University of Texas Rio Grande Valley

ScholarWorks @ UTRGV

Mathematical and Statistical Sciences Faculty Publications and Presentations

College of Sciences

5-24-2022

Quantization for a probability distribution generated by an infinite iterated function system

Lakshmi Roychowdhury

Mrinal Kanti Roychowdhury

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

Part of the Mathematics Commons

Commun. Korean Math. Soc. **37** (2022), No. 3, pp. 765–800
https://doi.org/10.4134/CKMS.c210266
pISSN: 1225-1763 / eISSN: 2234-3024

QUANTIZATION FOR A PROBABILITY DISTRIBUTION GENERATED BY AN INFINITE ITERATED FUNCTION SYSTEM

Lakshmi Roychowdhury and Mrinal Kanti Roychowdhury

Abstract. Quantization for probability distributions concerns the best approximation of a d-dimensional probability distribution P by a discrete probability with a given number n of supporting points. In this paper, we have considered a probability measure generated by an infinite iterated function system associated with a probability vector on $\mathbb{R}$. For such a probability measure P, an induction formula to determine the optimal sets of n-means and the nth quantization error for every natural number n is given. In addition, using the induction formula we give some results and observations about the optimal sets of n-means for all $n \geq 2$.

1. Introduction

Quantization is the process of converting a continuous analog signal into a digital signal of k discrete levels, or converting a digital signal of n levels into another digital signal of k levels, where $k < n$. It is must when analog quantities are represented, processed, stored, or transmitted by a digital system, or when data compression is required. It is a classic and still very active research topic in source coding and information theory. A good survey about the historical development of the theory has been provided by Gray and Neuhoff in [8]. For more applied aspects of quantization the reader is referred to the book of Gersho and Gray (see [4]). For mathematical treatment of quantization one may consult Graf-Luschgy's book (see [7]). Interested readers can also see [1, 5, 9, 16]. Let $\mathbb{R}^d$ denote the d-dimensional Euclidean space equipped with the Euclidean metric $\|\cdot\|$. Let P be a Borel probability measure on $\mathbb{R}^d$. Then, the nth quantization error for P, denoted by $V_n := V_n(P)$, is defined by

$$V_n(P) = \inf_{\alpha \in \mathcal{D}_n} \int \min_{a \in \alpha} \|x - a\|^2 dP(x),$$

Received August 6, 2021; Revised December 27, 2021; Accepted January 19, 2022.
2020 *Mathematics Subject Classification.* Primary 60Exx, 28A80, 94A34.
Key words and phrases. Probability measure, infinite iterated function system, optimal set, quantization error.
The research of the second author was supported by U.S. National Security Agency (NSA) Grant H98230-14-1-0320.

©2022 Korean Mathematical Society

where $\mathcal{D}_n := \{\alpha \subset \mathbb{R}^d : 1 \leq \text{card}(\alpha) \leq n\}$. The set α for which the infimum occurs and contains no more than n points is called an optimal set of n-means for P, and such a set exists if $\int \|x\|^2 dP(x) < \infty$ (see [5, 7, 9]). The set of all optimal sets of n-means for a probability measure P is denoted by $\mathcal{C}_n(P)$. It is known that for a Borel probability measure P if the support of P contains infinitely many elements, then an optimal set of n-means always has exactly n-elements (see [7, Theorem 4.12]). Let α be a finite set and $a \in \alpha$. Then, the *Voronoi cell*, or *Voronoi region* $M(a|\alpha)$ is the set of all elements in $\mathbb{R}^d$ whose distance to a is not greater than their distance to other elements in α, i.e.,

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

A Borel measurable partition $\{A_a : a \in \alpha\}$ of $\mathbb{R}^d$ is called a *Voronoi partition* of $\mathbb{R}^d$ with respect to α (and P) if $A_a \subset M(a|\alpha)$ (P-a.e.) for every $a \in \alpha$. The following proposition is known (see [4, 9]).

Proposition 1.1. *Let α be an optimal set of n-means, $a \in \alpha$, and $M(a|\alpha)$ be the Voronoi region generated by $a \in \alpha$. Then, for every $a \in \alpha$,* (i) $P(M(a|\alpha)) > 0$*,* (ii) $P(\partial M(a|\alpha)) = 0$*,* (iii) $a = E(X : X \in M(a|\alpha))$*, and* (iv) *$P$-almost surely the set $\{M(a|\alpha) : a \in \alpha\}$ forms a Voronoi partition of $\mathbb{R}^d$.*

Since for $a \in \alpha$, $a = E(X : X \in M(a|\alpha)) = \frac{1}{P(M(a|\alpha))} \int_{M(a|\alpha)} x dP(x)$, we can say that the elements in an optimal set of n-means are also the centroids of their own Voronoi regions with respect to the probability distribution P. For details in this regard one can see [3, 14].

Let M denote either the set $\{1, 2, \ldots, N\}$ for some positive integer $N \geq 2$, or the set $\mathbb{N}$ of natural numbers. A collection $\{S_j : j \in M\}$ of similarity mappings, or similitudes, on $\mathbb{R}^d$ with similarity ratios $\{s_j : j \in M\}$ is contractive if $\sup\{s_j : j \in M\} < 1$. If J is the limit set of the iterated function system, then it is known that J satisfies the following invariance relation (see [10–12]):

$$J = \bigcup_{j \in M} S_j(J).$$

The iterated function system $\{S_j : j \in M\}$ satisfies the *open set condition* (OSC) if there exists a bounded nonempty open set $U \subset \mathbb{R}^d$ such that $S_j(U) \subset U$ for all $j \in M$, and $S_i(U) \bigcap S_j(U) = \emptyset$ for $i, j \in M$ with $i \neq j$. Let $(p_j : j \in M)$ be a probability vector, with $p_j > 0$ for all $j \in M$. Then, there exists a unique Borel probability measure P on $\mathbb{R}^d$ (see [10–12], etc.) such that

$$P = \sum_{j \in M} p_j P \circ S_j^{-1},$$

where $P \circ S_j^{-1}$ denotes the image measure of P with respect to S_j for $j \in M$. Such a P has support the limit set J if M is finite, or the closure of J if M is infinite.

Let P be a Borel probability measure on $\mathbb{R}$ generated by the two contractive similarity mappings S_1 and S_2 associated with the probability vector $(\frac{1}{2}, \frac{1}{2})$

such that $S_1(x) = \frac{1}{3}x$ and $S_2(x) = \frac{1}{3}x + \frac{2}{3}$ for all $x \in \mathbb{R}$. Then, $P = \frac{1}{2}P \circ S_1^{-1} + \frac{1}{2}P \circ S_2^{-1}$ and it has support the classical Cantor set generated by S_1 and S_2. For this probability measure Graf and Luschgy gave a closed formula to determine the optimal sets of n-means and the nth quantization errors for all $n \geq 2$ (see [6]). Later for $n \geq 2$, L. Roychowdhury gave an induction formula to determine the optimal sets of n-means and the nth quantization errors for a probability distribution P on $\mathbb{R}$, given by $P = \frac{1}{4}P \circ S_1^{-1} + \frac{3}{4}P \circ S_2^{-1}$ which has support the Cantor set generated by S_1 and S_2, where $S_1(x) = \frac{1}{4}x$ and $S_2(x) = \frac{1}{2}x + \frac{1}{2}$ for all $x \in \mathbb{R}$ (see [13]). M. Roychowdhury (see [15]) gave an infinite extension of the result of Graf-Luschgy (see [6]). Çömez and Roychowdhury (see [2]) gave a closed formula to determine the optimal sets of n-means and the nth quantization error for a probability measure supported by a Cantor dust.

In this paper, we made an infinite extension of the work of L. Roychowdhury (see [13]). Let P be a Borel probability measure on $\mathbb{R}$ given by $P = \frac{1}{4}P \circ S_1^{-1} + \sum_{j=2}^{\infty} \frac{3}{2^{j+1}} P \circ S_j^{-1}$, i.e., P is generated by an infinite collection of similitudes $\{S_j\}_{j=1}^{\infty}$ associated with the probability vector $(\frac{1}{4}, \frac{3}{2^3}, \frac{3}{2^4}, \ldots)$ such that $S_j(x) = \frac{1}{2^{j+1}}x + 1 - \frac{1}{2^{j-1}}$ for all $x \in \mathbb{R}$, and for all $j \in \mathbb{N}$. For this probability measure, in this paper, we investigate the optimal sets of n-means and the nth quantization errors for all $n \in \mathbb{N}$. The arrangement of the paper is as follows: In Lemma 3.3 and Lemma 3.5, we obtain the optimal sets of n-means and the corresponding quantization errors for $n = 2$ and $n = 3$; Proposition 3.8, Proposition 3.13, Proposition 3.14, and Proposition 3.17 give some properties about the optimal sets of n-means and the nth quantization errors. In Theorem 3.1 we state and prove an induction formula to determine the optimal sets of n-means for all $n \geq 2$. In addition, using the induction formula we obtain some results and observations about the optimal sets of n-means which are given in Section 4; a tree diagram of the optimal sets of n-means for a certain range of n is also given.

2. Preliminaries

By a word ω over the set $\mathbb{N} = \{1, 2, 3, \ldots\}$ of natural numbers it is meant that $\omega := \omega_1\omega_2 \cdots \omega_k \in \mathbb{N}^k$ for some $k \geq 1$. Here k is called the length of the word ω and is denoted by $|\omega|$. A word of length zero is called the empty word and is denoted by $\emptyset$. Let $\mathbb{N}^*$ denote the set of all words over the alphabet $\mathbb{N}$ including the empty word $\emptyset$. For any two words $\omega := \omega_1\omega_2 \cdots \omega_k$ and $\tau := \tau_1\tau_2 \cdots \tau_m \in \mathbb{N}^*$, where $k, m \geq 1$, by $\omega\tau$ it is meant the concatenation of the words ω and τ, i.e., $\omega\tau = \omega_1\omega_2 \cdots \omega_k\tau_1\tau_2 \cdots \tau_m$. If $\omega := \omega_1\omega_2 \cdots \omega_k$, we write $\omega^- := \omega_1\omega_2 \cdots \omega_{k-1}$, where $k \geq 1$, i.e., ω^- is the word obtained from the word ω by deleting the last letter of ω. For $\omega \in \mathbb{N}^*$, by (ω, ∞) it is meant the set of all words $\omega^-(\omega_{|\omega|} + j)$, obtained by concatenation of the word ω^- with the word $\omega_{|\omega|} + j$ for $j \in \mathbb{N}$, i.e.,

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

Let $(p_j)_{j=1}^\infty$ be a probability vector such that $p_1 = \frac{1}{4}$ and $p_j = \frac{3}{2^{j+1}}$ for all $j \geq 2$. Let $\{S_j\}_{j=1}^\infty$ be an infinite collection of similitudes associated with the probability vector $(p_j)_{j=1}^\infty$ such that

$$S_j(x) = \frac{1}{2^{j+1}} x + 1 - \frac{1}{2^{j-1}}$$

for all $j \in \mathbb{N}$ and for all $x \in \mathbb{R}$. Then, as mentioned in the previous section, there exists a unique Borel probability measure P on $\mathbb{R}$ such that

$$P = \sum_{j=1}^\infty p_j P \circ S_j^{-1},$$

which has support lying in the closed interval $[0, 1]$. This paper deals with this probability measure P. For $\omega = \omega_1 \omega_2 \cdots \omega_n \in \mathbb{N}^n$, write

$$S_\omega := S_{\omega_1} \circ \cdots \circ S_{\omega_n}, \quad J_\omega := S_\omega(J), \quad s_\omega := s_{\omega_1} \cdots s_{\omega_n}, \quad p_\omega := p_{\omega_1} \cdots p_{\omega_n},$$

where $J := J_\emptyset = [0, 1]$. We also assume $p_\emptyset = 1$ and $s_\emptyset = 1$. Then, for any $\omega \in \mathbb{N}^*$, we write

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

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

Notice that for any $k \in \mathbb{N}$, $p_{(k,\infty)} = 1 - \sum_{j=1}^k p_j$, and for any word $\omega \in \mathbb{N}^*$, $p_{(\omega,\infty)} = p_{\omega^-} - \sum_{j=1}^{w_{|\omega|}} p_{\omega^- j}$. To avoid any confusion among the readers, we would like to mention that in the paper $dP(x)$ which is $P(dx)$ is identified as dP.

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

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

Proof. We know $P = \sum_{j=1}^\infty p_j P \circ S_j^{-1}$, and so by induction $P = \sum_{\omega \in \mathbb{N}^k} p_\omega P \circ S_\omega^{-1}$, and thus the lemma is yielded. ☐

Lemma 2.2. *Let X be a random variable with probability distribution P. Then, the expectation $E(X)$ and the variance $V := V(X)$ of the random variable X are given by*

$$E(X) = \frac{4}{7} \text{ and } V(X) = \frac{288}{3577} = 0.0805144.$$

Proof. Using Lemma 2.1, we have

$$\begin{aligned}E(X) &= \int x dP \\ &= \frac{1}{4}\int S_1(x)dP + \sum_{j=2}^{\infty}\frac{3}{2^{j+1}}\int S_j(x)dP \\ &= \frac{1}{16}\int x dP + \sum_{j=2}^{\infty}\frac{3}{2^{j+1}}\int\Big(\frac{1}{2^{j+1}}x + 1 - \frac{1}{2^{j-1}}\Big)dP \\ &= \frac{1}{16}E(X) + \frac{1}{16}E(X) + \frac{1}{2},\end{aligned}$$

which implies $E(X) = \frac{4}{7}$. Now,

$$\begin{aligned}E(X^2) &= \int x^2 dP \\ &= \frac{1}{4}\int(\frac{1}{4}x)^2 dP + \sum_{j=2}^{\infty}\frac{3}{2^{j+1}}\int\Big(\frac{1}{2^{j+1}}x + 1 - \frac{1}{2^{j-1}}\Big)^2 dP \\ &= \frac{1}{64}E(X^2) + \sum_{j=2}^{\infty}\frac{3}{2^{j+1}}\int\Big(\frac{1}{4^{j+1}}x^2 + \frac{2}{2^{j+1}}(1-\frac{1}{2^{j-1}})x + (1-\frac{1}{2^{j-1}})^2\Big)dP \\ &= \frac{1}{64}E(X^2) + \frac{3}{448}E(X^2) + \frac{1}{14}E(X) + \frac{5}{14} \\ &= \frac{5}{224}E(X^2) + \frac{39}{98},\end{aligned}$$

which yields $E(X^2) = \frac{208}{511}$. Thus, $V(X) = E(X^2) - (E(X))^2 = \frac{288}{3577} = 0.0805144$, which is the lemma. □

Lemma 2.3. *For any $k \geq 2$, we have*

$$E(X|X \in J_k \cup J_{k+1} \cup \cdots) = 1 - \frac{8}{7}\frac{1}{2^k}.$$

Proof. We have

$$\begin{aligned}E(X|X \in J_k \cup J_{k+1} \cup \cdots) &= \frac{1}{\sum_{j=k}^{\infty} p_j}\sum_{j=k}^{\infty} p_j S_j(\frac{4}{7}) \\ &= \frac{2^k}{3}\Big(\sum_{j=k}^{\infty}\frac{3}{2^{j+1}}(\frac{1}{2^{j+1}}\frac{4}{7} + 1 - \frac{1}{2^{j-1}})\Big),\end{aligned}$$

which after simplification yields $E(X|X \in J_k \cup J_{k+1} \cup \cdots) = 1 - \frac{8}{7}\frac{1}{2^k}$, which is the lemma. □

The following notes are in order.

Note 2.4. For $k \in \mathbb{N}$, we have $S_k(\frac{4}{7}) = \frac{1}{2^{k+1}}\frac{4}{7} + 1 - \frac{1}{2^{k-1}}$. Thus, by Lemma 2.3, for $k \in \mathbb{N}$,

$$E(X|X \in J_k \cup J_{k+1} \cup \cdots) = S_k(\frac{4}{7}) + \frac{1}{7}\frac{1}{2^{k-2}} = S_k(\frac{4}{7}) + \frac{8}{7}s_k.$$

Since for any $x_0 \in \mathbb{R}$, $\int (x-x_0)^2 dP = V(X) + (x_0 - E(X))^2$, we can deduce that the optimal set of one-mean is the expected value and the corresponding quantization error is the variance V of the random variable X. For $\omega \in \mathbb{N}^k$, $k \geq 1$, using Lemma 2.1, we have

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

Since S_j are similitudes, it is easy to see that $E(S_j(X)) = S_j(E(X))$ for $j \in \mathbb{N}$, and so by induction, $E(S_\omega(X)) = S_\omega(E(X))$ for $\omega \in \mathbb{N}^k$, $k \geq 1$.

Note 2.5. For words $\beta, \gamma, \ldots, \delta$ in $\mathbb{N}^*$, by $a(\beta, \gamma, \ldots, \delta)$ we denote the conditional expectation of the random variable X given that X is in $J_\beta \cup J_\gamma \cup \cdots \cup J_\delta$, i.e.,

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

Then, by Note 2.4, for $\omega \in \mathbb{N}^*$, we have

$$\begin{cases} a(\omega) = S_\omega(E(X)) = S_\omega(\frac{4}{7}), \text{ and} \\ a(\omega, \infty) = E(X|X \in J_{\omega^-(\omega_{|\omega|}+1)} \cup J_{\omega^-(\omega_{|\omega|}+2)} \cup \cdots) \\ \qquad = S_{\omega^-(\omega_{|\omega|}+1)}(\frac{4}{7}) + \frac{8}{7}s_{\omega^-(\omega_{|\omega|}+1)}. \end{cases} \tag{2}$$

Moreover, for any $\omega \in \mathbb{N}^*$ and for any $x_0 \in \mathbb{R}$, it is easy to see that

$$\begin{cases} \int_{J_\omega} (x-x_0)^2 dP = p_\omega \int (x-x_0)^2 d(P \circ S_\omega^{-1}) \\ \qquad = p_\omega \Big(s_\omega^2 V + (S_\omega(\frac{4}{7}) - x_0)^2\Big), \text{ and} \\ \int_{J_{(\omega,\infty)}} (x-x_0)^2 dP = \sum_{j=1}^\infty p_{\omega^-(\omega_{|\omega|}+j)} \Big(s^2_{\omega^-(\omega_{|\omega|}+j)} V + (S_{\omega^-(\omega_{|\omega|}+j)}(\frac{4}{7}) - x_0)^2\Big). \end{cases} \tag{3}$$

The expressions (2) and (3) are useful to obtain the optimal sets and the corresponding quantization errors with respect to the probability distribution P.

The following lemma plays a vital role in the paper.

Lemma 2.6. *Let P be the probability measure as defined before and let $\omega \in \mathbb{N}^k$, $k \geq 1$. Then,*

$$\int_{J_\omega} (x - a(\omega))^2 dP = p_\omega s_\omega^2 V, \text{ and}$$

$$\int_{J_{(\omega,\infty)}} (x - a(\omega,\infty))^2 dP = \begin{cases} \frac{43}{9} p_\omega s_\omega^2 V & \text{if } \omega_{|\omega|} \geq 2, \\ \frac{43}{3} p_\omega s_\omega^2 V & \text{if } \omega_{|\omega|} = 1. \end{cases}$$

Proof. In the first equation of (3) put $x_0 = a(\omega)$, and then $\int_{J_\omega} (x - a(\omega))^2 dP = p_\omega s_\omega^2 V$. In the second equation of (3), put $x_0 = a(\omega, \infty)$, and then

$$\begin{aligned} &\int_{J_{(\omega,\infty)}} (x - a(\omega,\infty))^2 dP \\ &= \sum_{j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}+j)} s^2_{\omega^-(\omega_{|\omega|}+j)} V \\ &\quad + \sum_{j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}+j)} \Big(S_{\omega^-(\omega_{|\omega|}+j)}(\frac{4}{7}) - a(\omega,\infty) \Big)^2. \end{aligned} \tag{4}$$

Putting the values of $a(\omega, \infty)$ from (2) we have

$$\begin{aligned} &S_{\omega^-(\omega_{|\omega|}+j)}(\frac{4}{7}) - a(\omega,\infty) \\ &= S_{\omega^-(\omega_{|\omega|}+j)}(\frac{4}{7}) - S_{\omega^-(\omega_{|\omega|}+1)}(\frac{4}{7}) - \frac{8}{7} s_{\omega^-(\omega_{|\omega|}+1)} \\ &= s_{\omega^-} \Big(S_{\omega_{|\omega|}+j}(\frac{4}{7}) - S_{\omega_{|\omega|}+1}(\frac{4}{7}) - \frac{8}{7} s_{\omega_{|\omega|}+1} \Big) \\ &= s_{\omega^-} \Big(\frac{1}{2^{\omega_{|\omega|}+j+1}} \frac{4}{7} - \frac{1}{2^{\omega_{|\omega|}+j-1}} - \frac{1}{2^{\omega_{|\omega|}+1+1}} \frac{4}{7} + \frac{1}{2^{\omega_{|\omega|}+1-1}} - \frac{8}{7} s_{\omega_{|\omega|}+1} \Big) \\ &= s_\omega \Big(\frac{1}{2^j} \frac{4}{7} - \frac{4}{2^j} - \frac{2}{7} + 2 - \frac{4}{7} \Big) = s_\omega \Big(\frac{8}{7} - \frac{24}{7} \frac{1}{2^j} \Big). \end{aligned}$$

Moreover, for any $j \geq 1$, $s_{\omega^-(\omega_{|\omega|}+j)} = s_\omega \frac{1}{2^j}$; and $p_{\omega^-(\omega_{|\omega|}+j)} = p_\omega \frac{1}{2^j}$ if $\omega_{|\omega|} \geq 2$, and $p_{\omega^-(\omega_{|\omega|}+j)} = p_\omega \frac{3}{2^j}$ if $\omega_{|\omega|} = 1$. Thus if $\omega_{|\omega|} \geq 2$, putting the corresponding values and making some simplification, we obtain

$$\sum_{j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}+j)} s^2_{\omega^-(\omega_{|\omega|}+j)} V = \frac{1}{7} p_\omega s_\omega^2 V \text{ and}$$

$$\begin{aligned} \sum_{j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}+j)} \Big(S_{\omega^-(\omega_{|\omega|}+j)}(\frac{4}{7}) - a(\omega,\infty) \Big)^2 &= p_\omega s_\omega^2 \sum_{j=1}^{\infty} \frac{1}{2^j} \Big(\frac{8}{7} - \frac{24}{7} \frac{1}{2^j} \Big)^2 \\ &= p_\omega s_\omega^2 V \frac{292}{63}, \end{aligned}$$

and then (4) yields $\int_{J_{(\omega,\infty)}}(x-a(\omega,\infty))^2dP=\frac{43}{9}p_\omega s_\omega^2V$. Similarly, if $\omega_{|\omega|}=1$, one can obtain $\int_{J_{(\omega,\infty)}}(x-a(\omega,\infty))^2dP=\frac{43}{3}p_\omega s_\omega^2V$. Thus, the lemma is yielded. □

Notation 2.7. For any $\omega\in\mathbb{N}^k$, $k\geq 1$, set

$$E(a(\omega)):=\int_{J_\omega}(x-a(\omega))^2dP \text{ and}$$
$$E(a(\omega,\infty)):=\int_{J_{(\omega,\infty)}}(x-a(\omega,\infty))^2dP. \tag{5}$$

Let us now prove the following lemma.

Lemma 2.8. *For any two nonempty words $\omega,\tau\in\mathbb{N}^*$ if $p_\omega=p_\tau$, then $s_\omega=s_\tau$.*

Proof. To prove the lemma, let us define a function c as follows:

$c:\mathbb{N}^*\setminus\{\emptyset\}\to\mathbb{N}\cup\{0\}$ such that $c(\omega)=\text{card}(\{\omega_i:\omega_i\neq 1,\ 1\leq i\leq|\omega|\})$.

Let $\omega,\tau\in\mathbb{N}^*$ with $\omega=\omega_1\omega_2\cdots\omega_k$ and $\tau=\tau_1\tau_2\cdots\tau_m$ for some $k,m\geq 1$. Then, $p_\omega=p_\tau$ implies

$$\frac{3^{c(\omega)}}{2^{\omega_1+\omega_2+\cdots+\omega_k+k}}=\frac{3^{c(\tau)}}{2^{\tau_1+\tau_2+\cdots+\tau_m+m}}$$

yielding $3^{c(\omega)-c(\tau)}=2^{(\omega_1+\omega_2+\cdots+\omega_k+k)-(\tau_1+\tau_2+\cdots+\tau_m+m)}$ and so, $c(\omega)=c(\tau)$ and $\omega_1+\omega_2+\cdots+\omega_k+k=\tau_1+\tau_2+\cdots+\tau_m+m$. Then,

$$s_\omega=\frac{1}{2^{\omega_1+\omega_2+\cdots+\omega_k+k}}=\frac{1}{2^{\tau_1+\tau_2+\cdots+\tau_m+m}}=s_\tau,$$

which is the lemma. □

In the next section we state and prove the main result of the paper.

3. Main result

The following theorem gives the main result of the paper.

Theorem 3.1. *For any $n\geq 2$, let $\alpha_n:=\{a(i):1\leq i\leq n\}$ be an optimal set of n-means, i.e., $\alpha_n\in\mathcal{C}_n:=\mathcal{C}_n(P)$. For $\omega\in\mathbb{N}^k$, $k\geq 1$, let $E(a(\omega))$ and $E(a(\omega,\infty))$ be defined by (5). Set*

$$\tilde{E}(a(i)):=\begin{cases}E(a(\omega)) \text{ if } a(i)=a(\omega) \text{ for some } \omega\in\mathbb{N}^*,\\ E(a(\omega,\infty)) \text{ if } a(i)=a(\omega,\infty) \text{ for some } \omega\in\mathbb{N}^*,\end{cases}$$

and $W(\alpha_n):=\{a(j):a(j)\in\alpha_n \text{ and } \tilde{E}(a(j))\geq\tilde{E}(a(i)) \text{ for all } 1\leq i\leq n\}$. Take any $a(j)\in W(\alpha_n)$, and write

$$\alpha_{n+1}(a(j)) :=\begin{cases}(\alpha_n\setminus\{a(j)\})\cup\{a(\omega^-(\omega_{|\omega|}+1)),\ a(\omega^-(\omega_{|\omega|}+1),\infty)\} \text{ if } a(j)=a(\omega,\infty),\\ (\alpha_n\setminus\{a(j)\})\cup\{a(\omega1),\ a(\omega1,\infty)\} \text{ if } a(j)=a(\omega).\end{cases}$$

Then, $\alpha_{n+1}(a(j))$ is an optimal set of $(n+1)$-means, and the number of such sets is given by

$$card\Big(\bigcup_{\alpha_n\in\mathcal{C}_n}\{\alpha_{n+1}(a(j)) : a(j)\in W(\alpha_n)\}\Big).$$

Remark 3.2. Once an optimal set of n-means is known, by using (3) and Lemma 2.6, the corresponding quantization error can easily be calculated.

To prove Theorem 3.1 we need some basic lemmas and propositions.

Lemma 3.3. *Let $\alpha=\{a_1,a_2\}$ be an optimal set of two-means, $a_1<a_2$. Then, $a_1=a(1)=\frac{1}{7}$, $a_2=a(1,\infty)=\frac{5}{7}$ and the quantization error is $V_2=\frac{69}{3577}=0.0192899$.*

Proof. Let us first consider the two-point set β given by $\beta=\{\frac{1}{7},\frac{5}{7}\}$. Since $S_1(1)<\frac{1}{2}(\frac{1}{7}+\frac{5}{7})<S_2(0)$, by Lemma 2.6, we have

$$\begin{aligned}\int\min_{b\in\beta}(x-b)^2dP&=\int_{J_1}(x-\frac{1}{7})^2dP+\int_{J_{(1,\infty)}}(x-\frac{5}{7})^2dP\\&=p_1s_1^2(1+\frac{43}{3})V=\frac{69}{3577}=0.0192899.\end{aligned}$$

Since V_2 is the quantization error for two-means, we have $V_2\leq 0.0192899$. Let $\alpha=\{a_1,a_2\}$ be an optimal set of two-means, $a_1<a_2$. Since a_1 and a_2 are the centroids of their own Voronoi regions, we have $0<a_1<a_2<1$. Suppose that $a_2\leq\frac{5}{8}$. Then,

$$V_2\geq\int_{J_3\cup J_4\cup J_5\cup J_6}(x-\frac{5}{8})^2dP=\frac{647055}{33488896}=0.0193215>V_2,$$

which leads to a contradiction. So, we can assume that $a_2>\frac{5}{8}$ implying $\frac{1}{2}(a_1+a_2)\geq\frac{1}{2}(0+\frac{5}{8})=\frac{5}{16}>\frac{1}{4}$. Thus, we see that the Voronoi region of a_2 does not contain any point from J_1, and $a_1\geq a(1)=\frac{1}{7}$. Suppose that $a_1\geq\frac{7}{16}$. Then, using (3), we have

$$\begin{aligned}V_2&\geq\int_{J_1}(x-a_1)^2dP\geq\int_{J_1}(x-\frac{7}{16})^2dP\\&=p_1\Big(s_1^2V+(S_1(\frac{4}{7})-\frac{7}{16})^2\Big)\\&=\frac{12015}{523264}=0.0229616>V_2\end{aligned}$$

which is a contradiction, and so $\frac{1}{7}\leq a_1<\frac{7}{16}$. We now show that $\frac{1}{2}(a_1+a_2)\leq\frac{1}{2}$. For the sake of contradiction assume that $\frac{1}{2}(a_1+a_2)>\frac{1}{2}$. Then, if $\frac{1}{2}(a_1+a_2)\geq\frac{5}{8}$, we have $a_1\geq E(X:X\in J_1\cup J_2)=\frac{2}{5}$, yielding

$$V_2\geq\int_{J_1\cup J_2}(x-\frac{2}{5})^2dP=\frac{171}{5840}=0.0292808>V_2,$$

which is a contradiction. Next, assume that $S_{2\sigma 1}(1) \leq \frac{1}{2}(a_1+a_2) \leq S_{2\sigma 2}(0)$ for some $\sigma \in \mathbb{N}^*$. For definiteness sake, take $\sigma = 1$, and so $S_{211}(1) \leq \frac{1}{2}(a_1+a_2) \leq S_{212}(0)$. Then, $a_1 = E(X : X \in J_1 \cup J_{211})$ and $a_2 = E(X : X \in J_{(211,\infty)} \cup J_{(21,\infty)} \cup J_{(2,\infty)})$ yielding

$$a_1 = \frac{P(J_1)S_1(\frac{4}{7}) + P(J_{211})S_{211}(\frac{4}{7})}{P(J_1)+P(J_{211})} = \frac{1363}{7840},$$

and

$$a_2 = \frac{p_{(211,\infty)}a(211,\infty) + p_{(21,\infty)}a(21,\infty) + p_{(2,\infty)}a(2,\infty)}{p_{(211,\infty)} + p_{(21,\infty)} + p_{(2,\infty)}} = \frac{5007}{6944},$$

where $p_{(211,\infty)} = p_{21}-p_{211}$, $p_{(21,\infty)} = p_2-p_{21}$, $p_{(2,\infty)} = 1-p_1-p_2$, $a(211,\infty) = S_{212}(\frac{4}{7}) + \frac{8}{7}s_{212}$, $a(21,\infty) = S_{22}(\frac{4}{7}) + \frac{8}{7}s_{22}$, and $a(2,\infty) = S_3(\frac{4}{7}) + \frac{8}{7}s_3$. Thus,

$$\begin{aligned} V_2 &\geq \int_{J_1 \cup J_{211}} (x - \frac{1363}{7840})^2 dP + \int_A (x - \frac{5007}{6944})^2 dP \\ &= \frac{648995235322779}{32296112614277120} = 0.0200952 > V_2, \end{aligned}$$

where $A = J_{212} \cup J_{213} \cup J_{22} \cup J_{23} \cup J_{24} \cup J_{25} \cup J_3 \cup J_4 \cup J_5 \cup J_6 \cup J_7 \cup J_8 \cup J_9 \cup J_{10}$, which gives a contradiction. Similarly, we can show that for any other choice of $\sigma \in \mathbb{N}^*$, the assumption $\frac{1}{2}(a_1+a_2) > \frac{1}{2}$ will give a contradiction. Thus, we have $\frac{1}{2}(a_1+a_2) \leq \frac{1}{2}$ implying $a_1 \leq a(1) = \frac{1}{7}$. Again, we have seen $a_1 \geq \frac{1}{7}$. Thus, we deduce that $a_1 = \frac{1}{7}$ and the Voronoi region of a_2 does not contain any point from J_1, i.e., $a_2 = a(1,\infty) = \frac{5}{7}$, and the corresponding quantization error is $V_2 = \frac{69}{3577} = 0.0192899$. This completes the proof of the lemma. □

Using the technique of Lemma 3.3, the following corollary can be proved.

Corollary 3.4. *For any $\omega \in \mathbb{N}^*$, the set $\{a(\omega 1), a(\omega 1, \infty)\}$ forms a unique optimal set two-means for the conditional measure of P on J_ω, and the set $\{a(\omega^-(\omega_{|\omega|}+1)), a(\omega^-(\omega_{|\omega|}+1), \infty)\}$ forms a unique optimal set of two-means for the conditional measure of P on $J_{(\omega,\infty)}$.*

Lemma 3.5. *Let α be an optimal set of three-means. Then,*

$$\alpha = \{a(1), a(2), a(2,\infty)\} = \{\frac{1}{7}, \frac{4}{7}, \frac{6}{7}\}$$

and the quantization error is $V_3 = \frac{57}{14308} = 0.00398379$.

Proof. Let us first consider a three-point set β given by $\beta := \{\frac{1}{7}, \frac{4}{7}, \frac{6}{7}\}$. Since $J_1 \subset M(\frac{1}{7}|\beta)$, $J_2 \subset M(\frac{4}{7}|\beta)$ and $J_{(2,\infty)} \subset M(\frac{6}{7}|\beta)$, by Lemma 2.6, we have

$$\begin{aligned} \int \min_{b \in \beta}(x-b)^2 dP &= \int_{J_1} (x-\frac{1}{7})^2 dP + \int_{J_2} (x - \frac{4}{7})^2 dP + \int_{J_{(2,\infty)}} (x-\frac{6}{7})^2 dP \\ &= p_1 s_1^2 V + p_2 s_2^2 V(1 + \frac{43}{9}) = \frac{57}{14308} = 0.00398379. \end{aligned}$$

Since V_3 is the quantization error for three-means, we have $V_3 \leq \frac{57}{14308} = 0.00398379$. Let α be an optimal set of three-means with $\alpha = \{a_1, a_2, a_3\}$, where $a_1 < a_2 < a_3$. Since the optimal points are the centroids of their own Voronoi regions, we have $0 < a_1 < a_2 < a_3 < 1$. If $a_1 > \frac{1}{4}$, then

$$V_3 \geq \int_{J_1} (x-a_1)^2 dP \geq \int_{J_1} (x-\frac{1}{4})^2 dP = \frac{135}{32704} = 0.00412794 > V_3,$$

which gives a contradiction, and so $a_1 \leq \frac{1}{4}$. If $a_3 < \frac{25}{32} = S_{32}(0)$, using (3), we see that

$$V_3 \geq \int_{J_{32} \cup J_{33} \cup \bigcup_{j=4}^{8} J_j} (x-\frac{25}{32})^2 dP = \frac{8764935}{2143289344} = 0.00408948 > V_3,$$

which leads to a contradiction, and so $\frac{25}{32} \leq a_3$. Suppose that $a_2 \leq \frac{1}{2} - \frac{1}{32} = \frac{15}{32}$. Then, as $\frac{1}{2}(\frac{15}{32} + \frac{25}{32}) = \frac{5}{8} = S_2(1)$, we have

$$V_3 \geq \int_{J_2} (x-\frac{15}{32})^2 dP = \frac{18525}{4186112} = 0.00442535 > V_3,$$

which is a contradiction. Assume that $\frac{15}{32} \leq a_2 < \frac{1}{2}$. Then, $\frac{1}{2}(a_1 + a_2) < \frac{1}{4}$ implying $a_1 \leq \frac{1}{2} - a_2 \leq \frac{1}{2} - \frac{15}{32} = \frac{1}{32} < \frac{4}{32} = S_{12}(0)$. Again $\frac{1}{2}(\frac{1}{2} + \frac{25}{32}) = \frac{41}{32} > \frac{5}{8} = S_2(1)$. Thus, we have

$$\begin{aligned} V_3 &\geq \int_{J_{12} \cup J_{13}} (x-\frac{1}{32})^2 dP + \int_{J_2} (x-\frac{1}{2})^2 dP \\ &= \frac{162087}{33488896} = 0.00484002 > V_3, \end{aligned}$$

which is a contradiction. So, we can assume that $\frac{1}{2} \leq a_2$. Suppose that $\frac{5}{8} + \frac{1}{32} = \frac{21}{32} \leq a_2$. Then, as $\frac{1}{4} < \frac{1}{2}(a(1) + \frac{21}{32}) < \frac{1}{2}$, we have

$$\begin{aligned} V_3 &\geq \int_{J_1} (x-a(1))^2 dP + \int_{J_2} (x-\frac{21}{32})^2 dP \\ &= \frac{129747}{29302784} = 0.0044278 > V_3, \end{aligned}$$

which yields a contradiction. Next, suppose that $\frac{5}{8} < a_2 \leq \frac{5}{8} + \frac{1}{32} = \frac{21}{32}$. Then, $\frac{1}{4} < \frac{1}{2}(a(1) + \frac{5}{8}) < \frac{1}{2}$. Moreover, $\frac{1}{2}(a_2 + a_3) > \frac{3}{4}$ implying $a_3 > \frac{3}{2} - a_2 \geq \frac{3}{2} - \frac{21}{32} = \frac{27}{32} > \frac{13}{16} = S_3(1)$ leading to the following two cases:

Case A. $\frac{27}{32} < a_3 \leq \frac{113}{128} = S_{41}(1)$.

Then, $\frac{1}{2}(\frac{21}{32} + \frac{27}{32}) = \frac{3}{4} = S_3(0)$, and so

$$\begin{aligned} V_3 &\geq \int_{J_1} (x-a(1))^2 dP + \int_{J_2} (x-\frac{5}{8})^2 dP \\ &\quad + \int_{J_3} (x-\frac{27}{32})^2 dP + \int_{J_5 \cup J_6 \cup J_7} (x-\frac{113}{128})^2 dP \\ &= \frac{60087981}{15003025408} = 0.00400506 > V_3, \end{aligned}$$

which gives a contradiction.

Case B. $S_{41}(1) = \frac{113}{128} \leq a_3$.

Then, $S_{31}(1) < \frac{1}{2}(\frac{21}{32} + \frac{113}{128}) < S_{32}(0)$, and so

$$\begin{aligned} V_3 &\geq \int_{J_1} (x - a(1))^2 dP + \int_{J_2} (x - \frac{5}{8})^2 dP \\ &\quad + \int_{J_{31}} (x - \frac{21}{32})^2 dP + \int_{J_{32} \cup J_{33}} (x - \frac{113}{128})^2 dP \\ &= \frac{63174099}{15003025408} = 0.00421076 > V_3, \end{aligned}$$

which leads to a contradiction.

Therefore, $\frac{1}{2} \leq a_2 \leq \frac{5}{8}$. Suppose that $S_{23}(0) = \frac{19}{32} \leq a_2 \leq \frac{5}{8}$. Then, the Voronoi region of a_2 does not contain any point from J_1, and $\frac{1}{2}(a_2 + a_3) > \frac{3}{4}$ implying $a_3 > \frac{3}{2} - a_2 \geq \frac{3}{2} - \frac{5}{8} = \frac{7}{8}$, otherwise the quantization error can strictly be reduced by moving the point a_2 to $a(2) = \frac{4}{7}$. Thus, we have

$$\begin{aligned} \min_{\frac{19}{32} \leq a_2 \leq \frac{5}{8}} \int_{J_2} (x - a_2)^2 dP &= p_2 \Big(s_2^2 V + \min_{\frac{19}{32} \leq a_2 \leq \frac{5}{8}} (S_2(\frac{4}{7}) - a_2)^2 \Big) \\ &= p_2 \Big(s_2^2 V + (a(2) - \frac{19}{32})^2 \Big) = \frac{2757}{4186112}. \end{aligned}$$

The following two cases can arise:

Case I. $\frac{7}{8} < a_3 \leq S_{42}(0) = \frac{57}{64}$.

Then, $\frac{1}{2}(\frac{5}{8} + \frac{7}{8}) = \frac{3}{4} = S_3(0)$. Write $A := J_{42} \cup J_{43} \cup \bigcup_{j=5}^{10} J_j$, and so

$$\begin{aligned} V_3 &\geq \int_{J_1} (x - a(1))^2 dP + \min_{\frac{19}{32} \leq a_2 \leq \frac{5}{8}} \int_{J_2} (x - a_2)^2 dP \\ &\quad + \int_{J_3} (x - \frac{7}{8})^2 dP + \int_A (x - \frac{57}{64})^2 dP \\ &= \frac{3839362137}{960193626112} = 0.00399853 > V_3, \end{aligned}$$

which gives a contradiction.

Case II. $S_{42}(0) = \frac{57}{64} < a_3$.

Then, $S_{311}(1) = \frac{193}{256} < \frac{1}{2}(\frac{5}{8} + \frac{57}{64}) = \frac{97}{128} = S_{312}(0)$. Write $A := \bigcup_{j=2}^{10} J_{31j} \cup \bigcup_{j=2}^{10} J_{3j} \cup J_{41}$. Thus,

$$\begin{aligned} V_3 &\geq \int_{J_1} (x - a(1))^2 dP + \min_{\frac{19}{32} \leq a_2 \leq \frac{5}{8}} \int_{J_2} (x - a_2)^2 dP \\ &\quad + \int_{J_{311}} (x - \frac{5}{8})^2 dP + \int_A (x - \frac{57}{64})^2 dP \end{aligned}$$

$$= \frac{1008051842887707}{251708997923504128} = 0.00400483 > V_3,$$

which leads to a contradiction.

Therefore, we can assume that $\frac{1}{2} \leq a_2 \leq \frac{19}{32} = S_{23}(0)$. Again, we have seen that $\frac{25}{32} \leq a_3 \leq 1$. Then, notice that the Voronoi region of a_2 does not contain any point from J_1. Moreover, $\frac{41}{64} = \frac{1}{2}(\frac{1}{2} + \frac{25}{32}) \leq \frac{1}{2}(a_2 + a_3) \leq \frac{1}{2}(\frac{19}{32} + 1) = \frac{51}{64}$ implying that the Voronoi region of a_3 does not contain any point from J_2. Suppose that the Voronoi region of a_2 contains points from $J_{(2,\infty)}$. Then, $\frac{1}{2}(a_2 + a_3) > \frac{3}{4}$, which implies $a_3 > \frac{3}{2} - a_2 \geq \frac{3}{2} - \frac{19}{32} = \frac{29}{32} = S_4(1)$. Moreover,

$$\min_{\frac{1}{2} \leq a_2 \leq \frac{19}{32}} \int_{J_2} (x - a_2)^2 dP = \int_{J_2} (x - a(2))^2 dP = p_2 s_2^2 V.$$

Thus, we see that

$$V_3 \geq \int_{J_1} (x - a(1))^2 dP + p_2 s_2^2 V + \int_{J_3 \cup J_4} (x - \frac{29}{32})^2 dP$$
$$= \frac{531801}{117211136} = 0.00453712 > V_3,$$

which gives a contradiction. Therefore, we can assume that the Voronoi region of a_2 does not contain any point from $J_{(2,\infty)}$. Thus, we have proved that $J_1 \subset M(a_1|\alpha)$, $J_2 \subset M(a_2|\alpha)$, and $J_3 \subset M(a_3|\alpha)$ yielding $a_1 = a(1) = \frac{1}{7}$, $a_2 = a(2) = \frac{4}{7}$, and $a_3 = a(2,\infty) = \frac{6}{7}$, and the corresponding quantization error is $V_3 = \frac{57}{14308} = 0.00398379$ (see Figure 1). Thus, the proof of the lemma is complete. □

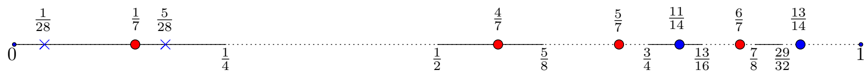

FIGURE 1. Optimal sets: of one-mean is $\{\frac{4}{7}\}$; of two-means is $\{\frac{1}{7}, \frac{5}{7}\}$; of three-means is $\{\frac{1}{7}, \frac{4}{7}, \frac{6}{7}\}$; of four-means is $\{\frac{1}{7}, \frac{4}{7}, \frac{11}{14}, \frac{13}{14}\}$; of five-means is $\{\frac{1}{28}, \frac{5}{28}, \frac{4}{7}, \frac{11}{14}, \frac{13}{14}\}$.

We need the following two lemmas to prove Proposition 3.8.

Lemma 3.6. *Let α_4 be an optimal set of four-means. Then, $\alpha_4 \cap J_1 \neq \emptyset$ and $\alpha_4 \cap J_{(1,\infty)} \neq \emptyset$, and α_4 does not contain any point from the open interval $(\frac{1}{4}, \frac{1}{2})$. Moreover, the Voronoi region of any point in $\alpha_4 \cap J_1$ does not contain any point from $J_{(1,\infty)}$ and the Voronoi region of any point in $\alpha_4 \cap J_{(1,\infty)}$ does not contain any point from J_1.*

Proof. Let $\alpha_4 := \{0 < a_1 < a_2 < a_3 < a_4 < 1\}$ be an optimal set of four-means. Consider the set $\beta := \{a(1), a(2), a(3), a(3,\infty)\}$ of four points. Then,

$$\int \min_{a \in \beta} (x - a)^2 dP = p_1 s_1^2 V + p_2 s_2^2 V + p_3 s_3^3 V (1 + \frac{43}{9}) = \frac{237}{114464} = 0.00207052.$$

Since V_4 is the quantization error for four-means, we have $V_4 \leq 0.00207052$. If $a_1 \geq \frac{13}{64} = S_{13}(1)$, we have

$$V_4 \geq \int_{J_{11}\cup J_{12}} (x - \frac{13}{64})^2 dP = \frac{20277}{9568256} = 0.00211919 > V_4,$$

which is a contradiction. So, we can assume that $a_1 \leq \frac{13}{64}$. Then, the Voronoi region of a_1 does not contain any point from $J_{(1,\infty)}$. If it does, then $\frac{1}{2}(a_1+a_2) > \frac{1}{2}$ implies $a_2 \geq 1 - a_1 \geq 1 - \frac{13}{64} = \frac{51}{64}$ which is a contradiction as

$$V_4 \geq \int_{J_1} (x - a(1))^2 dP + \int_{J_2} (x - \frac{51}{64}) dP = \frac{2436771}{117211136} = 0.0207896 > V_4.$$

If $a_4 \leq \frac{53}{64}$, then

$$V_4 \geq \int_{\bigcup_{j=4}^{10} J_j} (x - \frac{53}{64})^2 dP = \frac{292246431}{137170518016} = 0.00213053 > V_4,$$

which is a contradiction, and so $\frac{53}{64} < a_4$. If $a_2 \leq \frac{1}{4}$, then

$$\begin{aligned} V_4 &\geq \int_{J_2} (x - a(2))^2 dP + \int_{J_{(2,\infty)}} (x - a(2,\infty))^2 dP \\ &= (1 + \frac{43}{9}) p_2 s_2^2 V = \frac{39}{14308} = 0.00272575 > V_4, \end{aligned}$$

which gives a contradiction. So, we can assume that $\frac{1}{4} < a_2$. Suppose that $\frac{1}{4} < a_2 \leq \frac{3}{8}$. Then, $\frac{1}{2}(a_2 + a_3) > \frac{1}{2}$ yielding $a_3 > 1 - a_2 \geq 1 - \frac{3}{8} = \frac{5}{8}$. Thus, the following two cases can arise:

Case 1. $\frac{5}{8} < a_3 \leq \frac{43}{64}$.

Then, as $\frac{53}{64} < a_4$ and $\frac{1}{2}(\frac{43}{64} + \frac{53}{64}) = \frac{3}{4}$, we have

$$\begin{aligned} V_4 &\geq \int_{J_2} (x - \frac{5}{8})^2 dP + \int_{J_3} (x - \frac{53}{64})^2 dP + \int_{J_{(3,\infty)}} (x - a(3,\infty))^2 dP \\ &= \frac{521811}{234422272} = 0.00222594 > V_4, \end{aligned}$$

which is a contradiction.

Case 2. $\frac{43}{64} \leq a_3$.

Then, as $S_{212}(1) < \frac{1}{2}(\frac{3}{8} + \frac{43}{64}) = \frac{67}{128} = S_{213}(0)$, we have

$$\begin{aligned} V_4 &\geq \int_{J_{211}\cup J_{212}} (x - \frac{3}{8})^2 dP + \int_{J_{22}\cup J_{23}} (x - \frac{43}{64})^2 dP \\ &= \frac{6099}{2093056} = 0.00291392 > V_4, \end{aligned}$$

which leads to a contradiction.

Thus, a contradiction arises to our assumption $\frac{1}{4} < a_2 \leq \frac{3}{8}$. Suppose that $\frac{3}{8} \leq a_2 < \frac{1}{2}$. Then, $\frac{1}{2}(a_1 + a_2) < \frac{1}{4}$ implying $a_1 \leq \frac{1}{2} - a_2 \leq \frac{1}{2} - \frac{3}{8} = \frac{1}{8} < a(1)$, and

$$\min_{\{a_1 < \frac{1}{8} < \frac{3}{8} \leq a_2\}} \int_{J_1} \min_{a \in \{a_1, a_2\}} (x-a)^2 dP \geq \int_{J_1} (x - a(1))^2 dP = \frac{9}{7154}.$$

Since $\frac{53}{64} < a_4$, the following three cases can arise:

Case I. $a_3 \leq \frac{43}{64}$ and $\frac{53}{64} < a_4 \leq \frac{7}{8}$.
Then, as $\frac{1}{2}(\frac{43}{64} + \frac{53}{64}) = \frac{3}{4}$, we have

$$\begin{aligned} V_4 &\geq \int_{J_1} (x - a(1))^2 dP + \int_{J_3} (x - \frac{53}{64})^2 dP + \int_{J_4 \cup J_5 \cup J_6} (x - \frac{7}{8})^2 dP \\ &= \frac{126459}{58605568} = 0.0021578 > V_4, \end{aligned}$$

which gives a contradiction.

Case II. $a_3 \leq \frac{43}{64}$ and $\frac{7}{8} \leq a_4$.
Then, as $S_{31}(1) < \frac{1}{2}(\frac{43}{64} + \frac{7}{8}) < S_{32}(0)$,

$$\begin{aligned} V_4 &\geq \int_{J_1} (x - a(1))^2 dP + \int_{J_{31}} (x - \frac{43}{64})^2 dP + \int_{J_{32} \cup J_{33}} (x - \frac{7}{8})^2 dP \\ &= \frac{4458897}{1875378176} = 0.0023776 > V_4, \end{aligned}$$

which leads to a contradiction.

Case III. $\frac{43}{64} \leq a_3$.
Then, $S_{22}(1) < \frac{1}{2}(\frac{1}{2} + \frac{43}{64}) < S_{23}(0)$ yielding

$$\begin{aligned} V_4 &\geq \int_{J_1} (x - a(1))^2 dP + \int_{J_{21} \cup J_{22}} (x - \frac{1}{2})^2 dP + \int_{J_{23}} (x - \frac{43}{64})^2 dP \\ &= \frac{4496025}{1875378176} = 0.0023974 > V_4, \end{aligned}$$

which is a contradiction.

Thus, a contradiction arises to our assumption $\frac{3}{8} \leq a_2 < \frac{1}{2}$, and so we can assume $\frac{1}{2} \leq a_2$. Now, notice that $\frac{1}{2}(a_1 + a_2) \geq \frac{1}{2}(0 + \frac{1}{2}) = \frac{1}{4}$ yielding the fact that the Voronoi region of any point in $\alpha_4 \cap J_{(1,\infty)}$ does not contain any point from J_1. Moreover, we proved $a_1 < \frac{1}{4}$ and the Voronoi region of any point in $\alpha_4 \cap J_1$ does not contain any point from $J_{(1,\infty)}$. Thus, the proof of the lemma is complete. □

Lemma 3.7. *Let α_5 be an optimal set of five-means. Then, $\alpha_5 \cap J_1 \neq \emptyset$, $\alpha_5 \cap J_{(1,\infty)} \neq \emptyset$, and α_5 does not contain any point from the open interval $(\frac{1}{4}, \frac{1}{2})$. Moreover, the Voronoi region of any point in $\alpha_5 \cap J_1$ does not contain any point from $J_{(1,\infty)}$ and the Voronoi region of any point in $\alpha_5 \cap J_{(1,\infty)}$ does not contain any point from J_1.*

Proof. Let $\alpha_5 := \{0 < a_1 < a_2 < a_3 < a_4 < a_5 < 1\}$ be an optimal set of five-means. Consider the set $\beta := \{a(11), a(11,\infty), a(2), a(3), a(3,\infty)\}$ of five points. Then,

$$\int \min_{a\in\beta}(x-a)^2 dP = p_{11}s_{11}^2 V(1+\frac{43}{3}) + p_2 s_2^2 V + p_3 s_3^3 V(1+\frac{43}{9})$$
$$= \frac{255}{228928} = 0.00111389.$$

Since V_5 is the quantization error for five-means, we have $V_5 \leq 0.00111389$. If $a_5 \leq \frac{6}{7}$, then

$$V_5 \geq \int_{\bigcup_{j=4}^{10} J_j} (x-\frac{6}{7})^2 dP = \frac{1160604105}{960193626112} = 0.00120872 > V_5,$$

which is a contradiction, and so $\frac{6}{7} < a_5$. Suppose that $a_4 \leq \frac{11}{16}$. Consider the following two cases:

Case 1. $\frac{6}{7} \leq a_5 < \frac{7}{8}$.

Then, $S_{31}(1) < \frac{1}{2}(\frac{11}{16} + \frac{6}{7}) < \frac{25}{32} = S_{32}(0)$, yielding

$$V_5 \geq \int_{J_{31}} (x-\frac{11}{16})^2 dP + \int_{J_{32}\cup J_{33}} (x-\frac{6}{7})^2 dP + \int_{\bigcup_{j=4}^{6} J_j} (x-\frac{7}{8})^2 dP$$
$$= \frac{2290131}{1875378176} = 0.00122116 > V_5,$$

which leads to a contradiction.

Case 2. $\frac{7}{8} \leq a_5$.

Then, $S_{31}(1) < \frac{1}{2}(\frac{11}{16} + \frac{7}{8}) = \frac{25}{32} = S_{32}(0)$, yielding

$$V_5 \geq \int_{J_{31}} (x-\frac{11}{16})^2 dP + \int_{\bigcup_{j=2}^{10} J_{3j}} (x-\frac{7}{8})^2 dP$$
$$= \frac{651896011533}{561850441793536} = 0.00116027 > V_5,$$

which is a contradiction.

Hence, we can assume that $\frac{11}{16} < a_4$. If $a_3 \leq \frac{1}{4}$, then

$$V_5 \geq \int_{J_2} (x-a(2))^2 dP + \int_{J_{(2,\infty)}} (x-a(2,\infty))^2 dP$$
$$= (1+\frac{43}{9})p_2 s_2^2 V = \frac{39}{14308} = 0.00272575 > V_5,$$

which gives a contradiction. So, we can assume that $\frac{1}{4} < a_3$. Suppose that $\frac{1}{4} < a_3 < \frac{1}{2}$. The following two cases can arise:

Case (i). $\frac{1}{4} < a_3 \leq \frac{3}{8}$.

Then, $\frac{1}{2}(a_3+a_4) > \frac{1}{2}$ implying $a_4 > 1-a_3 \geq 1-\frac{3}{8} = \frac{5}{8}$, and so

$$V_5 \geq \int_{J_2} (x-\frac{5}{8})^2 dP = \frac{405}{261632} = 0.00154798 > V_5,$$

which is a contradiction.

Case (ii). $\frac{3}{8} \leq a_3 < \frac{1}{2}$.

Then, $\frac{1}{2}(a_2+a_3) < \frac{1}{4}$ implying $a_2 < 1-a_3 \leq \frac{1}{2}-\frac{3}{8} = \frac{1}{8}$. Moreover, as $\frac{11}{16} < a_4$, we have $S_{22}(1) < \frac{1}{2}(\frac{1}{2}+\frac{11}{16}) = \frac{19}{32} = S_{23}(0)$, and so

$$\begin{aligned} V_5 &\geq \int_{J_{12}} (x-\frac{1}{8})^2 dP + \int_{J_{21}\cup J_{22}} (x-\frac{1}{2})^2 dP + \int_{J_{23}} (x-\frac{11}{16})^2 dP \\ &= \frac{45399}{33488896} = 0.00135564 > V_5, \end{aligned}$$

which yields a contradiction.

Hence, we can assume that $\frac{1}{2} \leq a_3$. If $\frac{3}{8} \leq a_2$, then

$$\begin{aligned} V_5 &\geq \min_{\{a_1 < \frac{1}{8} < \frac{3}{8} \leq a_2\}} \int_{J_1} \min_{a \in \{a_1, a_2\}} (x-a)^2 dP \\ &\geq \int_{J_1} (x-a(1))^2 dP = \frac{9}{7154} = 0.00125804 > V_5, \end{aligned}$$

which is a contradiction. Suppose that $\frac{1}{4} < a_2 \leq \frac{3}{8}$. Then, $\frac{1}{2}(a_2+a_3) > \frac{1}{2}$ implying $a_3 > 1-a_2 \geq 1-\frac{3}{8} = \frac{5}{8}$, which yields

$$V_5 \geq \int_{J_2} (x-\frac{5}{8})^2 dP = \frac{405}{261632} = 0.00154798 > V_5,$$

leading to a contradiction. So, we can assume that $a_2 \leq \frac{1}{4}$. Thus, we have proved that $a_2 \leq \frac{1}{4}$ and $\frac{1}{2} \leq a_3$, yielding the fact that $\alpha_5 \cap J_1 \neq \emptyset$, $\alpha_5 \cap J_{(1,\infty)} \neq \emptyset$, and α_5 does not contain any point from the open interval $(\frac{1}{4}, \frac{1}{2})$. Since $\frac{1}{2}(a_2+a_3) \geq \frac{1}{2}(0+\frac{1}{2}) = \frac{1}{4}$, the Voronoi region of any point in $\alpha_5 \cap J_{(1,\infty)}$ does not contain any point from J_1. If the Voronoi region of a_2 contains points from $J_{(1,\infty)}$, then $\frac{1}{2}(a_2+a_3) > \frac{1}{2}$ implying $a_3 > 1-a_2 \geq 1-\frac{1}{4} = \frac{3}{4}$, and so

$$V_5 \geq \int_{J_2} (x-\frac{3}{4})^2 dP = \frac{813}{65408} = 0.0124297 > V_5,$$

which gives a contradiction. Thus, the proof of the lemma is complete. ☐

Proposition 3.8. *Let α_n be an optimal set of n-means for $n \geq 2$. Then, $\alpha_n \cap J_1 \neq \emptyset$ and $\alpha_n \cap J_{(1,\infty)} \neq \emptyset$, and α_n does not contain any point from the open interval $(\frac{1}{4}, \frac{1}{2})$. Moreover, the Voronoi region of any point in $\alpha_n \cap J_1$ does not contain any point from $J_{(1,\infty)}$ and the Voronoi region of any point in $\alpha_n \cap J_{(1,\infty)}$ does not contain any point from J_1.*

Proof. By Lemma 3.3, Lemma 3.5, Lemma 3.6, and Lemma 3.7, the proposition is true for $2 \leq n \leq 5$. We now prove the proposition for all $n \geq 6$. Let $\alpha_n := \{0 < a_1 < a_2 < \cdots < a_n < 1\}$ be an optimal set of n-means for $n \geq 6$. Consider the set of six points $\beta := \{a(11), a(11, \infty), a(21), a(21, \infty), a(3), a(3, \infty)\}$. Then, the distortion error is

$$\int \min_{a \in \beta} (x-a)^2 dP = (1+\frac{43}{3}) p_{11} s_{11}^2 V + (1+\frac{43}{3}) p_{21} s_{21}^2 V + (1+\frac{43}{9}) p_3 s_3^2 V$$
$$= \frac{1383}{1831424}.$$

Since, V_n is the quantization error for n-means for $n \geq 6$, we have $V_n \leq V_6 \leq \frac{1383}{1831424} = 0.00075515$. Proceeding in the similar way, as shown in the previous lemmas, we have $a_1 < \frac{1}{4}$ and $\frac{1}{2} < a_n$. Let $j = \max\{i : a_i < \frac{1}{2}\}$. Then, $a_j < \frac{1}{2}$. We show that $a_j \leq \frac{1}{4}$. Suppose that $\frac{1}{4} < a_j < \frac{1}{2}$. Then, the following two cases can arise:

Case 1. $\frac{3}{8} \leq a_j < \frac{1}{2}$.

Then, $\frac{1}{2}(a_{j-1} + a_j) < \frac{1}{4}$ implying $a_{j-1} < \frac{1}{2} - a_j \leq \frac{1}{2} - \frac{3}{8} = \frac{1}{8} = S_{12}(0)$ yielding

$$V_n \geq \int_{\bigcup_{j=2}^{10} J_{1j}} (x - \frac{1}{8})^2 dP = \frac{13986897}{17179869184} = 0.000814145 > V_n,$$

which is a contradiction.

Case 2. $\frac{1}{4} < a_j \leq \frac{3}{8}$.

Then, $\frac{1}{2}(a_j + a_{j+1}) > \frac{1}{2}$ implying $a_{j+1} > 1 - a_j \geq 1 - \frac{3}{8} = \frac{5}{8}$ yielding

$$V_n \geq \int_{J_2} (x - \frac{5}{8})^2 dP = \frac{405}{261632} = 0.00154798 > V_n,$$

which gives a contradiction.

Hence, we can assume that $a_j \leq \frac{1}{2}$. Thus, we have seen that $\alpha_n \cap J_1 \neq \emptyset$, $\alpha_n \cap J_{(1,\infty)} \neq \emptyset$, and α_n does not contain any point from the open interval $(\frac{1}{4}, \frac{1}{2})$. Since $\frac{1}{2}(a_j + a_{j+1}) \geq \frac{1}{2}(0 + \frac{1}{2}) = \frac{1}{4}$, the Voronoi region of any point in $\alpha_n \cap J_{(1,\infty)}$ does not contain any point from J_1. Suppose that the Voronoi region of a_j contains points from $J_{(1,\infty)}$. Then, $\frac{1}{2}(a_j + a_{j+1}) > \frac{1}{2}$ implying $a_{j+1} > 1 - a_2 \geq 1 - \frac{1}{4} = \frac{3}{4}$, and so

$$V_n \geq \int_{J_2} (x - \frac{3}{4})^2 dP = \frac{813}{65408} = 0.0124297 > V_n,$$

which is a contradiction. So, we can assume that the Voronoi region of any point in $\alpha_n \cap J_1$ does not contain any point from $J_{(1,\infty)}$. Thus, the proof of the proposition is complete. □

We need the following lemmas to prove Proposition 3.13.

Lemma 3.9. *Let $V(P, J_2, \{a, b\})$ be the quantization error due to the points a and b on the set J_2, where $\frac{1}{2} \leq a < b$ and $b = \frac{5}{8}$. Then, $a = a(21, 22)$ and*

$$V(P, J_2, \{a, b\}) = \int_{J_{21} \cup J_{22}} (x - a(21, 22))^2 dP + \int_{J_{(22,\infty)}} (x - \frac{5}{8})^2 dP = \frac{2403}{10465280}.$$

Proof. Consider the set $\{\frac{11}{20}, \frac{5}{8}\}$. Then, as $S_{22}(1) < \frac{1}{2}(\frac{11}{20} + \frac{5}{8}) < S_{23}(0)$, and $V(P, J_2, \{a, b\})$ is the quantization error due to the points a and b on the set J_2, we have

$$\begin{aligned} V(P, J_2, \{a, b\}) &\leq \int_{J_{21} \cup J_{22}} (x - \frac{11}{20})^2 dP + \int_{J_{(22,\infty)}} (x - \frac{5}{8})^2 dP \\ &= \frac{2403}{10465280} = 0.000229616. \end{aligned}$$

If $\frac{37}{64} = S_{22}(1) \leq a$, then

$$\begin{aligned} V(P, J_2, \{a, b\}) &\geq \int_{J_{21} \cup J_{22}} (x - S_{22}(1))^2 dP \\ &= \frac{6831}{19136512} = 0.000356962 > V(P, J_2, \{a, b\}), \end{aligned}$$

which is a contradiction, and so we can assume that $a < S_{22}(1) = \frac{37}{64}$. If the Voronoi region of b contains points from J_{22}, we must have $\frac{1}{2}(a + b) < \frac{37}{64}$ implying $a < \frac{37}{32} - b = \frac{37}{32} - \frac{5}{8} = \frac{17}{32} = S_{21}(1)$, and so

$$\begin{aligned} V(P, J_2, \{a, b\}) &> \int_{J_{22}} (x - \frac{17}{32})^2 dP + \int_{\overset{10}{\underset{j=3}{\cup}} J_{2j}} (x - \frac{5}{8})^2 dP \\ &= \frac{276910245}{962072674304} = 0.000287827, \end{aligned}$$

yielding $V(P, J_2, \{a, b\}) > 0.000287827 > V(P, J_2, \{a, b\})$, which leads to a contradiction. So, we can assume that the Voronoi region of b does not contain any point from J_{22} yielding $a \geq a(21, 22) = \frac{11}{20}$. If the Voronoi region of a contains points from J_{23}, we must have $\frac{1}{2}(a + \frac{5}{8}) > S_{23}(0) = \frac{19}{32}$ implying $a > \frac{19}{16} - \frac{5}{8} = \frac{9}{16} = S_{22}(0)$, and then

$$\begin{aligned} V(P, J_2, \{a, b\}) &> \int_{J_{21}} (x - \frac{9}{16})^2 dP + \int_{\overset{10}{\underset{j=3}{\cup}} J_{2j}} (x - \frac{5}{8})^2 dP \\ &= \frac{17716739853}{70231305224192} = 0.000252263, \end{aligned}$$

yielding $V(P, J_2, \{a, b\}) > 0.000252263 > V(P, J_2, \{a, b\})$, which leads to a contradiction. So, the Voronoi region of a does not contain any point from J_{23} yielding $a \leq a(21, 22)$. Again, we proved $a \geq a(21, 22)$. Thus, $a = a(21, 22)$ and

$$V(P, J_2, \{a, b\}) = \int_{J_{21} \cup J_{22}} (x - a(21, 22))^2 dP + \int_{J_{(22,\infty)}} (x - \frac{5}{8})^2 dP = \frac{2403}{10465280}.$$

Thus, the proof of the lemma is complete. □

Lemma 3.10. *Let α_6 be an optimal set of six-means. Then, card$(\alpha_6 \cap J_1) = 2$ and card$(\alpha_6 \cap J_{(1,\infty)}) = 4$. Moreover, card$(\alpha_6 \cap J_2) = 2$.*

Proof. Let $\alpha_6 := \{0 < a_1 < a_2 < a_3 < a_4 < a_5 < a_6 < 1\}$ be an optimal set of six-means. Consider the set of six points

$$\beta := \{a(11), a(11, \infty), a(21), a(21, \infty), a(3), a(3, \infty)\}.$$

Then, the distortion error is

$$\int \min_{a\in\beta}(x-a)^2 dP = (1 + \frac{43}{3})p_{11}s_{11}^2 V + (1 + \frac{43}{3})p_{21}s_{21}^2 V + (1 + \frac{43}{9})p_3 s_3^2 V$$
$$= \frac{1383}{1831424}.$$

Since, V_6 is the quantization error for six-means, we have $V_6 \leq \frac{1383}{1831424} = 0.00075515$. By Proposition 3.8, we have card$(\alpha_6 \cap J_1) \geq 1$ and card$(\alpha_6 \cap J_{(1,\infty)}) \geq 1$. Moreover, the Voronoi region of any point in $\alpha_6 \cap J_1$ does not contain any point from $J_{(1,\infty)}$ and the Voronoi region of any point in $\alpha_6 \cap J_{(1,\infty)}$ does not contain any point from J_1. Suppose that card$(\alpha_6 \cap J_{(1,\infty)}) = 2$, and then taking $\beta_2 = \{a(2), a(2, \infty)\}$ we see that

$$V_6 \geq \int_{J_2 \cup J_{(2,\infty)}} \min_{a\in\beta_2}(x-a)^2 dP = \int_{J_2}(x - a(2))^2 dP + \int_{J_{(2,\infty)}}(x - a(2, \infty))^2 dP$$
$$= \frac{39}{14308} = 0.00272575,$$

i.e., $V_6 \geq 0.00272575 > V_6$, which yields a contradiction. Next, assume that card$(\alpha_6 \cap J_{(1,\infty)}) = 3$, and then taking $\beta_2 = \{a(2), a(3), a(3, \infty)\}$, we see that

$$V_6 \geq \int_{J_2}(x - a(2))^2 dP + \int_{J_3}(x - a(3))^2 dP + \int_{J_{(3,\infty)}}(x - a(3, \infty))^2 dP$$
$$= \frac{93}{114464} = 0.000812483 > V_6,$$

which gives a contradiction. Thus, we can assume that card$(\alpha_6 \cap J_{(1,\infty)}) \geq 4$. If card$(\alpha_6 \cap J_1) = 1$, then,

$$V_6 \geq \int_{J_1}(x - a(1))^2 dP = \frac{9}{7154} = 0.00125804 > V_6,$$

which yields a contradiction, and so card$(\alpha_6 \cap J_1) \geq 2$. Therefore, we can assume that card$(\alpha_6 \cap J_1) = 2$ and card$(\alpha_6 \cap J_{(1,\infty)}) = 4$. We now show that card$(\alpha_6 \cap J_2) = 2$. By Proposition 3.8, the Voronoi region of any element in $\alpha_6 \cap J_1$ does not contain any point from $J_{(1,\infty)}$, and the Voronoi region of any element in $\alpha_6 \cap J_{(1,\infty)}$ does not contain any point from J_1. We have

$\alpha_6 \cap J_{(1,\infty)} = \{\frac{1}{2} \leq a_3 < a_4 < a_5 < a_6 < 1\}$. The distortion error contributed by the set $\beta \cap J_{(1,\infty)} = \{a(21), a(21,\infty), a(3), a(3,\infty)\}$ is given by

$$\int_{J_{(1,\infty)}} \min_{a\in\beta\cap J_{(1,\infty)}} (x-a)^2 dP = (1+\frac{43}{3})p_{21}s_{21}^2 V + (1+\frac{43}{9})p_3 s_3^2 V$$
$$= \frac{831}{1831424} = 0.000453745.$$

Let $V(P, \alpha_6 \cap J_{(1,\infty)})$ be the quantization error contributed by the set $\alpha_6 \cap J_{(1,\infty)}$ in the region $J_{(1,\infty)}$. Then, we must have $V(P, \alpha_6 \cap J_{(1,\infty)}) \leq 0.000453745$. If $a_6 \leq \frac{57}{64} = S_{42}(0)$, then

$$V(P, \alpha_6 \cap J_{(1,\infty)}) \geq \int_{\bigcup_{j=5}^{8} J_j} (x-\frac{57}{64})^2 dP = \frac{145935}{306184192} = 0.000476625$$
$$> V(P, \alpha_6 \cap J_{(1,\infty)}),$$

which yields a contradiction, and so $S_{42}(0) = \frac{57}{64} < a_6$. If $\frac{3}{4} < a_4$, then

$$V(P, \alpha_6 \cap J_{(1,\infty)}) \geq \int_{J_2} (x - a(2))^2 dP = \frac{27}{57232} = 0.000471764$$
$$> V(P, \alpha_6 \cap J_{(1,\infty)}),$$

which yields a contradiction. So, we can assume that $a_4 < \frac{3}{4}$. Suppose that $\frac{5}{8} < a_4 < \frac{3}{4}$. Then, the following two cases can arise:

Case 1. $\frac{11}{16} \leq a_4 < \frac{3}{4}$.
Then, $\frac{1}{2}(a_3 + a_4) < \frac{5}{8}$ implying $a_3 < \frac{5}{4} - a_4 \leq \frac{5}{4} - \frac{11}{16} = \frac{9}{16}$, and so

$$V(P, \alpha_6 \cap J_{(1,\infty)}) \geq \min_{\{a_3<\frac{9}{16}<\frac{11}{16}\leq a_4\}} \int_{J_2} \min_{a\in\{a_3,a_4\}} (x-a)^2 dP$$
$$\geq \int_{J_2} (x-a(2))^2 dP = \frac{27}{57232},$$

implying $V(P, \alpha_6 \cap J_{(1,\infty)}) \geq \frac{27}{57232} = 0.000471764 > V(P, \alpha_6 \cap J_{(1,\infty)})$, which gives a contradiction.

Case 2. $\frac{5}{8} < a_4 < \frac{11}{16}$.
Then, $\frac{1}{2}(a_4 + a_5) > \frac{3}{4}$ implying $a_5 > \frac{3}{2} - a_4 \geq \frac{3}{2} - \frac{11}{16} = \frac{13}{16}$. Then, the following two subcases can arise:

Subcase (i). $\frac{27}{32} \leq a_5$.
Then, $S_{31}(1) = \frac{49}{64} = \frac{1}{2}(\frac{11}{16} + \frac{27}{32}) < S_{32}(0)$, and so by Lemma 3.9,

$$V(P, \alpha_6 \cap J_{(1,\infty)}) \geq \int_{J_{21}\cup J_{22}} (x - a(21,22))^2 dP + \int_{J_{(22,\infty)}} (x-\frac{5}{8})^2 dP$$
$$+ \int_{J_{31}} (x-\frac{11}{16})^2 dP + \int_{J_{32}} (x-\frac{27}{32})^2 dP$$

$$= \frac{236721}{334888960} = 0.000706864 > V(P, \alpha_6 \cap J_{(1,\infty)}),$$

which gives a contradiction.

Subcase (ii). $\frac{13}{16} < a_5 < \frac{27}{32}$.

Then, $\frac{1}{2}(a_5 + a_6) > \frac{7}{8}$ implying $a_6 > \frac{7}{4} - a_5 \geq \frac{7}{4} - \frac{27}{32} = \frac{29}{32} = S_4(1)$. First, assume that $S_4(1) < a_6 < S_5(0) = \frac{15}{16}$. Then, using Lemma 3.9,

$$\begin{aligned} V(P, \alpha_6 \cap J_{(1,\infty)}) \geq & \int_{J_{21}\cup J_{22}} (x - a(21, 22))^2 dP + \int_{J_{(22,\infty)}} (x - \frac{5}{8})^2 dP \\ & + \int_{J_3} (x - \frac{13}{16})^2 dP + \int_{J_4} (x - \frac{29}{32})^2 dP + \int_{J_5 \cup J_6} (x - \frac{15}{16})^2 dP \\ = & \frac{11529}{23920640} = 0.000481969 > V(P, \alpha_6 \cap J_{(1,\infty)}), \end{aligned}$$

which leads to a contradiction. Next, assume that $S_5(0) = \frac{15}{16} \leq a_6$. Then, as $S_{42}(0) = \frac{57}{64} = \frac{1}{2}(\frac{27}{32} + \frac{15}{16})$, using Lemma 3.9, we have

$$\begin{aligned} V(P, \alpha_6 \cap J_{(1,\infty)}) \geq & \int_{J_{21}\cup J_{22}} (x - a(21, 22))^2 dP + \int_{J_{(22,\infty)}} (x - \frac{5}{8})^2 dP \\ & + \int_{J_3} (x - \frac{13}{16})^2 dP + \int_{J_{41}} (x - \frac{27}{32})^2 dP + \int_{J_{42}} (x - \frac{15}{16})^2 dP \\ = & \frac{700899}{1339555840} = 0.000523232 > V(P, \alpha_6 \cap J_{(1,\infty)}), \end{aligned}$$

which yields a contradiction.

Hence, by Case 1 and Case 2, we can assume that $a_4 \leq \frac{5}{8}$ yielding card$(\alpha_6 \cap J_2) = 2$. Thus, the proof of the proposition is complete. □

Lemma 3.11. *Let* α_7 *be an optimal set of seven-means. Then, either* (i) *card*$(\alpha_7 \cap J_1) = 3$ *and card*$(\alpha_7 \cap J_{(1,\infty)}) = 4$*, or* (ii) *card*$(\alpha_7 \cap J_1) = 2$ *and card*$(\alpha_7 \cap J_{(1,\infty)}) = 5$.

Proof. Let $\alpha_7 := \{0 < a_1 < a_2 < \cdots < a_7 < 1\}$ be an optimal set of seven-means. Consider the set of seven points

$$\beta := \{a(11), a(12), a(12, \infty), a(21), a(21, \infty), a(3), a(3, \infty)\}.$$

Then, the distortion error due to the set β is

$$\begin{aligned} \int \min_{a\in\beta} (x - a)^2 dP = & p_{11} s_{11}^2 V + (1 + \frac{43}{9}) p_{12} s_{12}^2 V \\ & + (1 + \frac{43}{3}) p_{21} s_{21}^2 V + (1 + \frac{43}{9}) p_3 s_3^2 V \\ = & \frac{135}{261632}. \end{aligned}$$

Since, V_7 is the quantization error for seven-means, we have $V_7 \leq \frac{135}{261632} = 0.000515992$. By Proposition 3.8, we have card$(\alpha_7 \cap J_1) \geq 1$ and card$(\alpha_7 \cap$

$J_{(1,\infty)}) \geq 1$. Moreover, the Voronoi region of any point in $\alpha_7 \cap J_1$ does not contain any point from $J_{(1,\infty)}$ and the Voronoi region of any point in $\alpha_7 \cap J_{(1,\infty)}$ does not contain any point from J_1. Suppose that card$(\alpha_7 \cap J_{(1,\infty)}) = 2$, and then taking $\beta_2 = \{a(2), a(2,\infty)\}$ we see that

$$V_7 \geq \int_{J_2 \cup J_{(2,\infty)}} \min_{a \in \beta_2} (x-a)^2 dP$$
$$= \int_{J_2} (x - a(2))^2 dP + \int_{J_{(2,\infty)}} (x - a(2,\infty))^2 dP = \frac{39}{14308} = 0.00272575,$$

i.e., $V_7 \geq 0.00272575 > V_7$, which yields a contradiction. Next, assume that card$(\alpha_7 \cap J_{(1,\infty)}) = 3$, and then taking $\beta_2 = \{a(2), a(3), a(3,\infty)\}$, we see that

$$V_7 \geq \int_{J_2} (x - a(2))^2 dP + \int_{J_3} (x - a(3))^2 dP + \int_{J_{(3,\infty)}} (x - a(3,\infty))^2 dP$$
$$= \frac{93}{114464} = 0.000812483 > V_7,$$

which gives a contradiction. Thus, we can assume that card$(\alpha_7 \cap J_{(1,\infty)}) \geq 4$. If card$(\alpha_7 \cap J_1) = 1$, then,

$$V_7 \geq \int_{J_1} (x - a(1))^2 dP = \frac{9}{7154} = 0.00125804 > V_7,$$

which gives a contradiction. So, we can assume that card$(\alpha_7 \cap J_1) \geq 2$. Thus, we have either (i) card$(\alpha_7 \cap J_1) = 3$ and card$(\alpha_7 \cap J_{(1,\infty)}) = 4$, or (ii) card$(\alpha_7 \cap J_1) = 2$ and card$(\alpha_7 \cap J_{(1,\infty)}) = 5$, which is the lemma. □

Lemma 3.12. *Let α_8 be an optimal set of eight-means. Then, card$(\alpha_8 \cap J_1) = 3$ and card$(\alpha_8 \cap J_{(1,\infty)}) = 5$.*

Proof. Let $\alpha_8 := \{0 < a_1 < a_2 < \cdots < a_8 < 1\}$ be an optimal set of eight-means. Consider the set of eight points

$$\beta := \{a(11), a(12), a(12,\infty), a(21), a(21,\infty), a(3), a(4), a(4,\infty)\}.$$

Then, the distortion error due to the set β is

$$\int \min_{a \in \beta} (x-a)^2 dP = p_{11} s_{11}^2 V + (1 + \frac{43}{9}) p_{12} s_{12}^2 V + (1 + \frac{43}{3}) p_{21} s_{21}^2 V$$
$$+ p_3 s_3^2 V + (1 + \frac{43}{9}) p_4 s_4^2 V$$
$$= \frac{507}{1831424}.$$

Since V_8 is the quantization error for eight-means, we have $V_8 \leq \frac{507}{1831424} = 0.000276834$. By Proposition 3.8, we have card$(\alpha_8 \cap J_1) \geq 1$ and card$(\alpha_8 \cap J_{(1,\infty)}) \geq 1$. Moreover, the Voronoi region of any point in $\alpha_8 \cap J_1$ does not contain any point from $J_{(1,\infty)}$ and the Voronoi region of any point in $\alpha_8 \cap J_{(1,\infty)}$

does not contain any point from J_1. Suppose that $\text{card}(\alpha_8 \cap J_{(1,\infty)}) = 2$, and then taking $\beta_2 = \{a(2), a(2,\infty)\}$ we see that

$$\begin{aligned} V_8 &\geq \int_{J_2 \cup J_{(2,\infty)}} \min_{a\in\beta_2}(x-a)^2 dP \\ &= \int_{J_2}(x-a(2))^2 dP + \int_{J_{(2,\infty)}}(x-a(2,\infty))^2 dP = \frac{39}{14308} = 0.00272575, \end{aligned}$$

i.e., $V_8 \geq 0.00272575 > V_8$, which yields a contradiction. Suppose that $\text{card}(\alpha_8 \cap J_{(1,\infty)}) = 3$, and then taking $\beta_3 = \{a(2), a(3), a(3,\infty)\}$, we see that

$$\begin{aligned} V_8 &\geq \int_{J_2}(x-a(2))^2 dP + \int_{J_3}(x-a(3))^2 dP + \int_{J_{(3,\infty)}}(x-a(3,\infty))^2 dP \\ &= \frac{93}{114464} = 0.000812483 > V_8, \end{aligned}$$

which gives a contradiction. Next, assume that $\text{card}(\alpha_8 \cap J_{(1,\infty)}) = 4$, and then taking

$$\beta_4 = \{a(21), a(21,\infty), a(3), a(3,\infty)\},$$

we see that

$$V_8 \geq (1+\frac{43}{3})p_{21}s_{21}^2 V + (1+\frac{43}{9})p_3 s_3^3 V = \frac{831}{1831424} = 0.000453745 > V_8,$$

which gives a contradiction. So, we can assume that $\text{card}(\alpha_8 \cap J_{(1,\infty)}) \geq 5$. If $\text{card}(\alpha_8 \cap J_1) = 1$, then,

$$V_8 \geq \int_{J_1}(x-a(1))^2 dP = \frac{9}{7154} = 0.00125804 > V_8,$$

which leads to a contradiction. If $\text{card}(\alpha_8 \cap J_1) = 2$, then taking $\beta_2 = \{a(11), a(11,\infty)\}$, we see that

$$V_8 \geq \int_{J_1} \min_{a\in\beta_2}(x-a)^2 dP = (1+\frac{43}{3})p_{11}s_{11}^2 V = \frac{69}{228928} = 0.000301405 > V_8,$$

which is a contradiction. So, we can assume that $\text{card}(\alpha_8 \cap J_1) \geq 3$. Since $\text{card}(\alpha_8 \cap J_1) \geq 3$ and $\text{card}(\alpha_8 \cap J_{(1,\infty)}) \geq 5$, we have $\text{card}(\alpha_8 \cap J_1) = 3$ and $\text{card}(\alpha_8 \cap J_{(1,\infty)}) = 5$, which is the lemma. □

Proposition 3.13. *Let α_n be an optimal set of n-means for P such that card$(\alpha_n \cap J_{(k,\infty)}) \geq 2$ for some $k \in \mathbb{N}$ and $n \in \mathbb{N}$. Then, $\alpha_n \cap J_{k+1} \neq \emptyset$, $\alpha_n \cap J_{(k+1,\infty)} \neq \emptyset$, and α_n does not contain any point from the open interval $(S_{k+1}(1), S_{k+2}(0))$. Moreover, the Voronoi region of any point in $\alpha_n \cap J_{k+1}$ does not contain any point from $J_{(k+1,\infty)}$ and the Voronoi region of any point in $\alpha_n \cap J_{(k+1,\infty)}$ does not contain any point from J_{k+1}.*

Proof. By Proposition 3.8, since α_n does not contain any point from $(\frac{1}{4}, \frac{1}{2})$, the Voronoi region of any point in $\alpha_n \cap J_1$ does not contain any point from $J_{(1,\infty)}$, and the Voronoi region of any point in $\alpha_n \cap J_{(1,\infty)}$ does not contain any point from J_1, to prove the proposition it is enough to prove it for $k = 1$, and then inductively the proposition will follow for all $k \geq 2$. Fix $k = 1$. By Lemma 3.5, it is clear that the proposition is true for $n = 3$. Let $\alpha_4 := \{0 < a_1 < a_2 < a_3 < a_4 < 1\}$ be an optimal set of four-means. In the proof of Lemma 3.6, we have seen that $\frac{1}{2} \leq a_2$ yielding $\alpha_4 \cap J_{(1,\infty)} = \{\frac{1}{2} \leq a_2 < a_3 < a_4 < 1\}$, i.e., card$(\alpha_4 \cap J_{(1,\infty)}) = 3 \geq 2$. We now prove the proposition for $n = 4$. Let $V(P, \alpha_4 \cap J_{(1,\infty)})$ be the quantization error contributed by the set $\alpha_4 \cap J_{(1,\infty)}$. The distortion error due to the set $\beta := \{a(2), a(3), a(3,\infty)\}$ of three points on $J_{(1,\infty)}$ is given by

$$\int_{J_{(1,\infty)}} \min_{a\in\beta}(x-a)^2 dP = p_2 s_2^2 V + (1 + \frac{43}{9}) p_3 s_3^2 V = \frac{93}{114464} = 0.000812483,$$

and so $V(P, \alpha_4 \cap J_{(1,\infty)}) \leq 0.000812483$. If $a_2 \geq \frac{39}{64} = S_{24}(0)$, then

$$\begin{aligned} V(P, \alpha_4 \cap J_{(1,\infty)}) &\geq \int_{J_{21}\cup J_{22}\cup J_{23}} (x - \frac{39}{64})^2 dP = \frac{269769}{267911168} = 0.00100693 \\ &> V(P, \alpha_4 \cap J_{(1,\infty)}), \end{aligned}$$

which is a contradiction. So, we can assume that $a_2 < \frac{39}{64}$. Suppose that $a_3 \leq \frac{5}{7}$. Then, as $S_3(1) = \frac{13}{16} < \frac{1}{2}(\frac{5}{7} + a(3,\infty)) < \frac{7}{8}$, we have

$$\begin{aligned} V(P, \alpha_4 \cap J_{(1,\infty)}) &\geq \int_{J_3} (x - \frac{5}{7})^2 dP + \int_{J_{(3,\infty)}} (x - a(3,\infty))^2 dP \\ &= \frac{297}{228928} = 0.00129735 \end{aligned}$$

implying $V(P, \alpha_4 \cap J_{(1,\infty)}) \geq 0.00129735 > V(P, \alpha_4 \cap J_{(1,\infty)})$, which is a contradiction. Next, suppose that $\frac{5}{7} \leq a_3 \leq \frac{3}{4}$. Then, as $S_2(1) < \frac{1}{2}(a(2) + \frac{5}{7})$ and $S_3(1) < \frac{1}{2}(\frac{3}{4} + a(3,\infty)) < \frac{7}{8} = S_4(0)$, we have

$$\begin{aligned} V(P, \alpha_4 \cap J_{(1,\infty)}) \geq &\int_{J_2} (x - a(2))^2 dP + \int_{J_3} (x - \frac{3}{4})^2 dP \\ &+ \int_{J_{(3,\infty)}} (x - a(3,\infty))^2 dP = \frac{963}{915712} \end{aligned}$$

yielding $V(P, \alpha_4 \cap J_{(1,\infty)}) \geq \frac{963}{915712} = 0.00105164 > V(P, \alpha_4 \cap J_{(1,\infty)})$, which gives a contradiction. Thus, we have $\frac{3}{4} < a_3$. Since $a_2 \leq \frac{39}{64} < \frac{5}{8}$ and $\frac{3}{4} < a_3$, the set $\alpha_4 \cap J_{(1,\infty)}$ does not contain any point from the open interval $(S_2(1), S_3(0))$. Since $\frac{1}{2}(a_2 + a_3) \geq \frac{1}{2}(\frac{1}{2} + \frac{3}{4}) = \frac{5}{8} = S_2(1)$, the Voronoi region of any point in $\alpha_4 \cap J_{(2,\infty)}$ does not contain any point from J_2. Suppose that the Voronoi region of any point in $\alpha_4 \cap J_2$ contains points from $J_{(2,\infty)}$. Then, $\frac{1}{2}(a_2 + a_3) > \frac{3}{4}$ implying $a_3 > \frac{3}{2} - a_2 \geq \frac{3}{2} - \frac{39}{64} = \frac{57}{64}$, and so

$$V_4 \geq \int_{J_3} (x - \frac{57}{64})^2 dP = \frac{10155}{4784128} = 0.00212264 > V_4,$$

which leads to a contradiction. Hence, the Voronoi region of any point in $\alpha_4 \cap J_2$ does not contain any point from $J_{(2,\infty)}$. Thus, the proposition is true for $n = 4$.

From the proof of Lemma 3.7, we see that if $\alpha_5 = \{0 < a_1 < a_2 < a_3 < a_4 < a_5 < 1\}$ is an optimal set of five-means, then $\alpha_5 \cap J_{(1,\infty)} = \{\frac{1}{2} \leq a_3 < a_4 < a_5 < 1\}$. Thus, the proof of the proposition for $n = 5$ follows exactly in the similar ways as the proof for $n = 4$ given above.

Now, we prove the proposition for $n = 6$. Let $\alpha_6 := \{0 < a_1 < a_2 < a_3 < a_4 < a_5 < a_6 < 1\}$ be an optimal set of six-means. Then, by Lemma 3.10, we know that $\text{card}(\alpha_6 \cap J_2) = 2$, and $\text{card}(\alpha_6 \cap J_{(1,\infty)}) = 4$. Thus, we see that $\alpha_6 \cap J_2 = \{a_3, a_4\} \neq \emptyset$ and $\alpha_6 \cap J_{(2,\infty)} = \{a_5, a_6\} \neq \emptyset$. As shown in the proof of Lemma 3.10, we have $\alpha_6 \cap J_{(1,\infty)} = \{\frac{1}{2} \leq a_3 < a_4 < a_5 < a_6 < 1\}$, and if $V(P, \alpha_6 \cap J_{(1,\infty)})$ is the quantization error contributed by the set $\alpha_6 \cap J_{(1,\infty)}$ in the region $J_{(1,\infty)}$, then we have $V(P, \alpha_6 \cap J_{(1,\infty)}) \leq 0.000453745$. We now show that the Voronoi region of any point in $\alpha_6 \cap J_2$ does not contain any point from $J_{(2,\infty)}$. If it does, then we must have $\frac{1}{2}(a_4 + a_5) > \frac{3}{4}$ implying $a_5 > \frac{3}{2} - a_4 \geq \frac{3}{2} - \frac{5}{8} = \frac{7}{8}$, and so

$$V(P, \alpha_6 \cap J_{(1,\infty)}) \geq \int_{J_3} (x - \frac{7}{8})^2 dP = \frac{813}{523264} = 0.00155371 > V(P, \alpha_6 \cap J_{(1,\infty)}),$$

which is a contradiction. Also, notice that the Voronoi region of any element from $\alpha_6 \cap J_{(2,\infty)}$ does not contain any point from J_2, if it does we must have $\frac{1}{2}(a_4 + a_5) < \frac{5}{8}$ implying $a_4 < \frac{5}{4} - a_5 \leq \frac{5}{4} - \frac{3}{4} = \frac{1}{2}$, which is a contradiction as $\frac{1}{2} \leq a_3 < a_4$.

Now, we prove the proposition for $n = 7$. Let $\alpha_7 := \{0 < a_1 < a_2 < \cdots < a_7 < 1\}$ be an optimal set of seven-means. By Lemma 3.11, first assume that $\text{card}(\alpha_7 \cap J_{(1,\infty)}) = 4$, i.e., $\frac{1}{2} \leq a_4$. Let $V(P, \alpha_7 \cap J_{(1,\infty)})$ be the quantization error contributed by the set $\alpha_7 \cap J_{(1,\infty)}$ in the region $J_{(1,\infty)}$. Let $\beta := \{a(11), a(12), a(12,\infty), a(21), a(21,\infty), a(3), a(3,\infty)\}$. The distortion error due to the set $\beta \cap J_{(1,\infty)} := \{a(21), a(21,\infty), a(3), a(3,\infty)\}$ is given by

$$\begin{aligned}\int_{J_{(1,\infty)}} \min_{a \in \beta \cap J_{(1,\infty)}} (x - a)^2 dP &= (1 + \frac{43}{3}) p_{21} s_{21}^2 V + (1 + \frac{43}{9}) p_3 s_3^2 V \\ &= \frac{831}{1831424} = 0.000453745,\end{aligned}$$

and so $V(P, \alpha_7 \cap J_{(1,\infty)}) \leq 0.000453745$. If $a_4 \geq \frac{77}{128} = S_{23}(1)$, then

$$\begin{aligned}V(P, \alpha_7 \cap J_{(1,\infty)}) &\geq \int_{J_{21} \cup J_{22} \cup J_{23}} (x - \frac{77}{128})^2 dP = \frac{852849}{1071644672} = 0.000795832 \\ &> V(P, \alpha_7 \cap J_{(1,\infty)}),\end{aligned}$$

which is a contradiction. So, we can assume that $a_4 < \frac{77}{128} = S_{23}(1)$. Suppose that $\frac{11}{16} \leq a_5$. Then, as $\frac{1}{2}(a(2) + a_5) \geq \frac{1}{2}(\frac{4}{7} + \frac{11}{16}) > \frac{5}{8}$, we have

$$\begin{aligned}V(P, \alpha_7 \cap J_{(1,\infty)}) &\geq \int_{J_2} (x - a(2))^2 dP = \frac{27}{57232} = 0.000471764 \\ &> V(P, \alpha_7 \cap J_{(1,\infty)}),\end{aligned}$$

which leads to a contradiction. So, we can assume that $a_5 \leq \frac{11}{16}$. Suppose that $\frac{5}{8} < a_5 \leq \frac{11}{16}$. Then, $\frac{1}{2}(a_5+a_6) > \frac{3}{4}$ implying $a_6 > \frac{3}{2}-a_6 \geq \frac{3}{2}-\frac{11}{16} = \frac{13}{16} = S_3(1)$. Then, the following two cases can arise:

Case (i). $\frac{27}{32} \leq a_6$.

Then, $S_{31}(1) = \frac{49}{64} = \frac{1}{2}(\frac{11}{16} + \frac{27}{32}) < S_{32}(0)$, and so by Lemma 3.9,

$$\begin{aligned} V(P, \alpha_7 \cap J_{(1,\infty)}) &\geq \int_{J_{21}\cup J_{22}} (x - a(21,22))^2 dP + \int_{J_{(22,\infty)}} (x - \frac{5}{8})^2 dP \\ &\quad + \int_{J_{31}} (x - \frac{11}{16})^2 dP + \int_{J_{32}} (x - \frac{27}{32})^2 dP \\ &= \frac{236721}{334888960} = 0.000706864 > V(P, \alpha_7 \cap J_{(1,\infty)}), \end{aligned}$$

which gives a contradiction.

Case (ii). $\frac{13}{16} < a_6 < \frac{27}{32}$.

Then, $\frac{1}{2}(a_6 + a_7) > \frac{7}{8}$ implying $a_7 > \frac{7}{4} - a_6 \geq \frac{7}{4} - \frac{27}{32} = \frac{29}{32} = S_4(1)$. First, assume that $S_4(1) < a_7 < S_5(0) = \frac{15}{16}$. Then, using Lemma 3.9,

$$\begin{aligned} V(P, \alpha_7 \cap J_{(1,\infty)}) &\geq \int_{J_{21}\cup J_{22}} (x - a(21,22))^2 dP + \int_{J_{(22,\infty)}} (x - \frac{5}{8})^2 dP \\ &\quad + \int_{J_3} (x - \frac{13}{16})^2 dP + \int_{J_4} (x - \frac{29}{32})^2 dP \\ &\quad + \int_{J_5\cup J_6} (x - \frac{15}{16})^2 dP \\ &= \frac{11529}{23920640} = 0.000481969 > V(P, \alpha_7 \cap J_{(1,\infty)}), \end{aligned}$$

which leads to a contradiction. Next, assume that $S_5(0) = \frac{15}{16} \leq a_7$. Then, as $S_{42}(0) = \frac{57}{64} = \frac{1}{2}(\frac{27}{32} + \frac{15}{16})$, using Lemma 3.9, we have

$$\begin{aligned} V(P, \alpha_7 \cap J_{(1,\infty)}) &\geq \int_{J_{21}\cup J_{22}} (x - a(21,22))^2 dP + \int_{J_{(22,\infty)}} (x - \frac{5}{8})^2 dP \\ &\quad + \int_{J_3} (x - \frac{13}{16})^2 dP + \int_{J_{41}} (x - \frac{27}{32})^2 dP \\ &\quad + \int_{J_{42}} (x - \frac{15}{16})^2 dP \\ &= \frac{700899}{1339555840} = 0.000523232 > V(P, \alpha_7 \cap J_{(1,\infty)}), \end{aligned}$$

which yields a contradiction.

Hence, by Case (i) and Case (ii), we can assume that $a_5 \leq \frac{5}{8}$. If $a_6 \leq \frac{3}{4}$, then as $\frac{13}{16} = S_3(1) = \frac{1}{2}(\frac{3}{4} + \frac{7}{8}) < \frac{1}{2}(\frac{3}{4} + a(3,\infty)) = \frac{1}{2}(\frac{3}{4} + \frac{13}{14}) < \frac{7}{8}$, we have

$$V_7 \geq \int_{J_3} (x - \frac{3}{4})^2 dP + \int_{J_{(3,\infty)}} (x - a(3,\infty)) dP$$

$$= \frac{531}{915712} = 0.000579877 > V_7,$$

which leads to a contradiction. So, we can assume that $\frac{3}{4} < a_6$. Thus, it is proved that $\alpha_7 \cap J_2 \neq \emptyset$, $\alpha_7 \cap J_{(2,\infty)} \neq \emptyset$, and α_7 does not contain any point from the open interval $(S_2(1), S_3(0))$. Since $\frac{1}{2}(a_5 + a_6) \geq \frac{1}{2}(\frac{1}{2} + \frac{3}{4}) = \frac{5}{8}$, the Voronoi region of any point in $\alpha_7 \cap J_{(2,\infty)}$ does not contain any point from J_2. If the Voronoi region of any point in $\alpha_7 \cap J_2$ contains points from $J_{(2,\infty)}$, we must have $\frac{1}{2}(a_5 + a_6) > \frac{3}{4}$ implying $a_6 > \frac{3}{2} - a_5 \geq \frac{3}{2} - \frac{5}{8} = \frac{7}{8}$, and so

$$V(P, \alpha_7 \cap J_{(1,\infty)}) \geq \int_{J_3} (x - \frac{7}{8})^2 dP = \frac{813}{523264} = 0.00155371 > V(P, \alpha_7 \cap J_{(1,\infty)}),$$

which is a contradiction. Thus, the Voronoi region of any point in $\alpha_7 \cap J_2$ does not contain any point from $J_{(2,\infty)}$ as well.

If we assume $\text{card}(\alpha_7 \cap J_{(1,\infty)}) = 5$, with the help of Lemma 3.11, similarly we can prove that the proposition is true. Notice that if we take $n = 8$, then by Lemma 3.12, we have $\text{card}(\alpha_8 \cap J_{(1,\infty)}) = 5$. Thus, the proof of the proposition for the case $n = 8$ is exactly same as the proof of the proposition for $n = 7$ with $\text{card}(\alpha_7 \cap J_{(1,\infty)}) = 5$.

Now, we prove the proposition for any $n \geq 9$. Let $\alpha_n := \{0 < a_1 < a_2 < \cdots < a_n < 1\}$ be an optimal set of n-means for any $n \geq 9$ such that $\text{card}(\alpha_n \cap J_{(1,\infty)}) \geq 2$. Let $V(P, \alpha_n \cap J_{(1,\infty)})$ be the quantization error contributed by the set $\alpha_n \cap J_{(1,\infty)}$ in the region $J_{(1,\infty)}$. Let

$$\beta := \{a(11), a(12), a(12, \infty), a(21), a(22), a(22, \infty), a(3), a(4), a(4, \infty)\}.$$

The distortion error due to the set

$$\beta \cap J_{(1,\infty)} := \{a(21), a(22), a(22, \infty), a(3), a(4), a(4, \infty)\}$$

is given by

$$\int_{J_{(1,\infty)}} \min_{a \in \beta \cap J_{(1,\infty)}} (x - a)^2 dP$$
$$= p_{21} s_{21}^2 V + (1 + \frac{43}{9}) p_{22} s_{22}^2 V + p_3 s_3^2 V + (1 + \frac{43}{9}) p_4 s_4^2 V = \frac{915}{7325696},$$

and so $V(P, \alpha_n \cap J_{(1,\infty)}) \leq \frac{915}{7325696} = 0.000124903$. Suppose that α_n does not contain any point from J_2. Since by Proposition 3.8, the Voronoi region of any point in $\alpha_n \cap J_1$ does not contain any point from $J_{(1,\infty)}$, we have

$$V(P, \alpha_n \cap J_{(1,\infty)}) \geq \int_{J_2} (x - \frac{5}{8})^2 dP = \frac{405}{261632} = 0.00154798 > V(P, \alpha_n \cap J_{(1,\infty)}),$$

which leads to a contradiction. So, we can assume that $\alpha_n \cap J_2 \neq \emptyset$. Let $j := \max\{i : a_i \leq \frac{5}{8} \text{ for all } 1 \leq i \leq n\}$, and so $a_j \leq \frac{5}{8}$. We now show that $a_{j+1} \geq \frac{3}{4}$. Suppose that $\frac{5}{8} < a_{j+1} < \frac{3}{4}$. Then, the following two cases can arise:

Case 1. $\frac{5}{8} < a_{j+1} \leq \frac{11}{16}$.

Then, $\frac{1}{2}(a_{j+1}+a_{j+2}) > \frac{3}{4}$ implying $a_{j+2} > \frac{3}{2} - a_{j+1} \geq \frac{3}{2} - \frac{11}{16} = \frac{13}{16}$, and so

$$V(P, \alpha_n \cap J_{(1,\infty)}) \geq \int_{J_3} (x - \frac{13}{16})^2 dP = \frac{405}{2093056} = 0.000193497$$
$$> V(P, \alpha_n \cap J_{(1,\infty)}),$$

which is contradiction.

Case 2. $\frac{11}{16} \leq a_{j+1} < \frac{3}{4}$.

Then, $\frac{1}{2}(a_j + a_{j+1}) < \frac{5}{8}$ implying $a_j < \frac{5}{4} - a_{j+1} \leq \frac{5}{4} - \frac{11}{16} = \frac{9}{16} = S_{22}(0)$, and so

$$V(P, \alpha_n \cap J_{(1,\infty)}) \geq \int_{J_{22}\cup J_{23}\cup J_{24}} (x - \frac{9}{16})^2 dP = \frac{99}{524288} = 0.000188828$$
$$> V(P, \alpha_n \cap J_{(1,\infty)}),$$

which gives a contradiction.

Thus, we have proved that $\alpha_n \cap J_2 \neq \emptyset$, $\alpha_n \cap J_{(2,\infty)} \neq \emptyset$, and α_n does not contain any point from the open interval $(S_2(1), S_3(0))$. Since $\frac{1}{2}(a_j + a_{j+1}) \geq \frac{1}{2}(\frac{1}{2} + \frac{3}{4}) = \frac{5}{8}$, the Voronoi region of any point in $\alpha_n \cap J_{(2,\infty)}$ does not contain any point from J_2. If the Voronoi region of any point in $\alpha_n \cap J_2$ contains points from $J_{(2,\infty)}$, we must have $\frac{1}{2}(a_j+a_{j+1}) > \frac{3}{4}$ implying $a_{j+1} > \frac{3}{2} - a_j \geq \frac{3}{2} - \frac{5}{8} = \frac{7}{8}$, and so

$$V(P, \alpha_n \cap J_{(1,\infty)}) \geq \int_{J_3} (x - \frac{7}{8})^2 dP = \frac{813}{523264} = 0.00155371$$
$$> V(P, \alpha_n \cap J_{(1,\infty)}),$$

which is a contradiction. Hence, the Voronoi region of any point in $\alpha_n \cap J_2$ does not contain any point from $J_{(2,\infty)}$. Thus, the proof of the proposition is complete. □

Proposition 3.14. *Let α_n be an optimal set of n-means for $n \geq 2$. Then, there exists a positive integer k such that $\alpha_n \cap J_j \neq \emptyset$ for all $1 \leq j \leq k$, and card$(\alpha_n \cap J_{(k,\infty)}) = 1$. Moreover, if $n_j :=$ card(α_j), where $\alpha_j := \alpha_n \cap J_j$, then $n = \sum_{j=1}^{k} n_j + 1$, with*

$$V_n = \begin{cases} p_1 s_1^2 V + \frac{43}{3} p_1 s_1^2 V & \text{if } k = 1, \\ \sum_{j=1}^{k} p_j s_j^2 V_{n_j} + \frac{43}{9} p_k s_k^2 V & \text{if } k \geq 2. \end{cases}$$

Proof. Proposition 3.8 says that if α_n is an optimal set of n-means for $n \geq 2$, then $\alpha_n \cap J_1 \neq \emptyset$, $\alpha_n \cap J_{(1,\infty)} \neq \emptyset$, and α_n does not contain any point from the open interval $(S_1(1), S_2(0))$. Proposition 3.13 says that if card$(\alpha_n \cap J_{(k,\infty)}) \geq 2$ for some $k \in \mathbb{N}$, then $\alpha_n \cap J_{k+1} \neq \emptyset$ and $\alpha_n \cap J_{(k+1,\infty)} \neq \emptyset$. Moreover, α_n does not take any point from the open interval $(S_{k+1}(1), S_{k+2}(0))$. Thus, by Induction Principle, we can say that if α_n is an optimal set of n-means for $n \geq 2$, then there exists a positive integer k such that $\alpha_n \cap J_j \neq \emptyset$ for all $1 \leq j \leq k$ and card$(\alpha_n \cap J_{(k,\infty)}) = 1$.

For a given $n \geq 2$, write $\alpha_j := \alpha_n \cap J_j$ and $n_j := \text{card}(\alpha_j)$. Since α_j are disjoints for $1 \leq j \leq k$, and α_n does not contain any point from the open intervals $(S_\ell(1), S_{\ell+1}(0))$ for $1 \leq \ell \leq k$, we have $\alpha_n = \bigcup_{j=1}^{k} \alpha_j \cup \{a(k,\infty)\}$ and $n = n_1 + n_2 + \cdots + n_k + 1$. Then, using Lemma 2.1, we deduce

$$\begin{aligned} V_n &= \int \min_{a\in\alpha_n} \|x-a\|^2 dP \\ &= \sum_{j=1}^{k} \int_{J_j} \min_{a\in\alpha_j} (x-a)^2 dP + \int_{J_{(k,\infty)}} (x - a(k,\infty))^2 dP \\ &= \sum_{j=1}^{k} p_j \int \min_{a\in\alpha_j} (x-a)^2 d(P\circ S_j^{-1}) + \int_{J_{(k,\infty)}} (x - a(k,\infty))^2 dP, \end{aligned}$$

which yields

$$V_n = \sum_{j=1}^{k} p_j s_j^2 \int \min_{a\in S_j^{-1}(\alpha_j)} (x-a)^2 dP + \frac{43}{9} p_k s_k^2 V. \tag{6}$$

We now show that $S_j^{-1}(\alpha_j)$ is an optimal set of n_j-means, where $1 \leq j \leq k$. If $S_j^{-1}(\alpha_j)$ is not an optimal set of n_j-means, then we can find a set $\beta \subset \mathbb{R}$ with $\text{card}(\beta) = n_j$ such that $\int \min_{b\in\beta} (x-b)^2 dP < \int \min_{a\in S_j^{-1}(\alpha_j)} (x-a)^2 dP$. But, then $S_j(\beta) \cup (\alpha_n \setminus \alpha_j)$ is a set of cardinality n such that

$$\int \min_{a\in S_j(\beta)\cup(\alpha_n\setminus\alpha_j)} (x-a)^2 dP < \int \min_{a\in\alpha_n} (x-a)^2 dP,$$

which contradicts the optimality of α_n. Thus, $S_j^{-1}(\alpha_j)$ is an optimal set of n_j-means for $1 \leq j \leq k$. Hence, by (6) we have

$$V_n = \sum_{j=1}^{k} p_j s_j^2 V_{n_j} + \frac{43}{9} p_k s_k^2 V.$$

Thus, the proof of the proposition is yielded. □

We need the following lemma to prove the main theorem (Theorem 3.1) of the paper.

Lemma 3.15. *For any $\omega \in \mathbb{N}^k$, $k \geq 1$, let $E(a(\omega))$ and $E(a(\omega,\infty))$ be given by (5). Then, for $\omega, \tau \in \mathbb{N}^k$, $k \geq 1$, we have*

(i) $E(a(\omega)) > E(a(\tau))$ *if and only if* $E(a(\omega 1)) + E(a(\omega 1,\infty)) + E(a(\tau)) < E(a(\omega)) + E(a(\tau 1)) + E(a(\tau 1,\infty))$;

(ii) $E(a(\omega)) > E(a(\tau,\infty))$ *if and only if* $E(a(\omega 1)) + E(a(\omega 1,\infty)) + E(a(\tau,\infty)) < E(a(\omega)) + E(a(\tau^-(\tau_{|\tau|}+1))) + E(a(\tau^-(\tau_{|\tau|}+1),\infty))$;

(iii) $E(a(\omega,\infty)) > E(a(\tau))$ *if and only if* $E(a(\omega^-(\omega_{|\omega|}+1))) + E(a(\omega^-(\omega_{|\omega|}+1),\infty)) + E(a(\tau)) < E(a(\omega,\infty)) + E(a(\tau 1)) + E(a(\tau 1,\infty))$;

(iv) $E(a(\omega,\infty)) > E(a(\tau,\infty))$ *if and only if* $E(a(\omega^-(\omega_{|\omega|}+1)))+E(a(\omega^-(\omega_{|\omega|}+1),\infty))+E(a(\tau,\infty)) < E(a(\omega,\infty))+E(a(\tau^-(\tau_{|\tau|}+1)))+E(a(\tau^-(\tau_{|\tau|}+1),\infty))$.

Proof. To prove (i), using Lemma 2.6, we see that

$$\begin{aligned} LHS &= E(a(\omega 1))+E(a(\omega 1,\infty))+E(a(\tau)) \\ &= p_{\omega 1}s_{\omega 1}^2 V(1+\frac{43}{3})+p_\tau s_\tau^2 V \\ &= \frac{1}{64}p_\omega s_\omega^2 V(1+\frac{43}{3})+p_\tau s_\tau^2 V, \\ RHS &= E(a(\omega))+E(a(\tau 1))+E(a(\tau 1,\infty)) \\ &= p_\omega s_\omega^2 V+\frac{1}{64}p_\tau s_\tau^2 V(1+\frac{43}{3}). \end{aligned}$$

Thus, $LHS < RHS$ if and only if $\frac{1}{64}p_\omega s_\omega^2 V(1+\frac{43}{3})+p_\tau s_\tau^2 V < p_\omega s_\omega^2 V+\frac{1}{64}p_\tau s_\tau^2 V(1+\frac{43}{3})$, which yields $p_\omega s_\omega^2 V > p_\tau s_\tau^2 V$, i.e., $E(a(\omega)) > E(a(\tau))$. Thus (i) is proved. To prove (ii), let us first assume $\tau_{|\tau|}=1$. Notice that $p_{\tau^-(\tau_{|\tau|}+1)} = p_{\tau^-}p_{\tau_{|\tau|}+1} = \frac{3}{2}p_\tau$, and $s_{\tau^-(\tau_{|\tau|}+1)} = s_{\tau^-}s_{\tau_{|\tau|}+1} = \frac{1}{2}s_\tau$, and then using Lemma 2.6, we have

$$\begin{aligned} LHS &= E(a(\omega 1))+E(a(\omega 1,\infty))+E(a(\tau,\infty)) \\ &= p_{\omega 1}s_{\omega 1}^2 V(1+\frac{43}{3})+\frac{43}{3}p_\tau s_\tau^2 V \\ &= \frac{1}{64}p_\omega s_\omega^2 V(1+\frac{43}{3})+\frac{43}{3}p_\tau s_\tau^2 V, \\ RHS &= E(a(\omega))+E(a(\tau^-(\tau_{|\tau|}+1)))+E(a(\tau^-(\tau_{|\tau|}+1),\infty)) \\ &= p_\omega s_\omega^2 V+p_{\tau^-(\tau_{|\tau|}+1)}s_{\tau^-(\tau_{|\tau|}+1)}^2 V(1+\frac{43}{9}) \\ &= p_\omega s_\omega^2 V+p_\tau s_\tau^2 V\frac{3}{8}(1+\frac{43}{9}). \end{aligned}$$

Thus, $LHS < RHS$ if and only if $\frac{1}{64}p_\omega s_\omega^2 V(1+\frac{43}{3})+\frac{43}{3}p_\tau s_\tau^2 V < p_\omega s_\omega^2 V+p_\tau s_\tau^2 V\frac{3}{8}(1+\frac{43}{9})$, which yields

$$p_\omega s_\omega^2 V > \frac{43}{3}p_\tau s_\tau^2 V\frac{\Big(\frac{43}{3}-\frac{3}{8}(1+\frac{43}{9})\Big)\frac{3}{43}}{1-\frac{1}{64}(1+\frac{43}{3})} > \frac{43}{3}p_\tau s_\tau^2 V,$$

i.e., $E(a(\omega)) > E(a(\tau,\infty))$. Thus, (ii) is proved under the assumption $\tau_{|\tau|}=1$. Similarly by taking $\tau_{|\tau|}\geq 2$, we can prove (ii). Thus, the proof of (ii) is complete. Proceeding in the similar way, (iii) and (iv) can be proved. This concludes the proof of the lemma. □

The following proposition gives some properties of $E(\omega)$ for $\omega\in\mathbb{N}^*$.

Proposition 3.16. *Let* ω,τ *be two nonempty words in* $\mathbb{N}^*$ *with* $p_\omega = p_\tau$. *Then, the quantization error satisfies the following conditions*:

(i) $E(a(\omega)) = E(a(\tau))$.
(ii) *If* $\omega_{|\omega|} = \tau_{|\tau|}$, *then* $E(a(\omega,\infty)) = E(a(\tau,\infty))$.
(iii) *If* $\omega_{|\omega|} \neq \tau_{|\tau|} = 1$, *then* $E(a(\omega,\infty)) = \frac{1}{3}E(a(\tau,\infty))$.
(iv) *If* $1 = \omega_{|\omega|} \neq \tau_{|\tau|}$, *then* $E(a(\omega,\infty)) = 3E(a(\tau,\infty))$.

Proof. (i) By Lemma 2.8, $p_\omega = p_\tau$ implies $s_\omega = s_\tau$, and so

$$E(a(\omega)) = p_\omega s_\omega^2 V = p_\tau s_\tau^2 V = E(a(\tau)).$$

(ii) Here two cases can arise: $\omega_{|\omega|} = \tau_{|\tau|} = 1$ or $\omega_{|\omega|} = \tau_{|\tau|} \geq 2$. In either case, using Lemma 2.6 one can see that $E(a(\omega,\infty)) = E(a(\tau,\infty))$.

(iii) If $\omega_{|\omega|} \neq \tau_{|\tau|} = 1$, then, $\omega_{|\omega|} \geq 2$ and $\tau_{|\tau|} = 1$, and so by Lemma 2.6 and Lemma 2.8, we get

$$E(a(\omega,\infty)) = \frac{43}{9} p_\omega s_\omega^2 V = \frac{1}{3}\frac{43}{3} p_\tau s_\tau^2 V = \frac{1}{3}E(a(\tau,\infty)).$$

Due to symmetry (iv) follows from (iii), and thus the proof of the proposition is complete. □

Proposition 3.17. *Let α_n be an optimal set of n-means for $n \geq 2$. Then, for $c \in \alpha_n$, we have $c = a(\omega)$, or $c = a(\omega,\infty)$ for some $\omega \in \mathbb{N}^*$.*

Proof. Let α_n be an optimal set of n-means for $n \geq 2$ such that $c \in \alpha_n$. By Proposition 3.13, there exists a positive integer k_1 such that $\alpha_n \cap J_{j_1} \neq \emptyset$ for $1 \leq j_1 \leq k_1$, and $\text{card}(\alpha_n \cap J_{(k_1,\infty)}) = 1$, and α_n does not contain any point from the open intervals $(S_\ell(1), S_{\ell+1}(0))$ for $1 \leq \ell \leq k_1$. If $c \in \alpha_n \cap J_{(k_1,\infty)}$, then $c = a(k_1,\infty)$. If $c \in \alpha_n \cap J_{j_1}$ for some $1 \leq j_1 \leq k_1$ with $\text{card}(\alpha_n \cap J_{j_1}) = 1$, then $c = a(j_1)$. Suppose that $c \in \alpha_n \cap J_{j_1}$ for some $1 \leq j_1 \leq k_1$ and $\text{card}(\alpha_n \cap J_{j_1}) \geq 2$. Then, as similarity mappings preserve the ratio of the distances of a point from any other two points, using Proposition 3.13 again, there exists a positive integer k_2 such that $\alpha_n \cap J_{j_1 j_2} \neq \emptyset$ for $1 \leq j_2 \leq k_2$, and $\text{card}(\alpha_n \cap J_{(j_1 k_2,\infty)}) = 1$, and α_n does not contain any point from the open intervals $(S_{j_1 \ell}(1), S_{j_1(\ell+1)}(0))$ for $1 \leq \ell \leq k_2$. If $c \in \alpha_n \cap J_{(j_1 k_2,\infty)}$, then $c = a(j_1 k_2,\infty)$. Suppose that $c \in \alpha_n \cap J_{j_1 j_2}$ for some $1 \leq j_2 \leq k_2$. If $\text{card}(\alpha_n \cap J_{j_1 j_2}) = 1$, then $c = a(j_1 j_2)$. If $\text{card}(\alpha_n \cap J_{j_1 j_2}) \geq 2$, proceeding inductively as before, we can find a word $\omega \in \mathbb{N}^*$, such that either $c \in \alpha_n \cap J_\omega$ with $\text{card}(\alpha_n \cap J_\omega) = 1$ implying $c = a(\omega)$, or $c \in \alpha_n \cap J_{(\omega,\infty)}$ with $\text{card}(\alpha_n \cap J_{(\omega,\infty)}) = 1$ implying $c = a(\omega,\infty)$. Thus, the proof of the proposition is complete. □

By Proposition 3.17, we can say that if α_n is an optimal set of n-means for any $n \geq 2$, then the error contributed by any element $c \in \alpha_n$ is given by $E(a(\omega))$ if $c = a(\omega)$, or by $E(a(\omega,\infty))$ if $c = a(\omega,\infty)$, where $\omega \in \mathbb{N}^*$. We are now ready to give the proof of Theorem 3.1.

Proof of Theorem 3.1. By Lemma 3.3 and Lemma 3.5, it is known that the optimal sets of two- and three-means are $\{a(1), a(1,\infty)\}$ and $\{a(1), a(2), a(2,\infty)\}$. Since

$$E(a(1,\infty)) = \frac{43}{3} p_1 s_1^2 V > p_1 s_1^2 V = E(a(1)),$$

the theorem is true for $n = 2$. For $n \geq 2$, let α_n be an optimal set of n-means. Let $\alpha_n := \{a(i) : 1 \leq i \leq n\}$. Let $\tilde{E}(a(i))$ and $W(\alpha_n)$ be defined as in the hypothesis. If $a(j) \notin W(\alpha_n)$, i.e., if $a(j) \in \alpha_n \setminus W(\alpha_n)$, then by Lemma 3.15, the error

$$\sum_{a(i)\in(\alpha_n\setminus\{a(j)\})} E(a(i)) + E(a(\omega^-(\omega_{|\omega|}+1))) + E(a(\omega^-(\omega_{|\omega|}+1), \infty))$$

if $a(j) = a(\omega, \infty)$, or

$$\sum_{a(i)\in(\alpha_n\setminus\{a(j)\})} E(a(i)) + E(a(\omega 1)) + E(a(\omega 1, \infty)) \text{ if } a(j) = a(\omega),$$

obtained in this case is strictly greater than the corresponding error obtained in the case when $a(j) \in W(\alpha_n)$. Hence for any $a(j) \in W(\alpha_n)$, the set $\alpha_{n+1}(a(j))$, where

$$\alpha_{n+1}(a(j)) = \begin{cases} (\alpha_n \setminus \{a(j)\}) \cup \{a(\omega^-(\omega_{|\omega|}+1)),\, a(\omega^-(\omega_{|\omega|}+1), \infty)\} \\ \text{ if } a(j) = a(\omega, \infty), \\ (\alpha_n \setminus \{a(j)\}) \cup \{a(\omega 1),\, a(\omega 1, \infty)\} \text{ if } a(j) = a(\omega), \end{cases}$$

is an optimal set of $(n+1)$-means, and the number of such sets is

$$\text{card}\Big(\bigcup_{\alpha_n\in\mathcal{C}_n} \{\alpha_{n+1}(a(j)) : a(j) \in W(\alpha_n)\}\Big).$$

Thus, the proof of the theorem is complete. □

4. Results and observations about optimal sets of *n*-means

The results and observations of this section are due to the induction formula given by Theorem 3.1.

Recall that the optimal set of one-mean consists of the expected value of the random variable X, and the corresponding quantization error is its variance. Let α_n be an optimal set of n-means, i.e., $\alpha_n \in \mathcal{C}_n$, and then for any $a \in \alpha_n$, we have $a = a(\omega)$ or $a = a(\omega, \infty)$ for some $\omega \in \mathbb{N}^*$. Theorem 3.1 implies that if $\text{card}(\mathcal{C}_n) = k$ and $\text{card}(\mathcal{C}_{n+1}) = m$, then either $1 \leq k \leq m$, or $1 \leq m \leq k$, for example from Figure 2, we see that the number of $\alpha_{15} = 1$, the number of $\alpha_{16} = 3$, the number of $\alpha_{17} = 3$, and the number of $\alpha_{18} = 1$. Thus, there exists a sequence $\{n_k\}_{k=1}^{\infty}$ of positive integers such that for all $n \geq 1$, we have $\text{card}(\mathcal{C}_n) = n_k$, and then we write

$$\mathcal{C}_n = \begin{cases} \{\alpha_n\} & \text{if } n_k = 1, \\ \{\alpha_{n,i} : 1 \leq i \leq n_k\} & \text{if } n_k \geq 2. \end{cases}$$

In addition, Theorem 3.1 implies that a single $\alpha \in \mathcal{C}_n$ can produce multiple distinct $\alpha \in \mathcal{C}_{n+1}$, and multiple distinct $\alpha \in \mathcal{C}_n$ can produce one common $\alpha \in \mathcal{C}_{n+1}$. For $\alpha \in \mathcal{C}_n$, by $\alpha \to \beta$, it is meant that $\beta \in \mathcal{C}_{n+1}$ and β is produced from α. Thus, from Figure 2, we see that

$$\{\alpha_{18} \to \alpha_{19,1}, \alpha_{18} \to \alpha_{19,2}, \alpha_{18} \to \alpha_{19,3}\},$$

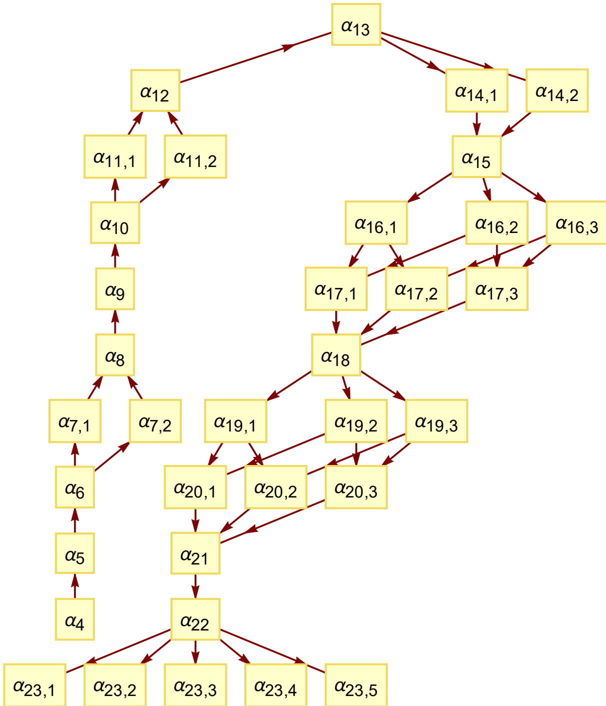

FIGURE 2. Tree diagram of the optimal sets from α_4 to α_{23}.

$$
\begin{aligned}
&\{\{\alpha_{19,1}\to\alpha_{20,1},\alpha_{19,1}\to\alpha_{20,2}\},\{\alpha_{19,2}\to\alpha_{20,1},\alpha_{19,2}\to\alpha_{20,3}\},\\
&\ \{\alpha_{19,3}\to\alpha_{20,2},\alpha_{19,3}\to\alpha_{20,3}\}\},\\
&\{\alpha_{20,1}\to\alpha_{21},\alpha_{20,2}\to\alpha_{21},\alpha_{20,3}\to\alpha_{21}\}.
\end{aligned}
$$

Again, we have

$$
\begin{aligned}
\alpha_{15}&=\{a(111),a(111,\infty),a(12),a(13),a(13,\infty),a(21),a(22),a(23),a(23,\infty),\\
&\quad a(31),a(32),a(32,\infty),a(4),a(5),a(5,\infty)\}\\
&\quad \text{with } V_{15}=\frac{27}{598016}=0.0000451493;\\
\alpha_{16,1}&=\{a(111),a(111,\infty),a(12),a(13),a(13,\infty),a(211),a(211,\infty),\\
&\quad a(22),a(23),a(23,\infty),a(31),a(32),a(32,\infty),a(4),a(5),a(5,\infty)\};\\
\alpha_{16,2}&=\{a(111),a(111,\infty),a(12),a(13),a(13,\infty),a(21),a(22),a(23),a(23,\infty),
\end{aligned}
$$

$$\begin{aligned}
&a(31), a(32), a(32,\infty), a(41), a(41,\infty), a(5), a(5,\infty)\}\\
\alpha_{16,3} &= \{a(111), a(111,\infty), a(121), a(121,\infty), a(13), a(13,\infty), a(21), a(22),\\
&\quad a(23), a(23,\infty), a(31), a(32), a(32,\infty), a(4), a(5), a(5,\infty)\}\\
&\quad \text{with } V_{16} = \frac{4635}{117211136} = 0.000039544;\\
\alpha_{17,1} &= \{a(111), a(111,\infty), a(12), a(13), a(13,\infty), a(211), a(211,\infty), a(22),\\
&\quad a(23), a(23,\infty), a(31), a(32), a(32,\infty), a(41), a(41,\infty), a(5), a(5,\infty)\};\\
\alpha_{17,2} &= \{a(111), a(111,\infty), a(121), a(121,\infty), a(13), a(13,\infty), a(211), a(211,\infty),\\
&\quad a(22), a(23), a(23,\infty), a(31), a(32), a(32,\infty), a(4), a(5), a(5,\infty)\},\\
\alpha_{17,3} &= \{a(111), a(111,\infty), a(121), a(121,\infty), a(13), a(13,\infty), a(21), a(22),\\
&\quad a(23), a(23,\infty), a(31), a(32), a(32,\infty), a(41), a(41,\infty), a(5), a(5,\infty)\}\\
&\quad \text{with } V_{17} = \frac{1989}{58605568} = 0.0000339388;\\
\alpha_{18} &= \{a(111), a(111,\infty), a(121), a(121,\infty), a(13), a(13,\infty), a(211),\\
&\quad a(211,\infty), a(22), a(23), a(23,\infty), a(31), a(32), a(32,\infty),\\
&\quad a(41), a(41,\infty), a(5), a(5,\infty)\}\\
&\quad \text{with } V_{18} = \frac{3321}{117211136} = 0.0000283335;
\end{aligned}$$

and so on.

References

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

[2] D. Çömez and M. K. Roychowdhury, *Quantization for uniform distributions of Cantor dusts on* $\mathbb{R}^2$, Topology Proc. **56** (2020), 195–218.

[3] Q. Du, V. Faber, and M. Gunzburger, *Centroidal Voronoi tessellations: applications and algorithms*, SIAM Rev. **41** (1999), no. 4, 637–676. `https://doi.org/10.1137/S0036144599352836`

[4] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academy publishers: Boston, 1992.

[5] R. M. Gray, J. C. Kieffer, and Y. Linde, *Locally optimal block quantizer design*, Inform. and Control **45** (1980), no. 2, 178–198. `https://doi.org/10.1016/S0019-9958(80)90313-7`

[6] S. Graf and H. Luschgy, *The quantization of the Cantor distribution*, Math. Nachr. **183** (1997), 113–133. `https://doi.org/10.1002/mana.19971830108`

[7] S. Graf and H. Luschgy, *Foundations of quantization for probability distributions*, Lecture Notes in Mathematics, 1730, Springer-Verlag, Berlin, 2000. `https://doi.org/10.1007/BFb0103945`

[8] R. M. Gray and D. L. Neuhoff, *Quantization*, IEEE Trans. Inform. Theory **44** (1998), no. 6, 2325–2383. `https://doi.org/10.1109/18.720541`

[9] A. György and T. Linder, *On the structure of optimal entropy-constrained scalar quantizers*, IEEE Trans. Inform. Theory **48** (2002), no. 2, 416–427. `https://doi.org/10.1109/18.978755`

[10] J. E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30** (1981), no. 5, 713–747. https://doi.org/10.1512/iumj.1981.30.30055

[11] R. D. Mauldin and M. Urbański, *Dimensions and measures in infinite iterated function systems*, Proc. London Math. Soc. (3) **73** (1996), no. 1, 105–154. https://doi.org/10.1112/plms/s3-73.1.105

[12] M. Moran, *Hausdorff measure of infinitely generated self-similar sets*, Monatsh. Math. **122** (1996), no. 4, 387–399. https://doi.org/10.1007/BF01326037

[13] L. Roychowdhury, *Optimal quantization for nonuniform Cantor distributions*, J. Interdiscip. Math. **22** (2019), 1325–1348.

[14] M. K. Roychowdhury, *Quantization and centroidal Voronoi tessellations for probability measures on dyadic Cantor sets*, J. Fractal Geom. **4** (2017), no. 2, 127–146. https://doi.org/10.4171/JFG/47

[15] M. K. Roychowdhury, *Optimal quantization for the Cantor distribution generated by infinite similutudes*, Israel J. Math. **231** (2019), no. 1, 437–466. https://doi.org/10.1007/s11856-019-1859-5

[16] P. L. Zador, *Asymptotic quantization error of continuous signals and the quantization dimension*, IEEE Trans. Inform. Theory **28** (1982), no. 2, 139–149. https://doi.org/10.1109/TIT.1982.1056490

Lakshmi Roychowdhury
School of Mathematical and Statistical Sciences
University of Texas Rio Grande Valley
1201 West University Drive
Edinburg, TX 78539, USA
Email address: lakshmiroychowdhury82@gmail.com

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