega}^{(1)}$ and $S_{\omega}^{(2)}$ are similarity mappings on \mathbb{R} with similarity ratios $s_{\omega}^{(1)}$ and $s_{\omega}^{(2)}$, respectively, such that $S_{\omega}^{(1)} = S_{(i_1, j_1)}^{(1)} \circ \cdots \circ S_{(i_k, j_k)}^{(1)}$ and $S_{\omega}^{(2)} = S_{(i_1, j_1)}^{(2)} \circ \cdots \circ S_{(i_k, j_k)}^{(2)}$. Thus, it follows that

$$s_{\omega}^{(1)} = s_{(i_1, j_1)}^{(1)} s_{(i_2, j_2)}^{(1)} \cdots s_{(i_k, j_k)}^{(1)} = r^{i_1 + i_2 + \cdots + i_k} \quad \text{and}$$

$$s_{\omega}^{(2)} = s_{(i_1, j_1)}^{(2)} s_{(i_2, j_2)}^{(2)} \cdots s_{(i_k, j_k)}^{(2)} = r^{j_1 + j_2 + \cdots + j_k}.$$

Moreover, we have $P(J_\omega) = p_\omega = p_{(i_1, j_1)} p_{(i_2, j_2)} \cdots p_{(i_k, j_k)} = \frac{1}{2^{i_1+i_2+\dots+i_k+j_1+j_2+\dots+j_k}}$. Let $X := (X_1, X_2)$ be a bivariate random variable with distribution P . Let P_1, P_2 be the marginal distributions of P , i.e., $P_1(A) = P(A \times \mathbb{R}) = P \circ \pi_1^{-1}(A)$ for all $A \in \mathfrak{B}$, and $P_2(B) = P(\mathbb{R} \times B) = P \circ \pi_2^{-1}(B)$ for all $B \in \mathfrak{B}$, where π_1, π_2 are projections given by $\pi_1(x_1, x_2) = x_1$ and $\pi_2(x_1, x_2) = x_2$ for all $(x_1, x_2) \in \mathbb{R}^2$. Here \mathfrak{B} is the Borel σ -algebra on \mathbb{R} . Then, X_1 has distribution P_1 and X_2 has distribution P_2 . Let $S_{(i,j)}^{-(1)}$ and $S_{(i,j)}^{-(2)}$ denote respectively the inverse images of the horizontal and vertical components of the transformations $S_{(i,j)}$ for all $i, j \in \mathbb{N}$. Then, the following lemma is known [16–18]:

Lemma 2. Let P_1 and P_2 be the marginal distributions of the probability measure P . Then,

$$P_1 = \sum_{i=1}^{\infty} \frac{1}{2^i} P_1 \circ S_{(i,j)}^{-(1)} \text{ and } P_2 = \sum_{j=1}^{\infty} \frac{1}{2^j} P_2 \circ S_{(i,j)}^{-(2)}.$$

Remark 1. Since $S_{(i,j)}^{(1)}$ and $S_{(i,j)}^{(2)}$ are similarity mappings, from Lemma 2, one can see that both the marginal distributions P_1 and P_2 are self-similar measures on \mathbb{R} generated by an infinite collection of similarities associated with the probability vector $(\frac{1}{2}, \frac{1}{2^2}, \dots)$.

Lemma 3. Let $E(X)$ and $V(X)$ denote the expectation and the variance of the random variable X . Then,

$$E(X) = (E(X_1), E(X_2)) = (\frac{1}{2}, \frac{1}{2}) \text{ and } V := V(X) = E\|X - (\frac{1}{2}, \frac{1}{2})\|^2 = \frac{1}{4}.$$

Proof. By Lemma 2, $P_1 = P_2 = \mu$, where μ is a unique Borel probability measure on \mathbb{R} such that

$$\mu = \sum_{k=1}^{\infty} \frac{1}{2^k} \mu \circ S_{(k,j)}^{-(1)} = \sum_{k=1}^{\infty} \frac{1}{2^k} \mu \circ S_{(i,k)}^{-(2)}.$$

Hence, $X_1 = X_2$, and by ([11], Lemma 2.2), $E(X_1) = E(X_2) = \frac{1}{2}$, and $V(X_1) = V(X_2) = \frac{1}{8}$, which implies 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. By using the standard rule of probability, for any $(a, b) \in \mathbb{R}^2$, we have $E\|X - (a, b)\|^2 = V + \|(a, b) - (\frac{1}{2}, \frac{1}{2})\|^2$, which yields that the optimal set of one-mean consists of the expected value and the corresponding quantization error is the variance V of the random variable X .

Lemma 4. Let $\omega \in \mathcal{I}^*$. Then,

- (i) $E(X|X \in J_{\omega(\infty, \infty)}) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}(\frac{1}{2}, \frac{1}{2}) + (s_\omega^{(1)} \frac{1}{2}(1-r), s_\omega^{(2)} \frac{1}{2}(1-r));$
- (ii) $E(X|X \in J_{\omega(\emptyset, \infty)}) = S_{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)}+1)}(\frac{1}{2}, \frac{1}{2}) + (0, s_\omega^{(2)} \frac{1}{2}(1-r)),$ and
- (iii) $E(X|X \in J_{\omega(\infty, \emptyset)}) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)})}(\frac{1}{2}, \frac{1}{2}) + (s_\omega^{(1)} \frac{1}{2}(1-r), 0).$

Proof. First prove (i). Because $P(J_{\omega(\infty, \infty)}) = p_{\omega(\infty, \infty)} = p_\omega$ and $p_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)} = p_\omega \frac{1}{2^{i+j}}$,

$$\begin{aligned} E(X|X \in J_{\omega(\infty, \infty)}) &= E(X|X \in \bigcup_{i,j=1}^{\infty} J_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}) \\ &= \frac{1}{P(J_{\omega(\infty, \infty)})} \sum_{i,j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)} S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}(\frac{1}{2}, \frac{1}{2}) = \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}} S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}(\frac{1}{2}, \frac{1}{2}). \end{aligned}$$

Notice that

$$\begin{aligned} & S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}\left(\frac{1}{2}, \frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) \\ &= \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right), S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{1}{2}\right) \right) - \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right), S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right) \right) \\ &= \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right), S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right) \right). \end{aligned}$$

Because

$$\begin{aligned} & S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) = s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) - s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(r^{\omega_{|\omega|}^{(1)}+i}\left(\frac{1}{2}\right) - r^{\omega_{|\omega|}^{(1)}+i-1} - r^{\omega_{|\omega|}^{(1)}+1}\left(\frac{1}{2}\right) + r^{\omega_{|\omega|}^{(1)}+1-1}\right) = s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}r^i - r^{i-1} - \frac{r}{2} + 1\right) \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(1 - \frac{r}{2}\right)\left(1 - r^{i-1}\right), \end{aligned}$$

and similarly

$$S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right) = s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(1 - \frac{r}{2}\right)\left(1 - r^{j-1}\right).$$

Hence, we have that

$$S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}\left(\frac{1}{2}, \frac{1}{2}\right) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) + (s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(u), s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}(v)), \text{ where } u = \left(\frac{2-r}{2}\right)\left(1 - r^{i-1}\right) \text{ and } v = \left(\frac{2-r}{2}\right)\left(1 - r^{j-1}\right). \text{ Therefore,}$$

$$\begin{aligned} E(X|X \in J_{\omega(\infty, \infty)}) &= S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) + \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}}(s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(u), s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}(v)) \\ &= S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) + (s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1-r}{2}\right), s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1-r}{2}\right)). \end{aligned}$$

Proofs of (ii) and (iii) are similar. \square

Note 1. For words $\beta, \gamma, \dots, \delta$ in \mathcal{I}^* , by $a(\beta, \gamma, \dots, \delta)$ we denote the conditional expectation of the random variable X given $J_{\beta} \cup J_{\gamma} \cup \dots \cup J_{\delta}$, i.e.,

$$a(\beta, \gamma, \dots, \delta) = E(X|X \in J_{\beta} \cup J_{\gamma} \cup \dots \cup J_{\delta}) = \frac{1}{P(J_{\beta} \cup \dots \cup J_{\delta})} \int_{J_{\beta} \cup \dots \cup J_{\delta}} (x_1, x_2) dP. \tag{1}$$

Then, for $\omega \in \mathcal{I}^*$,

$$\begin{cases} a(\omega) = S_{\omega}(E(X)) = S_{\omega}\left(\frac{1}{2}, \frac{1}{2}\right), \\ a(\omega(\emptyset, \infty)) = E(X|X \in J_{\omega(\emptyset, \infty)}), \\ a(\omega(\infty, \emptyset)) = E(X|X \in J_{\omega(\infty, \emptyset)}), \text{ and} \\ a(\omega(\infty, \infty)) = E(X|X \in J_{\omega(\infty, \infty)}). \end{cases} \tag{2}$$

Thus, by Lemma 4, if $\omega = (1, 1)$, then $a((1, 1)) = \left(\frac{r}{2}, \frac{r}{2}\right)$, $a((1, 1)(\infty, \emptyset)) = \left(1 - \frac{r}{2}, \frac{r}{2}\right)$, $a((1, 1)(\emptyset, \infty)) = \left(\frac{r}{2}, 1 - \frac{r}{2}\right)$, and $a((1, 1)(\infty, \infty)) = \left(1 - \frac{r}{2}, 1 - \frac{r}{2}\right)$. In addition,

$$\begin{cases} a((1, 1), (1, 1)(\infty, \emptyset)) = \left(\frac{1}{2}, \frac{r}{2}\right), \\ a((1, 1)(\emptyset, \infty), (1, 1)(\infty, \infty)) = \left(\frac{1}{2}, 1 - \frac{r}{2}\right), \\ a((1, 1), (1, 1)(\emptyset, \infty)) = \left(\frac{r}{2}, \frac{1}{2}\right), \\ a((1, 1)(\infty, \emptyset), (1, 1)(\infty, \infty)) = \left(1 - \frac{r}{2}, \frac{1}{2}\right). \end{cases} \tag{3}$$

Moreover, for $\omega \in \mathcal{I}^k, k \geq 1$, it is easy to see that

$$\begin{aligned} \int_{J_\omega} \|x - (a, b)\|^2 dP &= p_\omega \int \|(x_1, x_2) - (a, b)\|^2 dP \circ S_\omega^{-1} \\ &= p_\omega \left(s_\omega^{(1)2} V(X_1) + s_\omega^{(2)2} V(X_2) + \|S_\omega\left(\frac{1}{2}, \frac{1}{2}\right) - (a, b)\|^2 \right), \end{aligned} \tag{4}$$

where $s_\omega^{(k)2} := (s_\omega^{(k)})^2$ for $k = 1, 2$. The expressions (2) and (4) are useful to obtain the optimal sets and the corresponding quantization errors with respect to the probability distribution P .

For the rest of the article $r = \frac{1}{3}$ is assumed, which is the most important case due to its intimate connection with the standard Cantor system.

3. Optimal Sets of n -Means for $n = 2, 3$

In this section, we determine the optimal sets of two- and three-means, and their quantization errors.

Lemma 5. Let P be the affine measure on \mathbb{R}^2 and let $\omega \in \mathcal{I}^*$. Then,

$$\begin{aligned} \int_{J_{\omega(\infty, \infty)}} \|x - a(\omega(\infty, \infty))\|^2 dP &= \int_{J_{\omega(\emptyset, \infty)}} \|x - a(\omega(\emptyset, \infty))\|^2 dP \\ &= \int_{J_{\omega(\infty, \emptyset)}} \|x - a(\omega(\infty, \emptyset))\|^2 dP = \int_{J_\omega} \|x - a(\omega)\|^2 dP = p_\omega (s_\omega^{(1)2} + s_\omega^{(2)2}) \frac{1}{8}. \end{aligned}$$

Proof. Let us first prove $\int_{J_{\omega(\infty, \infty)}} \|x - a(\omega(\infty, \infty))\|^2 dP = p_\omega (s_\omega^{(1)2} + s_\omega^{(2)2}) \frac{1}{8}$. By Lemma 4, we have

$$\begin{aligned} \int_{J_{\omega(\infty, \infty)}} \|x - a(\omega(\infty, \infty))\|^2 dP &= \sum_{i,j=1}^\infty \int_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)} \|x - a(\omega(\infty, \infty))\|^2 dP \\ &= p_\omega \sum_{i,j=1}^\infty \frac{1}{2^{i+j}} \int \|S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}(x_1, x_2) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) \\ &\quad - (S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}, S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)})\|^2 dP. \end{aligned} \tag{5}$$

Note that $S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}(x_1, x_2) = \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1), S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}(x_2) \right)$ and $S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) = \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right), S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right) \right)$. Moreover, we have

$$\begin{aligned} &\left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) - s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \right)^2 \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)2} \left(S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}(x_1) - S_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)}\left(\frac{1}{2}\right) - s_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)} \right)^2 \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)2} \left(\left(S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}(x_1) - S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}\left(\frac{1}{2}\right) \right) + \left(S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}\left(\frac{1}{2}\right) - S_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)}\left(\frac{1}{2}\right) \right. \right. \\ &\quad \left. \left. - s_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)} \right) \right)^2. \end{aligned}$$

Now break the above expression by using the square formula and note the fact that

$$\int \left(S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1) - S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) \right)^2 dP_1 = s_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)2} V(X_1) = s_{(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)})}^{(1)2} \frac{1}{9^i} \frac{1}{8}, \text{ and}$$

$$\int \left(S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) - S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) \right) dP_1 = 0, \text{ and after some simplification we have}$$

$$\left(S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) - S_{(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) - s_{(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \right)^2 = s_{(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)})}^{(1)2} \frac{1}{4} \left(1 - \frac{5}{3^i}\right)^2.$$

Thus, it follows that

$$\int \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) - s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \right)^2 dP_1$$

$$= s_{\omega}^{(1)2} \left(\frac{1}{9^i} \frac{1}{8} + \frac{1}{4} \left(1 - \frac{5}{3^i}\right)^2 \right), \text{ and similarly}$$

$$\int \left(S_{\omega^-(\omega_{|\omega|}^{(2)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}(x_2) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right) - s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)} \right)^2 dP_2$$

$$= s_{\omega}^{(2)2} \left(\frac{1}{9^j} \frac{1}{8} + \frac{1}{4} \left(1 - \frac{5}{3^j}\right)^2 \right).$$

Therefore, (5) implies that

$$\int_{J_{\omega(\infty, \infty)}} \|x - a(\omega(\infty, \infty))\|^2 dP$$

$$= p_{\omega} \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}} \left(s_{\omega}^{(1)2} \left(\frac{1}{9^i} \frac{1}{8} + \frac{1}{4} \left(1 - \frac{5}{3^i}\right)^2 \right) + s_{\omega}^{(2)2} \left(\frac{1}{9^j} \frac{1}{8} + \frac{1}{4} \left(1 - \frac{5}{3^j}\right)^2 \right) \right) = p_{\omega} (s_{\omega}^{(1)2} + s_{\omega}^{(2)2}) \frac{1}{8}.$$

Other equalities of the statement are proved similarly. □

Lemma 6. Let P be the affine measure on \mathbb{R}^2 , and let $\{(a, p), (b, p)\}$ be a set of two points lying on the line $x_2 = p$ for which the distortion error is smallest. Then, $a = \frac{1}{6}$, $b = \frac{5}{6}$, $p = \frac{1}{2}$ and the distortion error is $\frac{5}{36}$.

Proof. Let $\beta = \{(a, p), (b, p)\}$. Because the points for which the distortion error is smallest are the centroids of their own Voronoi regions, by the properties of centroids, we have

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

which implies $pP(M((a, p)|\beta)) + pP(M((b, p)|\beta)) = \frac{1}{2}$, i.e, $p = \frac{1}{2}$. Thus, the boundary of the Voronoi regions is the line $x_1 = \frac{1}{2}$. Now, using the definition of conditional expectation,

$$\left(a, \frac{1}{2}\right) = E(X : X \in M((a, \frac{1}{2})|\beta)) = E(X : X \in \bigcup_{j=1}^{\infty} J_{(1,j)}) = \frac{1}{\sum_{j=1}^{\infty} p_{(1,j)}} \sum_{j=1}^{\infty} p_{(1,j)} S_{(1,j)}\left(\frac{1}{2}, \frac{1}{2}\right),$$

which implies $(a, \frac{1}{2}) = (\frac{1}{6}, \frac{1}{2})$ yielding $a = \frac{1}{6}$. Similarly, $b = \frac{5}{6}$. Then, the distortion error is

$$\int \min_{c \in \beta} \|x - c\|^2 dP = \int_{\bigcup_{j=1}^{\infty} J_{(1,j)}} \|x - (\frac{1}{6}, \frac{1}{2})\|^2 dP + \int_{\bigcup_{i=2, j=1}^{\infty} J_{(i,j)}} \|x - (\frac{5}{6}, \frac{1}{2})\|^2 dP = \frac{5}{72} + \frac{5}{72} = \frac{5}{36}.$$

This completes the proof the lemma. □

The following lemma provides us information on where to look for points of an optimal set of two-means.

Lemma 7. Let P be the affine measure on \mathbb{R}^2 . The points in an optimal set of two-means can not lie on an oblique line of the affine set.

Proof. In the affine set, among all the oblique lines that pass through the point $(\frac{1}{2}, \frac{1}{2})$, the line $x_2 = x_1$ has the maximum symmetry, i.e., with respect to the line $x_2 = x_1$ the affine set is geometrically symmetrical. Also, observe that, if two basic rectangles of similar geometrical shape lie in the opposite sides of the line $x_2 = x_1$, and are equidistant from the line $x_2 = x_1$, then they have the same probability (see Figure 1); hence, they are symmetrical with respect to the probability distribution P . Due to this, among all the pairs of two points which have the boundaries of the Voronoi regions oblique lines passing through the point $(\frac{1}{2}, \frac{1}{2})$, the two points which have the boundary of the Voronoi regions the line $x_2 = x_1$ will give the smallest distortion error. Again, we know the two 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 of the affine set with respect to the line $x_2 = x_1$ respectively. Then, from the definition of conditional expectation, we have

$$\begin{aligned} (a_1, b_1) = & 2 \left[\sum_{i=1, j=i+1}^{\infty} \frac{1}{2^{i+j}} S_{(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) + \sum_{k_1=1}^{\infty} \sum_{j=i+1}^{\infty} \frac{1}{2^{2k_1+i+j}} S_{(k_1, k_1)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) \right. \\ & + \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \sum_{j=i+1}^{\infty} \frac{1}{2^{2k_1+2k_2+i+j}} S_{(k_1, k_1)(k_2, k_2)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) \\ & \left. + \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \sum_{k_3=1}^{\infty} \sum_{j=i+1}^{\infty} \frac{1}{2^{2k_1+2k_2+2k_3+i+j}} S_{(k_1, k_1)(k_2, k_2)(k_3, k_3)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) + \dots \right] = \left(\frac{3}{10}, \frac{7}{10} \right), \end{aligned}$$

and

$$\begin{aligned} (a_2, b_2) = & 2 \left(\sum_{i=1}^{\infty} \sum_{j=1}^{i-1} \frac{1}{2^{i+j}} S_{(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) + \sum_{k_1=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{i-1} \frac{1}{2^{2k_1+i+j}} S_{(k_1, k_1)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) \right. \\ & + \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{i-1} \frac{1}{2^{2k_1+2k_2+i+j}} S_{(k_1, k_1)(k_2, k_2)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) \\ & \left. + \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \sum_{k_3=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{i-1} \frac{1}{2^{2k_1+2k_2+2k_3+i+j}} S_{(k_1, k_1)(k_2, k_2)(k_3, k_3)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) + \dots \right) = \left(\frac{7}{10}, \frac{3}{10} \right). \end{aligned}$$

Let $\beta = \{(\frac{3}{10}, \frac{7}{10}), (\frac{7}{10}, \frac{3}{10})\}$. Then, due to symmetry,

$$\int \min_{c \in \beta} \|x - c\|^2 dP = 2 \int_{M((\frac{3}{10}, \frac{7}{10}) | \beta)} \|x - (\frac{3}{10}, \frac{7}{10})\|^2 dP.$$

Write

$$\begin{aligned} A := & \left(\bigcup_{j=2}^4 J_{(1,1)(1,1)(1,1)(1,j)} \right) \cup \left(\bigcup_{j=2}^6 J_{(1,1)(1,1)(1,1)(1,j)} \right) \cup \left(\bigcup_{j=3}^5 J_{((1,1)(1,1)(1,1)(2,j))} \right) \cup \left(\bigcup_{j=2}^8 J_{(1,1)(1,1)(1,j)} \right) \\ & \cup \left(\bigcup_{j=3}^6 J_{(1,1)(1,1)(2,j)} \right) \cup J_{(1,1)(1,1)(3,4)} \cup \left(\bigcup_{j=2}^8 J_{(1,1)(1,j)} \right) \cup \left(\bigcup_{j=3}^7 J_{(1,1)(2,j)} \right) \cup \left(\bigcup_{j=4}^6 J_{(1,1)(3,j)} \right) \cup \left(\bigcup_{j=2}^{10} J_{(1,j)} \right) \\ & \cup \left(\bigcup_{j=3}^{10} J_{(2,j)} \right) \cup \left(\bigcup_{j=4}^{10} J_{(3,j)} \right) \cup \left(\bigcup_{j=5}^9 J_{(4,j)} \right) \cup \left(\bigcup_{j=6}^7 J_{(5,j)} \right). \end{aligned}$$

Because A is a proper subset of $M((\frac{3}{10}, \frac{7}{10})|\beta)$, we have $\int \min_{c \in \beta} \|x - c\|^2 dP > 2 \int_A \|x - (\frac{3}{10}, \frac{7}{10})\|^2 dP$. Now using (4), and then upon simplification, it follows that

$$\int \min_{c \in \beta} \|x - c\|^2 dP > 2 \int_A \|x - (\frac{3}{10}, \frac{7}{10})\|^2 dP = 0.13899,$$

which is larger than the distortion error $\frac{5}{36}$ obtained in Lemma 6. Hence, the points in an optimal set of two-means can not lie on an oblique line of the affine set. Thus, the assertion of the lemma follows. \square

Proposition 2. Let P be the affine measure on \mathbb{R}^2 . Then, 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 $\frac{5}{36}$.

Proof. By Lemma 7, it is known that the points in an optimal set of two-means cannot lie on an oblique line of the affine set. Thus, by Lemma 6, we see that $\{(\frac{1}{6}, \frac{1}{2}), (\frac{5}{6}, \frac{1}{2})\}$ forms an optimal set of two-means with quantization error $\frac{5}{36}$. Due to symmetry, $\{(\frac{1}{2}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ forms another optimal set of two-means (see Figure 2); thus, the assertion follows. \square

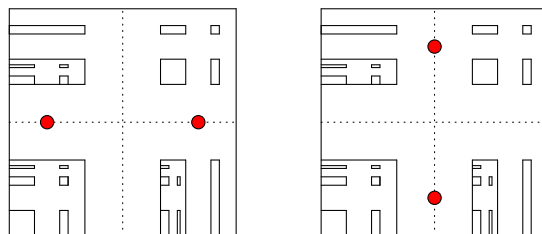


Figure 2. Optimal sets of two-means.

Proposition 3. Let P be the affine measure on \mathbb{R}^2 . Then, the set $\{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ forms an optimal set of three-means with quantization error $\frac{1}{12}$.

Proof. Let us first consider a three-point set β given by $\beta = \{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$. Then, by using Lemma 5 and Equation (4), we have

$$\begin{aligned} \int \min_{a \in \beta} \|x - a\|^2 dP &= \int_{J_{(1,1)}} \|x - (\frac{1}{6}, \frac{1}{6})\|^2 dP + \int_{J_{(1,1)(\infty, \emptyset)}} \|x - (\frac{5}{6}, \frac{1}{6})\|^2 dP \\ &+ \int_{J_{(1,1)(\emptyset, \infty)} \cup J_{(1,1)(\infty, \infty)}} \|x - (\frac{1}{2}, \frac{5}{6})\|^2 dP = \frac{1}{12}. \end{aligned}$$

Because V_3 is the quantization error for an optimal set of three-means, we have $\frac{1}{12} \geq V_3$. Let $\alpha = \{(a_i, b_i) : 1 \leq i \leq 3\}$ be an optimal set of three-means. Because the optimal points are the centroids of their own Voronoi regions, we have $\alpha \subset [0, 1] \times [0, 1]$. Let $A_1 = [0, \frac{1}{3}] \times [0, \frac{1}{3}]$, $A_2 = [\frac{2}{3}, 1] \times [0, \frac{1}{3}]$, $A_3 = [0, \frac{1}{3}] \times [\frac{2}{3}, 1]$, and $A_4 = [\frac{2}{3}, 1] \times [\frac{2}{3}, 1]$. Note that the centroids of A_1, A_2, A_3 and A_4 with respect to the probability distribution P are respectively $(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{6}, \frac{5}{6})$ and $(\frac{5}{6}, \frac{5}{6})$. Suppose that α does not contain any point from $\bigcup_{i=1}^4 A_i$. Then, we can assume that all the points of α are on the line $x_2 = \frac{1}{2}$, i.e., $\alpha = \{(a_i, \frac{1}{2}) : 1 \leq i \leq 3\}$ with $a_1 < a_2 < a_3$. If $a_1 > \frac{1}{3}$, quantization error can be strictly reduced by moving the point $(a_1, \frac{1}{2})$ to $(\frac{1}{3}, \frac{1}{2})$. So, we can assume that $a_1 \leq \frac{1}{3}$. Similarly, we can show that $a_3 \geq \frac{2}{3}$. Now, if $a_2 < \frac{1}{3}$, then $A_3 \cup A_4 \subset M((a_3, \frac{1}{2})|\alpha)$. Moreover, for

any $x = (x_1, x_2) \in J_{(1,1)(1,1)} \cup J_{(1,3)}$, we have $m(x) := \min_{c \in \alpha} \|(x_1, x_2) - c\|^2 \geq (\frac{7}{18})^2$ and so by (4) and Lemma 5, we obtain

$$\begin{aligned} \int m(x)^2 dP &= \int_{J_{(1,1)(1,1)} \cup J_{(1,3)}} m(x)^2 dP + \int_{J_{(1,1)(\infty, \emptyset)} \cup J_{(1,1)(\infty, \infty)}} m(x)^2 dP \\ &\geq \frac{1}{16} \left(\left(\frac{1}{81} + \frac{1}{81} \right) \frac{1}{8} + \left(\frac{7}{18} \right)^2 \right) + \frac{1}{16} \left(\left(\frac{1}{9} + \frac{1}{27^2} \right) \frac{1}{8} + \left(\frac{7}{18} \right)^2 \right) + \int_{J_{(1,1)(\infty, \emptyset)} \cup J_{(1,1)(\infty, \infty)}} \|x - (\frac{5}{6}, \frac{1}{2})\|^2 dP \\ &= \frac{1}{16} \left(\left(\frac{1}{81} + \frac{1}{81} \right) \frac{1}{8} + \left(\frac{7}{18} \right)^2 \right) + \frac{1}{16} \left(\left(\frac{1}{9} + \frac{1}{27^2} \right) \frac{1}{8} + \left(\frac{7}{18} \right)^2 \right) + \frac{5}{72} = \frac{1043}{11664} > V_3, \end{aligned}$$

which is a contradiction, and so $a_2 \geq \frac{1}{3}$ must be true. If $a_2 > \frac{2}{3}$, similarly we can show that a contradiction arises. So, $\frac{1}{3} < a_2 < \frac{2}{3}$. Next, suppose that $\frac{1}{2} \leq a_2 < \frac{2}{3}$. Then, we have $\frac{1}{2}(a_1 + a_2) \leq \frac{1}{3}$ which implies $a_1 \leq \frac{1}{6}$, for otherwise quantization error can be strictly reduced by moving a_2 to $(\frac{2}{3}, \frac{1}{2})$, contradicting the fact that α is an optimal set. Then, $\bigcup_{j=1}^{\infty} J_{(1,1)(1,j)} \cup \bigcup_{i=2, j=1}^{\infty} J_{(1,i)(1,j)} \subset M((a_1, \frac{1}{2}) | \alpha)$ and $E(X : X \in \bigcup_{j=1}^{\infty} J_{(1,1)(1,j)} \cup \bigcup_{i=2, j=1}^{\infty} J_{(1,i)(1,j)}) = (\frac{1}{18}, \frac{1}{2})$. So, for any $(x_1, x_2) \in \bigcup_{i=2, j=1}^{\infty} J_{(1,1)(i,j)} \cup \bigcup_{k=1, i=2, j=1}^{\infty} J_{(k,2)(i,j)}$, $\min_{c \in \alpha} \|(x_1, x_2) - c\|^2 \geq \|(x_1, x_2) - (\frac{1}{6}, \frac{1}{2})\|^2$. If $A = \bigcup_{j=1}^{\infty} J_{(1,1)(1,j)} \cup \bigcup_{i=2, j=1}^{\infty} J_{(1,i)(1,j)}$, $B = \bigcup_{i=2, j=1}^{\infty} J_{(1,1)(i,j)} \cup \bigcup_{k=1, i=2, j=1}^{\infty} J_{(k,2)(i,j)}$, $A' = \bigcup_{j=1}^{\infty} J_{(1,1)(1,j)}$ and $B' = \bigcup_{k=1, i=2, j=1}^{\infty} J_{(k,2)(i,j)}$, then

$$\begin{aligned} \int m(x)^2 dP &> \int_A \|(x_1, x_2) - (\frac{1}{18}, \frac{1}{2})\|^2 dP + \int_B \|(x_1, x_2) - (\frac{1}{6}, \frac{1}{2})\|^2 dP \\ &= 2 \int_{A'} \|x - (\frac{1}{18}, \frac{1}{2})\|^2 dP + \int_{\bigcup_{i=2, j=1}^{\infty} J_{(1,1)(i,j)}} \|x - (\frac{1}{6}, \frac{1}{2})\|^2 dP + \int_{B'} \|x - (\frac{1}{6}, \frac{1}{2})\|^2 dP \\ &= 2 \cdot \frac{41}{2592} + \frac{5}{288} + \frac{551}{14688} = \frac{953}{11016} > V_3, \end{aligned}$$

which is a contradiction. Similarly, if we assume $\frac{1}{3} \leq a_2 < \frac{1}{2}$, a contradiction will arise. Therefore, all the points in α can not lie on the line $x_2 = \frac{1}{2}$. Let (a_1, b_1) and (a_3, b_3) lie on the line $x_2 = \frac{1}{2}$, and (a_2, b_2) is above or below the horizontal line $x_2 = \frac{1}{2}$. If (a_2, b_2) is above the horizontal line, then the quantization error can be strictly reduced by moving (a_1, b_1) to A_1 and (a_3, b_3) to A_2 contradicting the fact that α is an optimal set. Similarly, if (a_2, b_2) is below the horizontal line, a contradiction will arise. All these contradictions arise due to our assumption that α does not contain any point from $\bigcup_{i=1}^4 A_i$. Hence, α contains at least one point from $\bigcup_{i=1}^4 A_i$. In order to complete the proof of the Proposition, first we will prove the following claim:

Claim 1. $card(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 2$.

For the sake of contradiction, assume that $card(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 1$. Then, without any loss of generality we assume that $(a_1, b_1) \in A_1$ and $(a_i, b_i) \notin A_2 \cup A_3 \cup A_4$ for $i = 2, 3$. Due to symmetry of the affine set with respect to the diagonal $x_2 = x_1$, we can assume that $(a_1, b_1) \in A_1$ lies on the diagonal $x_2 = x_1$; (a_2, b_2) and (a_3, b_3) are equidistant from the diagonal $x_2 = x_1$ and are in opposite sides of the diagonal $x_2 = x_1$. Now, consider the following cases:

Case 1. Assume that both (a_2, b_2) and (a_3, b_3) are below the diagonal $x_2 = 1 - x_1$, but not in $A_1 \cup A_2 \cup A_3$. Let (a_2, b_2) be above the diagonal $x_2 = x_1$ and (a_3, b_3) be below the diagonal $x_2 = x_1$. In that case, the quantization error can be strictly reduced by moving (a_2, b_2) to A_3 and (a_3, b_3) to A_2 which contradicts the optimality of α .

Case 2. Assume that both (a_2, b_2) and (a_3, b_3) are above the diagonal $x_2 = 1 - x_1$. Let (a_2, b_2) lie above the diagonal $x_2 = x_1$ and (a_3, b_3) lie below the diagonal $x_2 = x_1$. Then, due to symmetry we can assume that $(a_1, b_1) = (\frac{1}{6}, \frac{1}{6})$ which is the centroid of A_1 , $(a_2, b_2) = (\frac{1}{2}, \frac{5}{6})$ which is the midpoint of the line segment joining the centroids of A_3 and A_4 , $(a_3, b_3) = (\frac{5}{6}, \frac{1}{2})$ which is the midpoint of the line segment joining the centroids of A_2 and A_4 . Then,

$$\begin{aligned} \int m(x)^2 dP &= \int_{J_{(1,1)}} m(x)^2 dP + \int_{J_{(1,1)(\emptyset, \infty)}} m(x)^2 dP + \int_{J_{(1,1)(\infty, \emptyset)}} m(x)^2 dP + \int_{J_{(1,1)(\infty, \infty)}} m(x)^2 dP \\ &\geq \frac{1}{144} + \int_{J_{(1,1)(\emptyset, \infty)}} \|x - (\frac{1}{2}, \frac{5}{6})\|^2 dP + \int_{J_{(1,1)(\infty, \emptyset)}} \|x - (\frac{5}{6}, \frac{1}{2})\|^2 dP + \int_{\bigcup_{j=i+1}^{\infty} J_{(i,j)}} \|x - (\frac{1}{2}, \frac{5}{6})\|^2 dP \\ &= \frac{1}{144} + \frac{5}{144} + \frac{5}{144} + \frac{1381}{166320} = \frac{7043}{83160} > V_3, \end{aligned}$$

which is a contradiction. Thus, $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 1$ cannot hold.

Next, for the sake of contradiction, assume that $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 3$. Then, without any loss of generality we assume that $(a_1, b_1) \in A_3$, $(a_2, b_2) \in A_2$ and $(a_3, b_3) \in A_4$. Let A_{11} and A_{12} be the regions of A_1 which are respectively above and below the diagonal of A_1 passing through $(0, 0)$. Due to symmetry, we must have $A_3 \cup A_{11} \subset M((a_1, b_1)|\alpha)$ and $A_2 \cup A_{12} \subset M((a_2, b_2)|\alpha)$. Notice that $A_3 \cup A_{11} \subset M((a_1, b_1)|\alpha)$ implies

$$A_3 \cup \bigcup_{i=1, j=i+1} J_{(1,1)(i,j)} \cup \bigcup_{\substack{k=1, i=1 \\ j=i+1}} J_{(1,1)(k,k)(i,j)} \subset M((a_1, b_1)|\alpha),$$

and by using (1), we have

$$E(X : X \in A_3 \cup \bigcup_{i=1, j=i+1} J_{(1,1)(i,j)} \cup \bigcup_{\substack{k=1, i=1 \\ j=i+1}} J_{(1,1)(k,k)(i,j)}) = (\frac{1385}{9438}, \frac{6173}{9438}),$$

which shows that the point (a_1, b_1) falls below the line $x_2 = \frac{2}{3}$, which is a contradiction, as we assumed that $(a_1, b_1) \in A_3$. This contradiction arises due to our assumption that $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 3$. Hence, we conclude that $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 2$, which proves the claim.

By the claim, we assume that $(a_1, b_1) \in A_1$ and $(a_3, b_3) \in A_2$. Notice that A_1, A_2, A_3, A_4 are geometrically symmetric as well as their corresponding centroids are symmetrically distributed over the square $[0, 1] \times [0, 1]$. Without any loss of generality, we can assume that the optimal point (a_1, b_1) is the centroid of A_1 , i.e., $(a_1, b_1) = (\frac{1}{6}, \frac{1}{6})$. Then, due to symmetry with respect to the line $x_1 = \frac{1}{2}$, it follows that $(a_3, b_3) = \text{centroid of } A_2 = (\frac{5}{6}, \frac{1}{6})$, and (a_2, b_2) lies on $x_1 = \frac{1}{2}$ but above the line $x_2 = \frac{1}{2}$. Now, notice that

$$\min_{(a_3, b_3) \in [\frac{1}{3}, \frac{2}{3}] \times [\frac{2}{3}, 1]} \{ \|(\frac{1}{6}, \frac{5}{6}) - (a_3, b_3)\|^2 + \|(\frac{5}{6}, \frac{5}{6}) - (a_3, b_3)\|^2 \} = \frac{2}{9},$$

which occurs when $(a_3, b_3) = \text{center of } [\frac{1}{3}, \frac{2}{3}] \times [\frac{2}{3}, 1] = (\frac{1}{2}, \frac{5}{6})$. Moreover, the three points $(\frac{1}{6}, \frac{1}{6})$, $(\frac{5}{6}, \frac{1}{6})$ and $(\frac{1}{2}, \frac{5}{6})$ are the centroids of their own Voronoi regions. Thus, $\{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ forms an optimal set of three-means with quantization error $V_3 = \frac{1}{12}$. Hence, the proposition follows. \square

Remark 3. Due to symmetry, in addition to the optimal set given in Proposition 3, there are three more optimal sets of three-means with quantization error $V_3 = \frac{1}{12}$ (see Figure 3).

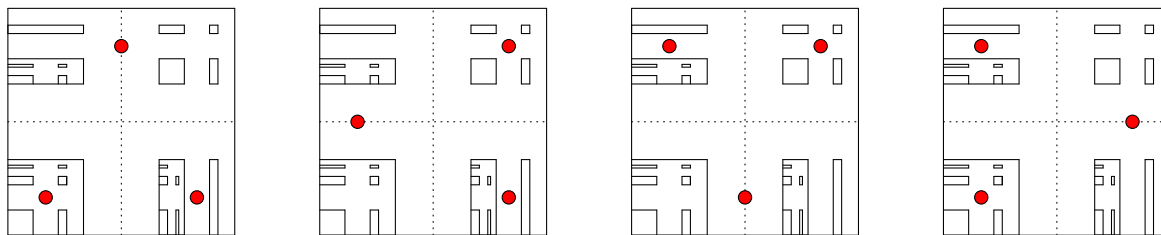


Figure 3. Optimal sets of three-means.

4. Affine Measures

In this section, we show that the affine measure P under consideration is the direct product of the Cantor distribution P_c .

For the rest of the article, by a word σ of length k over the alphabet $\{1, 2\}$, it is meant $\sigma := \sigma_1\sigma_2 \cdots \sigma_k \in \{1, 2\}^k, k \geq 1$. By a word of length zero it is meant the empty word \emptyset . $\{1, 2\}^*$ represents the set of all words over the alphabet $\{1, 2\}$ including the empty word \emptyset . Length of a word $\sigma \in \{1, 2\}^*$ is denoted by $|\sigma|$. If $\sigma = \sigma_1\sigma_2 \cdots \sigma_k$, we write $U_\sigma := U_{\sigma_1} \circ U_{\sigma_2} \circ \cdots \circ U_{\sigma_k}$. U_\emptyset represents the identity mapping on \mathbb{R} . By u_σ we represent the similarity ratio of U_σ . If X_c is the random variable with distribution P_c , then $E(X_c) = \frac{1}{2}$ and $V(X_c) = \frac{1}{8}$ [10]. For $\sigma \in \{1, 2\}^*$, write $A(\sigma) := U_\sigma(\frac{1}{2})$. Notice that for $\sigma \in \{1, 2\}^*$, we have $\frac{1}{2}(A(\sigma 1) + A(\sigma 2)) = A(\sigma)$, $u_\sigma = \frac{1}{3^{|\sigma|}}$, the contractive factor of U_σ , and for the empty word \emptyset , $A(\emptyset) = \frac{1}{2}$. For $\sigma \in \{1, 2\}^*$ define $A_\sigma := U_\sigma[0, 1]$. For any positive integer n , by 2^{*n} it is meant the concatenation of the symbol 2 with itself n -times successively, i.e., $2^{*n} = 222 \cdots (n \text{ times})$, with the convention that 2^{*0} is the empty word. For any positive integer k , by $\{1, 2\}^{k*2}$ it is meant the direct product of the set $\{1, 2\}^k$ with itself. By $\{1, 2\}^{0*2}$ it is meant the set $\{(\emptyset, \emptyset)\}$. Also, recall the notations defined in Section 2. Let us now introduce the map $F : \mathbb{N}^* \cup \{(\sigma, \infty) : \sigma \in \mathbb{N}^*\} \rightarrow \{1, 2\}^*$ such that

$$F(x) = \begin{cases} f(\sigma_1)f(\sigma_2) \cdots f(\sigma_{|\sigma|}) & \text{if } x = \sigma = \sigma_1\sigma_2 \cdots \sigma_{|\sigma|}, \\ f(\sigma_1)f(\sigma_2) \cdots f(\sigma_{|\sigma|}, \infty) & \text{if } x = (\sigma_1\sigma_2 \cdots \sigma_{|\sigma|}, \infty), \\ \emptyset & \text{if } x = \emptyset, \end{cases} \tag{6}$$

where $f : \mathbb{N} \cup \{(n, \infty) : n \in \mathbb{N}\} \rightarrow \{1, 2\}^* \setminus \{\emptyset\}$ is such that

$$f(x) = \begin{cases} 2^{*(n-1)}1 & \text{if } x = n \text{ for some } n \in \mathbb{N}, \\ 2^{*n} & \text{if } x = (n, \infty) \text{ for some } n \in \mathbb{N}. \end{cases}$$

The function f is one-to-one and onto, and consequently, F is also one-to-one and onto. For any $\sigma \in \mathbb{N}^*$, write $AF(\sigma) := A(F(\sigma))$ and $AF(\sigma, \infty) := A(F(\sigma, \infty))$.

The map F is instrumental in converting the infinitely generated affine measure P to a finitely generated affine measure $P_c \times P_c$. Furthermore, to improve the clarity of the arguments, we will write T_i for $S_{(i,j)}^{(1)}$, and T_j for $S_{(i,j)}^{(2)}$, where T_k for all $k \geq 1$ form an infinite collection of similarity mappings on \mathbb{R} such that $T_k(x) = \frac{1}{3^k}x + 1 - \frac{1}{3^{k-1}}$ for all $x \in \mathbb{R}$. Thus, if $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_n, j_n)$, then $S_\omega^{(1)} = T_{i_1} \circ \cdots \circ T_{i_n} = T_{i_1 i_2 \cdots i_n}$ and $S_\omega^{(2)} = T_{j_1} \circ \cdots \circ T_{j_n} = T_{j_1 j_2 \cdots j_n}$ for all $n \geq 1$. Again, T_\emptyset is the identity mapping on \mathbb{R} .

Lemma 8. Let T_k for $k \geq 1$ be the infinite collection of similitudes defined above, and U_1 and U_2 be the similitudes generating the Cantor set. Then, for any $\sigma \in \mathbb{N}^*$ and $x \in \mathbb{R}$, we have $T_\sigma(x) = U_{F(\sigma)}(x)$.

Proof. If $\sigma = 1$, then $T_1(x) = \frac{1}{3}x = U_1(x) = U_{F(1)}(x)$ for any $x \in \mathbb{R}$. Assume that the lemma is true if $\sigma = k$ for some positive integer k , i.e., $T_k(x) = U_{F(k)}(x)$. Then,

$$\begin{aligned} U_{F(k+1)}(x) &= U_{2^{*k}1}(x) = U_{2^{*(k-1)}21}(x) = U_{2^{*(k-1)}}U_{21}(x) = U_{2^{*(k-1)}}\left(\frac{1}{9}x + \frac{2}{3}\right) \\ &= U_{2^{*(k-1)}1}\left(3\left(\frac{1}{9}x + \frac{2}{3}\right)\right) = U_{F(k)}\left(\frac{1}{3}x + 2\right) = T_k\left(\frac{1}{3}x + 2\right) = \frac{1}{3^k}\left(\frac{1}{3}x + 2\right) + 1 - \frac{1}{3^{k-1}} \\ &= \frac{1}{3^{k+1}}x + 1 - \frac{1}{3^k} = T_{k+1}(x). \end{aligned}$$

Thus, by the Principle of Mathematical Induction, $T_k(x) = U_{F(k)}(x)$ for all $k \in \mathbb{N}$. Again, for any $\tau, \delta \in \mathbb{N}^*$, by (6), it follows that $F(\sigma\delta) = F(\sigma)F(\delta)$. Hence, for any $\sigma = \sigma_1\sigma_2 \cdots \sigma_n \in \mathbb{N}^*$, $n \geq 1$, we have

$$T_\sigma(x) = T_{\sigma_1} \circ T_{\sigma_2} \circ \cdots \circ T_{\sigma_n}(x) = U_{F(\sigma_1)} \circ U_{F(\sigma_2)} \circ \cdots \circ U_{F(\sigma_n)}(x) = U_{F(\sigma)}(x),$$

which completes the proof. \square

Lemma 9. Let $\omega \in \mathcal{I}^*$, and F be the function as defined in (6). Then for $r = 1, 2$, we have $AF(\omega^{(r)}) = S_\omega^{(r)}\left(\frac{1}{2}\right)$, and $AF(\omega^{(r)}, \infty) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}\right) + s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(r)}$.

Proof. By Lemma 8, we have

$$AF(\omega^{(1)}) = U_{F(\omega^{(1)})}\left(\frac{1}{2}\right) = T_{\omega^{(1)}}\left(\frac{1}{2}\right) = S_\omega^{(1)}\left(\frac{1}{2}\right), \text{ and similarly } AF(\omega^{(2)}) = S_\omega^{(2)}\left(\frac{1}{2}\right).$$

Without any loss of generality, we can assume $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)$ for $k \geq 1$. Then,

$$\begin{aligned} AF(\omega^{(1)}, \infty) &= U_{F(i_1i_2 \cdots i_k, \infty)}\left(\frac{1}{2}\right) = U_{F(i_1i_2 \cdots i_{k-1})} \circ U_{F(i_k, \infty)}\left(\frac{1}{2}\right) = U_{F(i_1i_2 \cdots i_{k-1})} \circ U_{2^{*i_k}}\left(\frac{1}{2}\right) \\ &= U_{F(i_1i_2 \cdots i_{k-1})} \circ U_{2^{*i_k}1}\left(U_1^{-1}\left(\frac{1}{2}\right)\right) = U_{F(i_1i_2 \cdots i_{k-1})} \circ U_{F(i_k+1)}\left(\frac{3}{2}\right) = U_{F(i_1i_2 \cdots i_{k-1}(i_k+1))}\left(\frac{3}{2}\right) \\ &= T_{i_1i_2 \cdots i_{k-1}(i_k+1)}\left(\frac{3}{2}\right) = S_{\omega^-(i_k+1, j_k+1)}^{(1)}\left(\frac{3}{2}\right). \end{aligned}$$

Because, $S_{(i_k+1, j_k+1)}^{(1)}\left(\frac{3}{2}\right) - S_{(i_k+1, j_k+1)}^{(1)}\left(\frac{1}{2}\right) = \frac{1}{3^{i_k+1}}\frac{3}{2} + 1 - \frac{1}{3^{i_k}} - \frac{1}{3^{i_k+1}}\frac{1}{2} - 1 + \frac{1}{3^{i_k}} = \frac{1}{3^{i_k+1}}$, we have

$$\begin{aligned} S_{\omega^-(i_k+1, j_k+1)}^{(1)}\left(\frac{3}{2}\right) - S_{\omega^-(i_k+1, j_k+1)}^{(1)}\left(\frac{1}{2}\right) &= s_{\omega^-(i_k+1, j_k+1)}^{(1)}\left(S_{(i_k+1, j_k+1)}^{(1)}\left(\frac{3}{2}\right) - S_{(i_k+1, j_k+1)}^{(1)}\left(\frac{1}{2}\right)\right) = s_{\omega^-(i_k+1, j_k+1)}^{(1)}\frac{1}{3^{i_k+1}} \\ &= s_{\omega^-(i_k+1, j_k+1)}^{(1)} = s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}, \text{ which yields} \end{aligned}$$

$$\begin{aligned} AF(\omega^{(1)}, \infty) &= S_{\omega^-(i_k+1, j_k+1)}^{(1)}\left(\frac{3}{2}\right) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) + s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}. \text{ Similarly,} \\ AF(\omega^{(2)}, \infty) &= S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right) + s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}. \quad \square \end{aligned}$$

Remark 4. By Lemmas 4 and 9, for any $\omega \in \mathcal{I}^*$, we have

$$\begin{aligned} a(\omega) &= (AF(\omega^{(1)}), AF(\omega^{(2)})), \quad a(\omega(\infty, \infty)) = (AF(\omega^{(1)}, \infty), AF(\omega^{(2)}, \infty)), \\ a(\omega(\infty, \emptyset)) &= (AF(\omega^{(1)}, \infty), AF(\omega^{(2)})), \quad \text{and } a(\omega(\emptyset, \infty)) = (AF(\omega^{(1)}), AF(\omega^{(2)}, \infty)). \end{aligned}$$

The following example illustrates the outcome of the lemma above.

Example 1. $a((1, 1)) = (AF(1), AF(1)) = (A(1), A(1)) = (\frac{1}{6}, \frac{1}{6}),$
 $a((1, 1)(\infty, \emptyset)) = (AF(1, \infty), AF(1)) = (A(2), A(1)) = (\frac{5}{6}, \frac{1}{6}),$
 $a((1, 1)(\emptyset, \infty)) = (AF(1), AF(1, \infty)) = (A(1), A(2)) = (\frac{1}{6}, \frac{5}{6}),$
 $a((1, 1)(\infty, \infty)) = (AF(1, \infty), AF(1, \infty)) = (A(2), A(2)) = (\frac{5}{6}, \frac{5}{6}),$
 $a((1, 1)(1, 1)) = (AF(11), AF(11)) = (A(11), A(11)) = (\frac{1}{18}, \frac{1}{18}),$
 $a((1, 1)(1, 1)(\infty, \emptyset)) = (AF(11, \infty), AF(11)) = (A(12), A(11)) = (\frac{5}{18}, \frac{1}{18}),$
 $a((1, 1)(1, 1)(\emptyset, \infty)) = (AF(11), AF(11, \infty)) = (A(11), A(12)) = (\frac{1}{18}, \frac{5}{18}),$ and
 $a((1, 1)(1, 1)(\infty, \infty)) = (AF(11, \infty), AF(11, \infty)) = (A(12), A(12)) = (\frac{5}{18}, \frac{5}{18}),$ etc.

Lemma 10. Let $\mu = \sum_{k=1}^{\infty} \frac{1}{2^k} \mu \circ T_k^{-1}$. Then, for any $\sigma \in \mathbb{N}^*$, we have $\mu(T_\sigma[0, 1]) = P_c(A_{F(\sigma)})$, where $P_c := \frac{1}{2}P_c \circ U_1^{-1} + \frac{1}{2}P_c \circ U_2^{-1}$.

Proof. Without any loss of generality, let $\sigma = i_1 i_2 \dots i_k$ for any $k \geq 1$. See that $F(\sigma) = F(i_1)F(i_2) \dots F(i_k)$, and thus $|F(\sigma)| = |F(i_1)| + |F(i_2)| + \dots + |F(i_k)| = i_1 + i_2 + \dots + i_k$. Consequently,

$$\mu(T_\sigma[0, 1]) = \frac{1}{2^{i_1+i_2+\dots+i_k}} = \frac{1}{2^{|F(\sigma)|}} = P_c(A_{F(\sigma)}),$$

which proves the lemma. \square

Proposition 4. Let P be the affine measure. Then, $P = P_c \times P_c$, where P_c is the Cantor distribution.

Proof. Borel σ -algebra on the affine set is generated by all sets of the form $J_{(\delta, \tau)}$ for $(\delta, \tau) \in \mathcal{I}^*$, where $J_{(\delta, \tau)} = S_{(\delta, \tau)}([0, 1] \times [0, 1])$. Notice that

$$J_{(\delta, \tau)} = T_\delta[0, 1] \times T_\tau[0, 1] = U_{F(\delta)}[0, 1] \times U_{F(\tau)}[0, 1] = A_{F(\delta)} \times A_{F(\tau)}.$$

Again, the sets of the form A_α , where $\alpha \in \{1, 2\}^*$, generate the Borel σ -algebra on the Cantor set C . Thus, we see that the Borel σ -algebra of the affine set is the same as the product of the Borel σ -algebras on the Cantor set. Moreover, for any $(\delta, \tau) \in \mathcal{I}^*$, by Remark 1 and Lemma 10, we have

$$P(J_{(\delta, \tau)}) = \mu(T_\delta[0, 1])\mu(T_\tau[0, 1]) = P_c(A_{F(\delta)})P_c(A_{F(\tau)}) = (P_c \times P_c)(A_{F(\delta)} \times A_{F(\tau)}).$$

Hence, the proposition follows. \square

Remark 5. By Proposition 4, it follows that the optimal sets of n -means for P are the same as the optimal sets n -means for the product measure $P_c \times P_c$ on the affine set. Moreover, for $k \geq 1$ we can write

$$P = P_c \times P_c = \sum_{(\sigma, \tau) \in \{1, 2\}^{k*2}} \frac{1}{4^k} (P_c \times P_c) \circ (U_\sigma, U_\tau)^{-1},$$

where for $(x_1, x_2) \in \mathbb{R}^2$, $(U_\sigma, U_\tau)^{-1}(x_1, x_2) = (U_\sigma^{-1}(x_1), U_\tau^{-1}(x_2))$.

5. Optimal Sets of n -Means for all $n \geq 4$

In this section, we will derive closed formulas to determine the optimal sets of n -means and the n th quantization error for all $n \geq 4$. For $(\sigma, \tau) \in \{1, 2\}^{k*2}$, write $A_{(\sigma, \tau)} := A_\sigma \times A_\tau$ and $U_{(\sigma, \tau)} := (U_\sigma, U_\tau)$.

Lemma 11. Let α be an optimal set of n -means with $n \geq 4$. Then, $\alpha \cap A_{(i, j)} \neq \emptyset$ for all $1 \leq i, j \leq 2$.

Proof. Let α be an optimal set of n -means for $n \geq 4$. As the optimal points are the centroids of their own Voronoi regions we have $\alpha \subset A_\emptyset \times A_\emptyset := [0, 1] \times [0, 1]$.

Consider the four-point set β given by $\beta = \{(A(i), A(j)) : 1 \leq i, j \leq 2\}$. Then,

$$\int \min_{c \in \beta} \|x - c\|^2 dP = \sum_{i,j=1}^2 \int_{A(i,j)} \|x - (A(i), A(j))\|^2 d(P_c \times P_c) = \sum_{i,j=1}^2 \frac{1}{4} \left(\frac{1}{9} + \frac{1}{9}\right) \frac{1}{8} = \frac{1}{36}.$$

Because V_4 is the quantization error of four-means, we have $\frac{1}{36} \geq V_4 \geq V_n$.

Assume that α does not contain any point from $\bigcup_{i,j=1}^2 A(i,j)$. We know that

$$\sum_{(a,b) \in \alpha} (a,b)P(M(a,b)|\alpha) = \left(\frac{1}{2}, \frac{1}{2}\right). \tag{7}$$

If all the points of α are below the line $x_2 = \frac{1}{2}$, i.e., if $b < \frac{1}{2}$ then by (7), we see that $\frac{1}{2} = \sum_{(a,b) \in \alpha} bP(M(a,b)|\alpha) < \sum_{(a,b) \in \alpha} \frac{1}{2}P(M(a,b)|\alpha) = \frac{1}{2}$, which is a contradiction. Similarly, it follows that if all the points of α are above the line $x_2 = \frac{1}{2}$, or left of the line $x_1 = \frac{1}{2}$, or right of the line $x_1 = \frac{1}{2}$, a contradiction will arise.

Next, suppose that all the points of α are on the line $x_2 = \frac{1}{2}$. We will consider two cases: $n = 4$ and $n > 4$. When $n = 4$, let $\alpha = \{(a_i, \frac{1}{2}) : 1 \leq i \leq 4\}$ with $a_i < a_j$ for $i < j$. Due to symmetry, we can assume that the boundary of the Voronoi regions of the points $(a_1, \frac{1}{2}), (a_2, \frac{1}{2}), (a_3, \frac{1}{2})$, and $(a_4, \frac{1}{2})$ are respectively $x_1 = \frac{1}{6}, x_1 = \frac{1}{2}$, and $x_1 = \frac{5}{6}$ yielding $\alpha = \{(\frac{1}{18}, \frac{1}{2}), (\frac{5}{18}, \frac{1}{2}), (\frac{13}{18}, \frac{1}{2}), (\frac{17}{18}, \frac{1}{2})\}$, and then writing $B := A_{(11,11)} \cup A_{(11,12)} \cup A_{(11,21)} \cup A_{(11,22)}$, by symmetry we have

$$\begin{aligned} \int \min_{c \in \alpha} \|x - c\|^2 dP &= 4 \int_B \|x - (\frac{1}{18}, \frac{1}{2})\|^2 d(P_c \times P_c) \\ &= 8 \int_{A_{(11,11)}} \|x - (\frac{1}{18}, \frac{1}{2})\|^2 d(P_c \times P_c) + 8 \int_{A_{(11,12)}} \|x - (\frac{1}{18}, \frac{1}{2})\|^2 d(P_c \times P_c) \\ &= 8\left(\frac{65}{5184} + \frac{17}{5184}\right) = \frac{41}{324} > V_4, \end{aligned}$$

which is a contradiction. We consider the case $n > 4$. Because for any $(x_1, x_2) \in \bigcup_{i,j=1}^2 A_{ij}$, $\min_{c \in \alpha} \|(x_1, x_2) - c\|^2 \geq \frac{1}{36}$, we have

$$\int \min_{c \in \alpha} \|x - c\|^2 dP = \sum_{i,j=1}^2 \int_{A(i,j)} \min_{c \in \alpha} \|x - c\|^2 d(P_c \times P_c) \geq \sum_{i,j=1}^2 \int_{A(i,j)} \frac{1}{36} d(P_c \times P_c) = \frac{1}{36},$$

which implies $\frac{1}{36} \geq V_4 > V_n$, a contradiction. Thus, we see that all the points of α can not lie on $x_2 = \frac{1}{2}$. Similarly, all the points of α can not lie on $x_1 = \frac{1}{2}$.

Notice that the lines $x_1 = \frac{1}{2}$ and $x_2 = \frac{1}{2}$ partition the square $[0, 1] \times [0, 1]$ into four quadrants with center $(\frac{1}{2}, \frac{1}{2})$. If $n = 4k$ for some positive integer k , due to symmetry, we can assume that each quadrant contains k -points from the set α . But then, any of the k points in the quadrant containing a basic rectangle $A(i,j)$ can be moved to $A(i,j)$ which strictly reduce the quantization error, and it gives a contradiction as we assumed that the set α is an optimal set of n -means and α does not contain any point from $A(i,j)$ for $1 \leq i, j \leq 2$.

If $n = 4k + 1, 4k + 2$, or $n = 4k + 3$, then, again due to symmetry, each quadrant gets at least k points. Then, as in the case $n = 4k$, here also, one can strictly reduce the quantization error by moving a point in the quadrant containing a basic rectangle $A(i,j)$ to $A(i,j)$ for $1 \leq i, j \leq 2$, which is a contradiction.

Thus, we have proved that $\alpha \cap A(i,j) \neq \emptyset$ for all $1 \leq i, j \leq 2$. \square

Lemma 12. Let α be an optimal set of n -means with $n \geq 4$. Then, $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

Proof. By Lemma 11, we know that $\alpha \cap A_{(i,j)} \neq \emptyset$ for all $1 \leq i, j \leq 2$. Now, we will prove the statement by considering four distinct cases:

Case 1: $n = 4k$ for some integer $k \geq 1$.

In this case, due to symmetry, we can assume that α contains k points from each of $A_{(i,j)}$, otherwise, quantization error can be reduced by redistributing the points of α equally among $A_{(i,j)}$ for $1 \leq i, j \leq 2$, and so $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

Case 2: $n = 4k + 1$ for some integer $k \geq 1$.

In this case, again due to symmetry, we can assume that α contains k points from each of $A_{(i,j)}$, and if possible, one point, say (a, b) , from $A_{(\emptyset, \emptyset)} \setminus \bigcup_{i,j=1}^2 A_{(i,j)}$. By symmetry, one can assume that (a, b) is the midpoint of the line segment joining any two centroids of the basic rectangles $A_{(i,j)}$ for $1 \leq i, j \leq 2$. Let us first take $(a, b) = (\frac{1}{2}, \frac{1}{2})$ which is the center of the affine set. For simplicity, we first assume $k = 1$, i.e., $n = 5$. Then, α contains only one point from each of $A_{(i,j)}$. Let (a_1, b_1) be the point that α takes from $A_{(1,1)}$. As $(\frac{1}{2}, \frac{1}{2})$ lies on the diagonal $x_2 = x_1$, due to symmetry we can also assume that (a_1, b_1) lies on the diagonal $x_2 = x_1$. By Proposition 1, we have $P(M((\frac{1}{2}, \frac{1}{2})|\alpha)) > 0$. This yields that $\frac{1}{2}((a_1, b_1) + (\frac{1}{2}, \frac{1}{2})) < (\frac{1}{3}, \frac{1}{3})$ which implies $a_1 < \frac{1}{6}$ and $b_1 < \frac{1}{6}$. Then, we see that

$$\frac{1}{36} = V_4 \approx V_5 = 4 \int_{A_{(1,1)}} \min_{c \in \{(a_1, b_1), (\frac{1}{2}, \frac{1}{2})\}} \|x - c\|^2 dP > \int_{c \in \beta} \min \|x - c\|^2 dP = \frac{2}{81} \geq V_5,$$

where $\beta = \{(\frac{1}{18}, \frac{1}{18}), (\frac{1}{18}, \frac{5}{18}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{6}, \frac{5}{6}), (\frac{5}{6}, \frac{5}{6})\}$, which is a contradiction. Similarly, if we take (a, b) as the midpoint of a line segments joining the centroids of any two adjacent basic rectangles $A_{(i,j)}$ for $1 \leq i, j \leq 2$, contradiction arises. Proceeding in the similar way, by taking $k = 2, 3, \dots$, we see that contradiction arises at each value k takes. Therefore, $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

Case 3: $n = 4k + 2$ for some integer $k \geq 1$.

In this case, due to symmetry, we can assume that α contains k points from each of $A_{(i,j)}$, and if possible, two points, say (a_1, b_1) and (a_2, b_2) , from $A_{(\emptyset, \emptyset)} \setminus \bigcup_{i,j=1}^2 A_{(i,j)}$. Then, by symmetry, we can assume that (a_1, b_1) lies on the midpoint of the line segment joining the centroids of $A_{(1,1)}, A_{(2,1)}$; and (a_2, b_2) lies on the midpoint of the line segment joining the centroids of $A_{(1,2)}$ and $A_{(2,2)}$. As in Case 2, this leads to a contradiction. Thus, $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

Case 4: $n = 4k + 3$ for some integer $k \geq 1$. Due to symmetry, in this case, we can assume that each of $A_{(1,1)}$ and $A_{(2,1)}$ gets $k + 1$ points; each of $A_{(1,2)}$ and $A_{(2,2)}$ gets k points. The remaining one point lies on the midpoint of the line segment joining the centroids of $A_{(1,2)}$ and $A_{(2,2)}$. But, in that case, proceeding as in Case 2, we can show that a contradiction arises. Thus, $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

We have shown that in all possible cases $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$; hence, the lemma follows. \square

Corollary 1. The set $\{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{6}, \frac{5}{6}), (\frac{5}{6}, \frac{5}{6})\}$ is a unique optimal set of four-means of the affine measure P with quantization error $V_4 = \frac{1}{36}$ (see Figure 4).

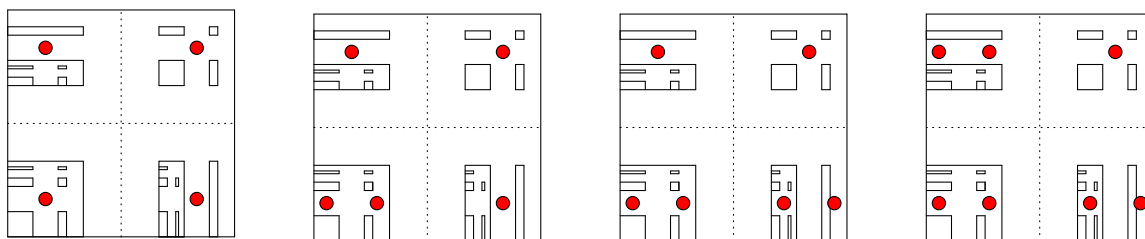


Figure 4. Optimal sets of n -means for $4 \leq n \leq 7$. Optimal set of 4-means is unique; on the other hand, optimal sets of n -means for $n = 5, 6, 7$ are not unique.

Remark 6. Let α be an optimal set of n -means, and $n_{ij} = \text{card}(\beta_{ij})$ where $\beta_{ij} = \alpha \cap A_{(i,j)}$ for $1 \leq i, j \leq 2$. Then, $0 \leq |n_{ij} - n_{pq}| \leq 1$ for $1 \leq i, j, p, q \leq 2$.

Lemma 13. Let $n \geq 4$ and α be an optimal set of n -means for the product measure $P_c \times P_c$. For $1 \leq i, j \leq 2$, set $\beta_{ij} := \alpha \cap A_{(i,j)}$, and let $n_{ij} = \text{card}(\beta_{ij})$. Then, $U_{(i,j)}^{-1}(\beta_{ij})$ is an optimal set of n_{ij} -means, and $V_n = \sum_{i,j=1}^2 \frac{1}{36} V_{n_{ij}}$.

Proof. For $n \geq 4$, by Lemma 11, we have $\alpha = \bigcup_{i,j=1}^2 \beta_{ij}$, $n = \sum_{i,j=1}^2 n_{ij}$, and so

$$V_n = \sum_{i,j=1}^2 \int_{A_{(i,j)}} \min_{a \in \beta_{ij}} \|x - a\|^2 d(P_c \times P_c).$$

If $U_{(1,1)}^{-1}(\beta_{11})$ is not an optimal set of n_{11} -means for $P_c \times P_c$, then there exists a set $\gamma_{11} \subset \mathbb{R}^2$ with $\text{card}(\gamma_{11}) = n_{11}$ such that $\int \min_{a \in \gamma_{11}} \|x - a\|^2 d(P_c \times P_c) < \int \min_{a \in U_{(1,1)}^{-1}(\beta_{11})} \|x - a\|^2 d(P_c \times P_c)$. But then, $\delta := U_{(1,1)}(\gamma_{11}) \cup \beta_{12} \cup \beta_{21} \cup \beta_{22}$ is a set of cardinality n and it satisfies $\int \min_{a \in \delta} \|x - a\|^2 d(P_c \times P_c) < \int \min_{a \in \alpha} \|x - a\|^2 d(P_c \times P_c)$, contradicting the fact that α is an optimal set of n -means for $P_c \times P_c$. Similarly, it can be proved that $U_{(1,2)}^{-1}(\beta_{12})$, $U_{(2,1)}^{-1}(\beta_{21})$, and $U_{(2,2)}^{-1}(\beta_{22})$ are optimal sets of n_{12} -, n_{21} -, and n_{22} -means respectively. Thus,

$$V_n = \sum_{i,j=1}^2 \frac{1}{4} \int \min_{a \in \beta_{ij}} \|x - a\|^2 d((P_c \times P_c) \circ U_{(i,j)}^{-1}) = \sum_{i,j=1}^2 \frac{1}{36} \int \min_{a \in U_{(i,j)}^{-1}(\beta_{ij})} \|x - a\|^2 dP = \sum_{i,j=1}^2 \frac{1}{36} V_{n_{ij}},$$

which gives the lemma. \square

Proposition 5. Let $n \in \mathbb{N}$ be such that $n = 4^{\ell(n)}$ for some positive integer $\ell(n)$. Then, the set

$$\alpha_{4^{\ell(n)}} := \bigcup_{(\sigma, \tau) \in \{1,2\}^{\ell(n)*2}} \{(A(\sigma), A(\tau))\}$$

forms a unique optimal set of n -means for the affine measure P with quantization error

$$V_{4^{\ell(n)}} = \frac{1}{4} \frac{1}{9^{\ell(n)}}.$$

Proof. We will prove the statement by induction. By Corollary 1, it is true if $\ell(n) = 1$. Let us assume that it is true for $n = 4^k$ for some positive integer k . We now show that it is also true if $n = 4^{k+1}$. Let β be an optimal set of 4^{k+1} -means. Set $\beta_{ij} := \beta \cap A_{(i,j)}$ for $1 \leq i, j \leq 2$. Then, by Lemmas 11 and 13, $U_{(i,j)}^{-1}(\beta_{ij})$ is an optimal set of 4^k -means, and so $U_{(i,j)}^{-1}(\beta_{ij}) = \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1,2\}^{k*2}\}$ which implies $\beta_{ij} = \{(A(i\sigma), A(j\tau)) : (\sigma, \tau) \in \{1,2\}^{k*2}\}$. Thus, $\beta = \bigcup_{i,j=1}^2 \beta_{ij} = \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1,2\}^{(k+1)*2}\}$ is

an optimal set of 4^{k+1} -means. Because $(A(\sigma), A(\tau))$ is the centroid of $A_{(\sigma,\tau)}$ for each $(\sigma, \tau) \in \mathcal{I}^{k+1}$, the set β is unique. Now, by Lemma 13, we have the quantization error as

$$V_{k+1} = \sum_{i,j=1}^2 \frac{1}{36} V_k = \frac{1}{9} \cdot \frac{1}{4} \cdot \frac{1}{9^k} = \frac{1}{4 \cdot 9^{k+1}}.$$

Thus, by induction, the proof of the proposition is complete. \square

Definition 1. For $n \in \mathbb{N}$ with $n \geq 4$ let $\ell(n)$ be the unique natural number with $4^{\ell(n)} < n \leq 2 \cdot 4^{\ell(n)}$. For $I \subset \{1, 2\}^{\ell(n)*2}$ with $\text{card}(I) = n - 4^{\ell(n)}$ let $\alpha_n(I)$ be the set defined as follows:

$$\alpha_n(I) = \bigcup_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \{(A(\sigma), A(\tau))\} \cup \left(\bigcup_{(\sigma,\tau) \in I} \{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\} \right).$$

Remark 7. In Definition 1, instead of choosing the set $\{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\}$, one can choose $\{(A(\sigma), A(\tau 1)), (A(\sigma), A(\tau 2))\}$, i.e., the set associated with each $(\sigma, \tau) \in I$ can be chosen in two different ways. Moreover, the subset I can be chosen from $\{1, 2\}^{\ell(n)*2}$ in $4^{\ell(n)} C_{n-4^{\ell(n)}}$ ways. Hence, the number of the sets $\alpha_n(I)$ is $2^{\text{card}(I)} \cdot 4^{\ell(n)} C_{n-4^{\ell(n)}}$.

The following example illustrates Definition 1.

Example 2. Let $n = 5$. Then, $\ell(n) = 1$, $I \subset \{1, 2\}^{*2}$ with $\text{card}(I) = 1$, and so

$$\begin{aligned} \alpha_5(\{(1, 1)\}) &= \{(A(1), A(2)), (A(2), A(1)), (A(2), A(2))\} \cup \{(A(11), A(1)), (A(12), A(1))\} \\ &= \left\{ \left(\frac{1}{6}, \frac{5}{6}\right), \left(\frac{5}{6}, \frac{1}{6}\right), \left(\frac{5}{6}, \frac{5}{6}\right) \right\} \cup \left\{ \left(\frac{1}{18}, \frac{1}{6}\right), \left(\frac{5}{18}, \frac{1}{6}\right) \right\}, \end{aligned}$$

or,

$$\begin{aligned} \alpha_5(\{(1, 1)\}) &= \{(A(1), A(2)), (A(2), A(1)), (A(2), A(2))\} \cup \{(A(1), A(11)), (A(1), A(12))\} \\ &= \left\{ \left(\frac{1}{6}, \frac{5}{6}\right), \left(\frac{5}{6}, \frac{1}{6}\right), \left(\frac{5}{6}, \frac{5}{6}\right) \right\} \cup \left\{ \left(\frac{1}{6}, \frac{1}{18}\right), \left(\frac{1}{6}, \frac{5}{18}\right) \right\}. \end{aligned}$$

Similarly, one can get six more sets by taking $I = \{(1, 2)\}$, $\{(2, 1)\}$, or $\{(2, 2)\}$, i.e., the number of the sets $\alpha_n(I)$ in this case is $2^{\text{card}(I)} \cdot 4^{\ell(n)} C_{n-4^{\ell(n)}} = 8$.

Proposition 6. Let $n \geq 4$ and $\alpha_n(I)$ be the set as defined in Definition 1. Then, $\alpha_n(I)$ forms an optimal set of n -means with quantization error

$$V_n = \frac{1}{4} \frac{1}{36^{\ell(n)}} \left(2 \cdot 4^{\ell(n)} - n + \frac{5}{9} (n - 4^{\ell(n)}) \right).$$

Proof. We have $n = 4^{\ell(n)} + k$ where $1 \leq k \leq 4^{\ell(n)}$. Set $\beta_{ij} = \alpha \cap A_{ij}$ with $n_{ij} = \text{card}(\beta_{ij})$ for $1 \leq i, j \leq 2$. Let us prove it by induction. We first assume $k = 1$. By Lemmas 11 and 13, we can assume that each of $U_{(i,j)}^{-1}(\beta_{ij})$ for $i = 2$ and $j = 1, 2$, are optimal sets of $4^{\ell(n)-1}$ -means and $U_{(1,1)}^{-1}(\beta_{11})$ is an optimal set of $(4^{\ell(n)-1} + 1)$ -means. Thus, for $i = 2$ and $j = 1, 2$, we can write

$$\begin{aligned} U_{(i,j)}^{-1}(\beta_{ij}) &= \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{(\ell(n)-1)*2}\}, \text{ and} \\ U_{(1,1)}^{-1}(\beta_{11}) &= \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{(\ell(n)-1)*2} \setminus \{\tau\}\} \cup U_{\tau}(\alpha_2), \end{aligned}$$

for some $\tau \in \{1, 2\}^{(\ell(n)-1)*2}$, where α_2 is an optimal set of two-means. Thus,

$$\alpha_n(\{(1, 1)\tau\}) = \bigcup_{i,j=1}^2 \beta_{ij} = \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus \{(1, 1)\tau\}\} \cup U_{(1,1)\tau}(\alpha_2),$$

for some $\tau \in \{1, 2\}^{(\ell(n)-1)*2}$, where α_2 is an optimal set of two-means. Notice that instead of choosing $U_{(1,1)}^{-1}(\beta_{11})$ as an optimal set of $(4^{\ell(n)-1} + 1)$ -means, one can choose any one from $U_{(i,j)}^{-1}(\beta_{ij})$ for $i = 2, j = 1, 2$, as an optimal set of $(4^{\ell(n)-1} + 1)$ -means. Hence, for $n = 4^{\ell(n)} + 1$, one can write

$$\alpha_n(I) = \bigcup_{i,j=1}^2 \beta_{ij} = \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus \{\tau\}\} \cup U_\tau(\alpha_2),$$

where $I = \{\tau\}$ for some $\tau \in \{1, 2\}^{\ell(n)*2}$ as an optimal set of n -means. Thus, we see that the proposition is true if $n = 4^{\ell(n)} + k$. Similarly, one can prove that the proposition is true for any $1 \leq k \leq 4^{\ell(n)}$. Then, the quantization error is

$$\begin{aligned} V_n &= \min_{(a,b) \in \alpha_n(I)} \|x - (a, b)\|^2 dP = \sum_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \int_{A_\sigma \times A_\tau} \|x - (A(\sigma), A(\tau))\|^2 d(P_c \times P_c) \\ &\quad + \sum_{(\sigma,\tau) \in I} \sum_{i=1}^2 \int_{A_{\sigma i} \times A_\tau} \|x - (A(\sigma i), A(\tau))\|^2 d(P_c \times P_c) \\ &= \sum_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \frac{1}{4^{\ell(n)}} (u_\sigma^2 + u_\tau^2) \frac{1}{8} + \sum_{(\sigma,\tau) \in I} \sum_{i=1}^2 \frac{1}{4^{\ell(n)}} \frac{1}{2} (u_{\sigma i}^2 + u_\tau^2) \frac{1}{8} \\ &= \sum_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \frac{1}{4^{\ell(n)}} (u_\sigma^2 + u_\tau^2) \frac{1}{8} + \sum_{(\sigma,\tau) \in I} \frac{1}{4^{\ell(n)}} \left(\frac{1}{9} u_\sigma^2 + u_\tau^2\right) \frac{1}{8}. \end{aligned}$$

Because $\text{card}(\{1, 2\}^{\ell(n)*2} \setminus I) = 2 \cdot 4^{\ell(n)} - n$, $\text{card}(I) = n - 4^{\ell(n)}$, $u_\sigma = u_\tau = \frac{1}{3^{\ell(n)}}$, upon simplification, we have $V_n = \frac{1}{4} \frac{1}{36^{\ell(n)}} \left(2 \cdot 4^{\ell(n)} - n + \frac{5}{9}(n - 4^{\ell(n)})\right)$. Thus, the proof of the proposition is complete. \square

Definition 2. For $n \in \mathbb{N}$ with $n \geq 4$ let $\ell(n)$ be the unique natural number with $2 \cdot 4^{\ell(n)} < n < 4^{\ell(n)+1}$. For $I \subset \{1, 2\}^{\ell(n)*2}$ with $\text{card}(I) = n - 2 \cdot 4^{\ell(n)}$ let $\alpha_n(I)$ be the set defined as follows:

$$\begin{aligned} \alpha_n(I) &= \bigcup_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\} \\ &\quad \cup \left(\bigcup_{(\sigma,\tau) \in I} \{(A(\sigma 1), A(\tau 1)), (A(\sigma 1), A(\tau 2)), (A(\sigma 2), A(\tau))\} \right). \end{aligned}$$

Remark 8. In Definition 2, instead of choosing the set $\{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\}$, one can choose $\{(A(\sigma), A(\tau 1)), (A(\sigma), A(\tau 2))\}$. Instead of choosing the set $\{(A(\sigma 1), A(\tau 1)), (A(\sigma 1), A(\tau 2)), (A(\sigma 2), A(\tau))\}$, one can choose either the set $\{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau 1)), (A(\sigma 2), A(\tau 2))\}$, or $\{(A(\sigma 1), A(\tau 1)), (A(\sigma 2), A(\tau 1)), (A(\sigma), A(\tau 2))\}$, or $\{(A(\sigma), A(\tau 1)), (A(\sigma 1), A(\tau 2)), (A(\sigma 2), A(\tau 2))\}$, i.e., the set corresponding to each $(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I$ can be chosen in two different ways, and the set corresponding to each $(\sigma, \tau) \in I$ can be chosen in four different ways. Because $\text{card}(\{1, 2\}^{\ell(n)*2} \setminus I) = 4^{\ell(n)} - (n - 2 \cdot 4^{\ell(n)}) = 3 \cdot 4^{\ell(n)} - n$ and the subset I can be chosen from $\{1, 2\}^{\ell(n)*2}$ in $4^{\ell(n)} C_{n-2 \cdot 4^{\ell(n)}}^{\ell(n)}$ ways, the number of the sets $\alpha_n(I)$ is $2^{3 \cdot 4^{\ell(n)} - n} \cdot 4^{\ell(n) \text{card}(I)} \cdot 4^{\ell(n)} C_{n-2 \cdot 4^{\ell(n)}}^{\ell(n)}$.

We now give an example illustrating Definition 2.

Example 3. Let $n = 9$. Then, $\ell(n) = 1$, $I \subset \{1,2\}^{*2}$ with $\text{card}(I) = 1$. Take $I = \{(1,1)\}$. Then,

$$\begin{aligned} \alpha_9(\{(1,1)\}) &= \{(A(11), A(2)), (A(12), A(2)), (A(21), A(2)), (A(22), A(2)), (A(21), A(1)), \\ &\quad (A(22), A(1))\} \cup \{(A(11), A(1)), (A(12), A(11)), (A(12), A(12))\} \\ &= \left\{ \left(\frac{1}{18}, \frac{5}{6}\right), \left(\frac{5}{18}, \frac{5}{6}\right), \left(\frac{13}{18}, \frac{5}{6}\right), \left(\frac{17}{18}, \frac{5}{6}\right), \left(\frac{13}{18}, \frac{1}{6}\right), \left(\frac{17}{18}, \frac{1}{6}\right) \right\} \\ &\quad \cup \left\{ \left(\frac{1}{18}, \frac{1}{6}\right), \left(\frac{5}{18}, \frac{1}{18}\right), \left(\frac{5}{18}, \frac{5}{18}\right) \right\}. \end{aligned}$$

Note that each of $\alpha_9(\{(1,1)\})$, $\alpha_9(\{(1,2)\})$, $\alpha_9(\{(2,1)\})$, $\alpha_9(\{(2,2)\})$ can be chosen in 32 ways, i.e., the numbers of the sets $\alpha_9(I)$ in this case is $4 \cdot 32 = 128$. Moreover, by using the formula in Remark 8, we have

$$2^{3 \cdot 4^{\ell(n)} - n} \cdot 4^{\text{card}(I)} \cdot 4^{\ell(n)} C_{n-2 \cdot 4^{\ell(n)}} = 128.$$

Proposition 7. Let $n \geq 4$ and $\alpha_n(I)$ be the set as defined in Definition 2. Then, $\alpha_n(I)$ forms an optimal set of n -means with quantization error

$$V_n = \frac{1}{36^{\ell(n)+1}} (9 \cdot 4^{\ell(n)} - 2n).$$

Proof. We have $n = 2 \cdot 4^{\ell(n)} + k$ where $1 \leq k < 2 \cdot 4^{\ell(n)}$. Set $\beta_{ij} = \alpha \cap A_{ij}$ with $n_{ij} = \text{card}(\beta_{ij})$ for $1 \leq i, j \leq 2$. Let us prove it by induction. We first assume $k = 1$. By Lemmas 11 and 13, we can assume that each of $U_{(i,j)}^{-1}(\beta_{ij})$ for $i = 2$ and $j = 1, 2$, are optimal sets of $2 \cdot 4^{\ell(n)-1}$ -means and $U_{(1,1)}^{-1}(\beta_{11})$ is an optimal set of $(2 \cdot 4^{\ell(n)-1} + 1)$ -means. Thus, for $i = 2$ and $j = 1, 2$, we can write

$$\begin{aligned} U_{(i,j)}^{-1}(\beta_{ij}) &= \{U_{(\sigma,\tau)}(\alpha_2) : (\sigma, \tau) \in \{1,2\}^{(\ell(n)-1)*2}\}, \text{ and} \\ U_{(1,1)}^{-1}(\beta_{11}) &= \{U_{(\sigma,\tau)}(\alpha_2) : (\sigma, \tau) \in \{1,2\}^{(\ell(n)-1)*2} \setminus \{\tau\}\} \cup U_{\tau}(\alpha_3), \end{aligned}$$

for some $\tau \in \{1,2\}^{(\ell(n)-1)*2}$, where α_3 is an optimal set of three-means. Thus

$$\alpha_n(\{(1,1)\tau\}) = \bigcup_{i,j=1}^2 \beta_{ij} = \{U_{(\sigma,\tau)}(\alpha_2) : (\sigma, \tau) \in \{1,2\}^{\ell(n)*2} \setminus \{(1,1)\tau\}\} \cup U_{(1,1)\tau}(\alpha_3),$$

for some $\tau \in \{1,2\}^{(\ell(n)-1)*2}$, where α_3 is an optimal set of three-means. Notice that instead of choosing $U_{(1,1)}^{-1}(\beta_{11})$ as an optimal set of $(2 \cdot 4^{\ell(n)-1} + 1)$ -means, one can choose any one from $U_{(i,j)}^{-1}(\beta_{ij})$ for $i = 2, j = 1, 2$, as an optimal set of $(2 \cdot 4^{\ell(n)-1} + 1)$ -means. Hence, for $n = 2 \cdot 4^{\ell(n)} + 1$, one can write

$$\alpha_n(I) = \bigcup_{i,j=1}^2 \beta_{ij} = \{U_{(\sigma,\tau)}(\alpha_2) : (\sigma, \tau) \in \{1,2\}^{\ell(n)*2} \setminus \{\tau\}\} \cup U_{\tau}(\alpha_3),$$

where $I = \{\tau\}$ for some $\tau \in \{1,2\}^{\ell(n)*2}$ as an optimal set of n -means. Thus, we see that the proposition is true if $n = 2 \cdot 4^{\ell(n)} + 1$. Similarly, one can prove that the proposition is true for any $1 \leq k < 2 \cdot 4^{\ell(n)}$. Thus, writing $\alpha_2 = \{(A(1), A(\emptyset)), (A(2), A(\emptyset))\}$, and $\alpha_3 = \{(A(1), A(1)), (A(1), A(2)), (A(2), A(\emptyset))\}$, we have, in general,

$$\begin{aligned} \alpha_n(I) &= \bigcup_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\} \\ &\quad \cup \left(\bigcup_{(\sigma,\tau) \in I} \{(A(\sigma 1), A(\tau 1)), (A(\sigma 1), A(\tau 2)), (A(\sigma 2), A(\tau))\} \right), \end{aligned}$$

where $I \subset \{1, 2\}^{\ell(n)*2}$ with $\text{card}(I) = k$ for some $1 \leq k < 2 \cdot 4^{\ell(n)}$. Then, we obtain the quantization error as

$$\begin{aligned} V_n &= \min_{(a,b) \in \beta_n(I)} \|x - (a, b)\|^2 dP = \sum_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I} \sum_{i=1}^2 \int_{A_{\sigma i} \times A_{\tau}} \|x - (A(\sigma i), A(\tau))\|^2 d(P_c \times P_c) \\ &\quad + \sum_{(\sigma, \tau) \in I} \left(\sum_{j=1}^2 \int_{A_{\sigma 1} \times A_{\tau j}} \|x - (A(\sigma 1), A(\tau j))\|^2 d(P_c \times P_c) \right. \\ &\quad \left. + \int_{A_{\sigma 2} \times A_{\tau}} \|x - (A(\sigma 2), A(\tau))\|^2 d(P_c \times P_c) \right) \\ &= \sum_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I} \sum_{i=1}^2 \frac{1}{4^{\ell(n)}} \frac{1}{2} (u_{\sigma i}^2 + u_{\tau}^2) \frac{1}{8} + \sum_{(\sigma, \tau) \in I} \frac{1}{4^{\ell(n)}} \left(\sum_{j=1}^2 \frac{1}{4} (u_{\sigma 1}^2 + u_{\tau j}^2) \frac{1}{8} + \frac{1}{2} (u_{\sigma 2}^2 + u_{\tau}^2) \frac{1}{8} \right) \\ &= \sum_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I} \frac{1}{4^{\ell(n)}} \left(\frac{1}{9} u_{\sigma}^2 + u_{\tau}^2 \right) \frac{1}{8} + \sum_{(\sigma, \tau) \in I} \frac{1}{4^{\ell(n)}} (u_{\sigma}^2 + 5u_{\tau}^2) \frac{1}{72}. \end{aligned}$$

Because $\text{card}(\{1, 2\}^{\ell(n)*2} \setminus I) = 3 \cdot 4^{\ell(n)} - n$, $\text{card}(I) = n - 2 \cdot 4^{\ell(n)}$, $u_{\sigma} = u_{\tau} = \frac{1}{3^{\ell(n)}}$, upon simplification, we have $V_n = \frac{1}{36^{\ell(n)+1}} (9 \cdot 4^{\ell(n)} - 2n)$. Thus, the proof of the proposition is complete. \square

6. Quantization Dimension and Quantization Coefficient for P

The techniques employed in the previous sections also provide closed formulas for the quantization errors involved at each step. Such closed formulas are amenable for direct calculation of the quantization dimension and the quantization coefficient for the probability distribution involved. Hence, in this section we will calculate the quantization dimension $D(P)$ of the probability distribution P , and the accumulation points for the $D(P)$ -dimensional quantization coefficients. By Propositions 5-7, the n th quantization error V_n is given by

$$V_n = \begin{cases} \frac{1}{4} \frac{1}{36^{\ell(n)}} \left(2 \cdot 4^{\ell(n)} - n + \frac{5}{9} (n - 4^{\ell(n)}) \right) & \text{if } 4^{\ell(n)} \leq n \leq 2 \cdot 4^{\ell(n)}, \\ \frac{1}{36^{\ell(n)+1}} (9 \cdot 4^{\ell(n)} - 2n) & \text{if } 2 \cdot 4^{\ell(n)} < n < 4^{\ell(n)+1}. \end{cases} \tag{8}$$

Proposition 8. *The quantization dimension $D(P)$ of the probability distribution P exists and equals $\frac{\log 4}{\log 3}$.*

Proof. By (8), for $4^{\ell(n)} \leq n \leq 2 \cdot 4^{\ell(n)}$, it follows that $V_{2 \cdot 4^{\ell(n)}} \leq V_n \leq V_{4^{\ell(n)}}$, i.e.,

$$\frac{5}{36} 9^{-\ell(n)} \leq V_n \leq \frac{1}{4} 9^{-\ell(n)},$$

and so

$$\frac{2\ell(n) \log 4}{-\log \frac{5}{36} + \ell(n) \log 9} \leq \frac{2 \log n}{-\log V_n} < \frac{2 \log 2 + 2\ell(n) \log 4}{-\log \frac{1}{4} + \ell(n) \log 9}.$$

Thus, we deduce that

$$\lim_{n \rightarrow \infty} \frac{2 \log n}{-\log V_n} = \frac{\log 4}{\log 3}.$$

Similarly, for $2 \cdot 4^{\ell(n)} < n < 4^{\ell(n)+1}$, we also obtain the same limit. Hence,

$$D(P) = \lim_{n \rightarrow \infty} \frac{2 \log n}{-\log V_n} = \frac{\log 4}{\log 3}.$$

Thus, the proof of the proposition is complete. \square

Proposition 9. Let $\beta := D(P)$ be the quantization dimension of P . Then, the β -dimensional quantization coefficient for P does not exist, and the accumulation points of $\{n^{\frac{2}{\beta}} V_n\}_{n \in \mathbb{N}}$ lie in the closed interval $[\frac{1}{12}, \frac{5}{4}]$.

Proof. Recall the sequence of quantization errors $\{V_n\}_{n=4}^\infty$ given by (8). Again, notice that $4^{\frac{1}{\beta}} = 3$. Along the sequence $\{4^{\ell(n)}\}_{n \in \mathbb{N}}$, we have $\lim_{n \rightarrow \infty} (4^{\ell(n)})^{\frac{2}{\beta}} V_{4^{\ell(n)}} = \frac{1}{4}$. Similarly, along the sequence $\{2 \cdot 4^{\ell(n)}\}_{n \in \mathbb{N}}$, we have $\lim_{n \rightarrow \infty} (2 \cdot 4^{\ell(n)})^{\frac{2}{\beta}} V_{2 \cdot 4^{\ell(n)}} = \frac{5}{12}$. Consequently, $\lim_{n \rightarrow \infty} n^{\frac{2}{\beta}} V_n$ does not exist. Now, we calculate the range for the accumulation points of $\{n^{\frac{2}{\beta}} V_n\}_{n \in \mathbb{N}}$. The following two cases can arise:

Case 1. $4^{\ell(n)} \leq n \leq 2 \cdot 4^{\ell(n)}$.

In this case, we have $V_{2 \cdot 4^{\ell(n)}} \leq V_n \leq V_{4^{\ell(n)}}$, implying $(4^{\ell(n)})^{\frac{2}{\beta}} V_{2 \cdot 4^{\ell(n)}} \leq n^{\frac{2}{\beta}} V_n \leq (2 \cdot 4^{\ell(n)})^{\frac{2}{\beta}} V_{4^{\ell(n)}}$. Because

$$\lim_{n \rightarrow \infty} (4^{\ell(n)})^{\frac{2}{\beta}} V_{2 \cdot 4^{\ell(n)}} = \frac{5}{36}, \text{ and } \lim_{n \rightarrow \infty} (2 \cdot 4^{\ell(n)})^{\frac{2}{\beta}} V_{4^{\ell(n)}} = \frac{3}{4},$$

it follows that along such subsequences, we have $\liminf_n n^{\frac{2}{\beta}} V_n = \frac{5}{36} < \frac{3}{4} = \limsup_n n^{\frac{2}{\beta}} V_n$.

Case 2. $2 \cdot 4^{\ell(n)} < n < 4^{\ell(n)+1}$.

In this case, we have $V_{4^{\ell(n)+1}} < V_n < V_{2 \cdot 4^{\ell(n)}}$, implying

$$(2 \cdot 4^{\ell(n)})^{\frac{2}{\beta}} V_{4^{\ell(n)+1}} < n^{\frac{2}{\beta}} V_n < (4^{\ell(n)+1})^{\frac{2}{\beta}} V_{2 \cdot 4^{\ell(n)}}.$$

Because

$$\lim_{n \rightarrow \infty} (2 \cdot 4^{\ell(n)})^{\frac{2}{\beta}} V_{4^{\ell(n)+1}} = \frac{1}{12}, \text{ and } \lim_{n \rightarrow \infty} (4^{\ell(n)+1})^{\frac{2}{\beta}} V_{2 \cdot 4^{\ell(n)}} = \frac{5}{4},$$

it follows that $\liminf_n n^{\frac{2}{\beta}} V_n = \frac{1}{12} < \frac{5}{4} = \limsup_n n^{\frac{2}{\beta}} V_n$.

By Case 1 and Case 2, for $n \in \mathbb{N}$, we see that

$$\liminf_n n^{\frac{2}{\beta}} V_n = \frac{1}{12} < \frac{5}{4} = \limsup_n n^{\frac{2}{\beta}} V_n,$$

which yields the fact that the accumulation points of $\{n^{\frac{2}{\beta}} V_n\}_{n \in \mathbb{N}}$ lie in the closed interval $[\frac{1}{12}, \frac{5}{4}]$. Thus, the proof of the proposition is complete. \square

7. Discussion and Concluding Remarks

Motivation. As it has been mentioned in Introduction, the main motivation for this article is completion of the programme initiated in [14]. In the meantime, we extend the results in [12] to the setting of infinite affine transformations. Analogously to [10], this completes the programme of providing complete quantization for affine measures on \mathbb{R}^2 .

Observations and Remarks. Quantization of continuous random signals (or random variables and processes) is an important part of digital representation of analog signals for various coding techniques (e.g., source coding, data compression, archiving, restoration). The oldest example of quantization in statistics is rounding off. Sheppard (see [19]) was the first who analyzed rounding off for estimating densities by histograms. Any real number x can be rounded off (or quantized) to the nearest integer, say $q(x) = [x]$, with a resulting quantization error $e(x) = x - q(x)$. Hence, the restored signal may differ from the original one and some information can be lost. Thus, in quantization of a continuous set of values there is always a distortion (also known as noise or error) between the original set of values and the quantized set of values. The main goal in quantization theory is finding a set of quantizers with minimum distortion, which has been extensively investigated by numerous authors [2,20–24]. A different approach for uniform scalar quantization is

developed in [25], where the correlation properties of a Gaussian process are exploited to evaluate the asymptotic behavior of the random quantization rate for uniform quantizers. General quantization problems for Gaussian processes in infinite-dimensional functional spaces are considered in [26]. In estimating weighted integrals of time series with no quadratic mean derivatives, by means of samples at discrete times, it is known that the rate of convergence of mean-square error is reduced from n^{-2} to $n^{-1.5}$ when the samples are quantized (see [27]). For smoother time series, with $k = 1, 2, \dots$ quadratic mean derivatives, the rate of convergence is reduced from n^{-2k-2} to n^{-2} when the samples are quantized, which is a very significant reduction (see [28]). The interplay between sampling and quantization is also studied in [28], which asymptotically leads to optimal allocation between the number of samples and the number of levels of quantization. Quantization also seems to be a promising tool in recent development in numerical probability (see, e.g., [29]).

By Proposition 1 the points in an optimal set are the centroids of their own Voronoi regions. Consequently, the points in an optimal set are an evenly spaced distribution of sites in the domain with minimum distortion error with respect to a given probability measure and is therefore very useful in many fields, such as clustering, data compression, optimal mesh generation, cellular biology, optimal quadrature, coverage control and geographical optimization; for more details one can see [7,30]. In addition, it has applications in energy-efficient distribution of base stations in a cellular network [31–33]. In both geographical and cellular applications the distribution of users is highly complex and often modeled by a fractal [34,35].

Future Directions. k -means clustering is a method of vector quantization, originally from signal processing, that aims to partition n observations, or the underlying data set into k clusters in which each observation belongs to the cluster with the nearest mean, also known as cluster center or cluster centroid. For a given k and a given probability distribution in a dataset there can be two or more different sets of k -means clusters: for example, with respect to a uniform distribution the unit square $\{(x_1, x_2) : |x_1| \leq 1, |x_2| \leq 1\}$ has four different sets of two-means clusters with cluster centers $\{(\frac{1}{2}, \frac{1}{2}), (-\frac{1}{2}, -\frac{1}{2})\}$, $\{(-\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, -\frac{1}{2})\}$, $\{(-\frac{1}{2}, 0), (\frac{1}{2}, 0)\}$, and $\{(0, \frac{1}{2}), (0, -\frac{1}{2})\}$. Among these only $\{(-\frac{1}{2}, 0), (\frac{1}{2}, 0)\}$, and $\{(0, \frac{1}{2}), (0, -\frac{1}{2})\}$ form two different optimal sets of two-means. In other words, we can say that for a given k , among the multiple sets of k -means clusters, the centers of a set with the smallest distortion error form an optimal set of k -means. Thus, it is much more difficult to calculate an optimal set of k -means than to calculate a set of k -means clusters. There are several work done in the direction of k -means clustering. On the other hand, there is not much work in the direction of finding optimal sets of k -means clusters, and the work in this paper is an addition in this direction.

The probability measure P considered in this study has identical marginal distributions, which is instrumental in determining optimal sets of 2-, 3-, and 4-means accurately. Besides, it enables us to bridge infinitely generated affine measures with finitely generated ones, and consequently, connect optimal sets of n -means for P and $P_C \times P_C$. It would be interesting to investigate if similar results can be achieved when P is induced by different infinite probability vectors $\{p_{ij}\}$ than considered in this article.

Author Contributions: The work in this paper is completely new. Both the authors equally contributed in writing the draft of this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bucklew, J.A.; Wise, G.L. Multidimensional asymptotic quantization with r th power distortion measures. *IEEE Trans. Inf. Theory* **1982**, *28*, 239–247. [[CrossRef](#)]
2. Gray, R.; Neuhoff, D. Quantization. *IEEE Trans. Inform. Theory* **1998**, *44*, 2325–2383. [[CrossRef](#)]
3. Abaya, E.F.; Wise, G.L. Some remarks on the existence of optimal quantizers. *Stat. Probab. Lett.* **1984**, *2*, 349–351. [[CrossRef](#)]
4. Gray, R.M.; Kieffer, J.C.; Linde, Y. Locally optimal block quantizer design. *Inf. Control.* **1980**, *45*, 178–198. [[CrossRef](#)]
5. György, A.; Linder, T. On the structure of optimal entropy-constrained scalar quantizers. *IEEE Trans. Inf. Theory* **2002**, *48*, 416–427. [[CrossRef](#)]
6. Graf, S.; Luschgy, H. *Foundations of Quantization for Probability Distributions*; Lecture Notes in Mathematics; Springer: Berlin/Heidelberg, Germany, 2000; Volume 1730.
7. Du, Q.; Faber, V.; Gunzburger, M. Centroidal Voronoi Tessellations: Applications and Algorithms. *Siam Rev.* **1999**, *41*, 637–676. [[CrossRef](#)]
8. Roychowdhury, M.K. Quantization and centroidal Voronoi tessellations for probability measures on dyadic Cantor sets. *J. Fractal Geom.* **2017**, *4*, 127–146. [[CrossRef](#)]
9. Gersho, A.; Gray, R.M. *Vector Quantization and Signal Compression*; Kluwer Academy Publishers: Boston, MA, USA, 1992.
10. Graf, S.; Luschgy, H. The Quantization of the Cantor Distribution. *Math. Nachr.* **1997**, *183*, 113–133. [[CrossRef](#)]
11. Roychowdhury, L. Optimal quantization for nonuniform Cantor distributions. *J. Interdiscip. Math.* **2019**, *22*, 1325–1348. [[CrossRef](#)]
12. Çömez, D.; Roychowdhury, M.K. Quantization for uniform distributions of Cantor dusts on \mathbb{R}^2 . *Topol. Proc.* **2020**, *56*, 195–218.
13. Roychowdhury, M.K. Optimal quantization for the Cantor distribution generated by infinite similitudes. *Isr. J. Math.* **2019**, *231*, 437–466. [[CrossRef](#)]
14. Mihailescu, E.; Roychowdhury, M.K. Quantization coefficients in infinite systems. *Kyoto J. Math.* **2015**, *55*, 857–873. [[CrossRef](#)]
15. Graf, S.; Luschgy, H. The quantization dimension of self-similar probabilities. *Math. Nachr.* **2002**, *241*, 103–109. [[CrossRef](#)]
16. Hutchinson, J. Fractals and self-similarity. *Indiana Univ. J.* **1981**, *30*, 713–747. [[CrossRef](#)]
17. Moran, M. Hausdorff measure of infinitely generated self-similar sets. *Monatsh. Math.* **1996**, *122*, 387–399. [[CrossRef](#)]
18. Mauldin, D.; Urbański, M. Dimensions and measures in infinite iterated function systems. *Proc. Lond. Math. Soc.* **1996**, *73*, 105–154. [[CrossRef](#)]
19. Sheppard, W.F. On the calculation of the most probable values of frequency constants for data arranged according to equidistant divisions of a scale. *Proc. Lond. Math. Soc.* **1897**, *1*, 353–380. [[CrossRef](#)]
20. Cambanis, S.; Gerr, N. A simple class of asymptotically optimal quantizers. *IEEE Trans. Inform. Theory* **1983**, *29*, 664–676. [[CrossRef](#)]
21. Gray, R.M.; Linder, T. Mismatch in high rate entropy constrained vector quantization. *IEEE Trans. Inform. Theory* **2003**, *49*, 1204–1217. [[CrossRef](#)]
22. Li, J.; Chaddha, N.; Gray, R.M. Asymptotic performance of vector quantizers with a perceptual distortion measure. *IEEE Trans. Inform. Theory* **1999**, *45*, 1082–1091.
23. Shykula, M.; Seleznev, O. Stochastic structure of asymptotic quantization errors. *Stat. Prob. Lett.* **2006**, *76*, 453–464. [[CrossRef](#)]
24. Zador, P.L. Asymptotic quantization error of continuous signals and the quantization dimensions. *IEEE Trans. Inform. Theory* **1982**, *28*, 139–148. [[CrossRef](#)]
25. Shykula, M.; Seleznev, O. *Uniform Quantization of Random Processes*; Univ. Umeå Research Report; Umeå University: Umeå, Sweden, 2004; pp. 1–16.
26. Luschgy, H.; Pagès, G. Functional quantization of Gaussian processes. *J. Funct. Anal.* **2002**, *196*, 486–531. [[CrossRef](#)]
27. Bucklew, J.A.; Cambanis, S. Estimating random integrals from noisy observations: Sampling designs and their performance. *IEEE Trans. Inform. Theory* **1988**, *34*, 111–127. [[CrossRef](#)]
28. Benhenni, K.; Cambanis, S. The effect of quantization on the performance of sampling designs. *IEEE Trans. Inform. Theory* **1998**, *44*, 1981–1992. [[CrossRef](#)]
29. Pagès, G.; Pham, H.; Printemps, J. Optimal quantization methods and applications to numerical problems in finance. In *Handbook of Computational and Numerical Methods in Finance*; Rachev, S., Ed.; Birkhäuser Boston: Boston, MA, USA, 2004; pp. 253–297.
30. Okabe, A.; Boots, B.; Sugihara, K.; Chiu, S.N. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2000.
31. Hao, Y.; Chen, M.; Hu, L.; Song, J.; Volk, M.; Humar, I. Wireless Fractal Ultra-Dense Cellular Networks. *Sensors* **2017**, *17*, 841. [[CrossRef](#)]
32. Kaza, K.R.; Kshirsagar, K.; Rajan, K.S. A bi-objective algorithm for dynamic reconfiguration of mobile networks. In Proceedings of the IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 10–15 June 2012; pp. 5741–5745.
33. Song, Y. Cost-Effective Algorithms for Deployment and Sensing in Mobile Sensor Networks. Ph.D. Thesis, University of Connecticut, Storrs, CT, USA, 2014.
34. Abundo, C.; Bodnar, T.; Driscoll, J.; Hatton, I.; Wright, J. City population dynamics and fractal transport networks. In *Proceedings of the Santa Fe Institute's CSSS2013*; Santa Fe Institute: Santa Fe, NM, USA, 2013.
35. Lu, Z.; Zhang, H.; Southworth, F.; Crittenden, J. Fractal dimensions of metropolitan area road networks and the impacts on the urban built environment. *Ecol. Indic.* **2016**, *70*, 285–296. [[CrossRef](#)]