# On the Minimum Average Distortion of Quantizers with Index-Dependent Distortion Measures

Erdem Koyuncu, *Member, IEEE*, and Hamid Jafarkhani, *Fellow, IEEE*

***Abstract*—In many applications, one is interested in optimally deploying a network of unidentical agents to a certain area of interest; examples include heterogeneous sensor networks and cellular networks. Such deployment problems can equivalently be formulated as quantizer design problems where different distortion measures should be associated with different quantization indices. In this paper, we consider the case where the distortion measure is the $r$th power $(r \geq 1)$ of the distance between the reproduction point and the source sample weighted by a factor that varies from one index to another. For a uniform distribution of source samples, we determine the corresponding optimal scalar quantizers and their average distortions. For non-uniform distributions and vector quantization, we provide a high-resolution analysis of the minimum possible average distortion. Applications to sensor and cellular networks are also discussed together with numerical design examples.**

## I. Introduction

### A. Index-Dependent Distortion Measures

Let $n$ be a positive integer, $\mathcal{I}_n \triangleq \{1, \ldots, n\}$ be a quantizer index set, $X \triangleq \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$ be a quantizer codebook, and $U \triangleq \{S_1, \ldots, S_n\}$ be a partition of $\mathbb{R}^d$ to Borel measurable quantization cells $S_i$, $i \in \mathcal{I}_n$. For a source sample $q \in \mathbb{R}^d$, we consider an $n$-level $d$-dimensional vector quantizer $(X, U)$ given by the encoder mapping $\alpha(q) \triangleq i$ whenever $q \in S_i$, $i \in \mathcal{I}_n$, and the decoder mapping $\beta(i) \triangleq x_i$, $i \in \mathcal{I}_n$.

Let $r \geq 1$, and $\omega_i$, $i \in \mathcal{I}_n$ be a sequence of positive weights. We associate the distortion measure $h_i(q, x) \triangleq \omega_i \|q - x\|^r$ to quantizer index $i \in \mathcal{I}_n$. Correspondingly, the distortion of quantizing the source sample $q$ using $(X, U)$ can be expressed as $h_{\alpha(q)}(q, \beta(\alpha(q)))$, or equivalently as $\omega_i \|q - x_i\|^r$ whenever $q \in S_i$. The average distortion with $(X, U)$ is then

$$D(X, U) \triangleq \mathrm{E}\left[h_{\alpha(q)}(q, \beta(\alpha(q)))\right] \tag{1}$$

$$= \sum_{i \in \mathcal{I}_n} \omega_i \int_{S_i} \|q - x_i\|^r f(q) \mathrm{d}q, \tag{2}$$

where $\mathrm{E}[\cdot]$ is the expected value, and $f(q)$ is the probability density function of the source random variable $Q$. We assume $\mathrm{E}\|Q\|^r < \infty$ so that $D(X, U)$ is finite for any $(X, U)$.

In the formulation above, the distortion measure varies across different quantization indices. We thus say that the distortion measure is quantization index dependent, or simply, *index-dependent*. In contrast, in most quantization problems, the distortion measure associated to every quantization index is the same: The distortion of quantizing the source sample $q$ via $(X, U)$ is given by $h(q, \beta(\alpha(q)))$ for some unique *index-independent* distortion measure $h$.

The main goal of this paper is to determine the $n \to \infty$ asymptotic performance of optimal quantizers that minimize the average distortion in (2) given the weights $\omega_i$, $i \in \mathcal{I}_n$. In fact, in addition to accomplishing this goal for a general $d$ and $f$, for the special case of one-dimensional uniform source distributions, we will find an optimal quantizer for every $n$.

### B. Multiplicatively-Weighted Voronoi Diagrams

Regarding the minimization of the average distortion in (2), given $i \in \mathcal{I}_n$, let us first define the Voronoi cell

$$V_i(X) \triangleq \{q : \omega_i \|q - x_i\|^r \leq \omega_j \|q - x_j\|^r, \forall j \in \mathcal{I}_n\}. \tag{3}$$

Ties are broken in favor of the smaller index. The tie-breaking rule ensures that each $V_i(X)$ is a Borel set. The collection

$$W(X) \triangleq \{V_i(X) : i \in \mathcal{I}_n\} \tag{4}$$

is commonly referred to as the (multiplicatively) weighted Voronoi diagram [1] generated by $x_i$, $i \in \mathcal{I}_n$ with weights $\omega_i$, $i \in \mathcal{I}_n$. Given any codebook $X$, it can be shown that the Voronoi regions $V_i(X)$, $i \in \mathcal{I}_n$ are the optimal encoding regions in the sense that we have

$$D(X, U) \geq D(X, W(X)), \forall U. \tag{5}$$

A proof of (5) will be provided later on. We thus call $(X, W(X))$ the optimal quantizer given $X$.

In the special equally-weighted case $\omega_1 = \cdots = \omega_n$, the Voronoi cells are always convex polytopes and a given cell $V_i(X)$ can be thought as the intersection of the half-planes $\{q : \|q - x_i\| \leq \|q - x_j\|\}$, $j \neq i$. In general, $V_i(X)$ is the intersection of the sets $\{q : \omega_i \|q - x_i\|^r \leq \omega_j \|q - x_j\|^r\}$, $j \in \mathcal{I}_n - \{i\}$, each of which is either a half-plane (when $\omega_i = \omega_j$), a closed $d$-ball (when $\omega_i > \omega_j$), or the complement of a closed $d$-ball (when $\omega_i < \omega_j$). A two-dimensional weighted Voronoi diagram example is illustrated in Fig. 1. As can be observed in Fig. 1, unlike Voronoi diagrams generated by points with equal weights, the cells of a weighted Voronoi diagram are in general non-convex, non-star-shaped,[1] or even disconnected. This also implies that the corresponding optimal quantizer encoders cannot be

Erdem Koyuncu is with the Department of Electrical and Computer Engineering, University of Illinois at Chicago. Hamid Jafarkhani is with the Center for Pervasive Communications and Computing, University of California, Irvine. E-mails: ekoyuncu@uic.edu, hamidj@uci.edu.

This work was presented in part at the IEEE Data Compression Conference, Mar. 2016.

This work was supported in part by the DARPA GRAPHS program Award N66001-14-1-4061.

[1] A set $\mathbb{R}^d$ is called a star-shaped set if there exists $a \in A$ such that for every point $b \in A$, the line segment that joins $a$ and $b$ is contained in $A$.

implemented using the standard Euclidean nearest-neighbor encoding rule. It is known that optimal entropy-constrained quantizers may similarly have disconnected Voronoi cells in certain cases [2]. Certain structured quantizers such as polar quantizers [3], [4] and spherical quantizers [5] also utilize non-convex quantization cells.

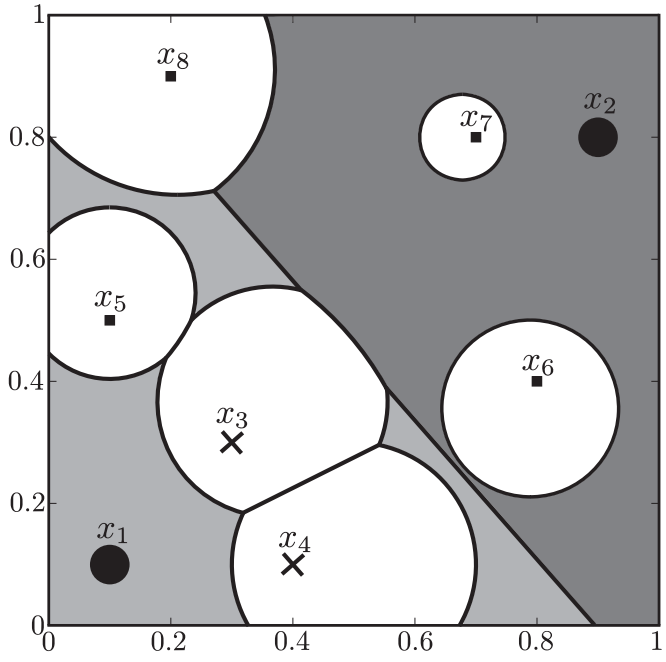


Fig. 1. The optimal encoding regions (Voronoi cells) $W(X)$ corresponding to the reproduction points $x_1 = (0.1, 0.1)$, $x_2 = (0.9, 0.8)$, $x_3 = (0.3, 0.3)$, $x_4 = (0.4, 0.1)$, $x_5 = (0.1, 0.5)$, $x_6 = (0.8, 0.4)$, $x_7 = (0.7, 0.8)$, $x_8 = (0.2, 0.9)$ and weights $\omega_1 = \omega_2 = 1$, $\omega_3 = \omega_4 = 4$, $\omega_5 = \cdots = \omega_8 = 10$, when $r = 2$. The large solid black disks, the crosses, and the squares represent the locations of the reproduction points whose indices have weights 1, 4, and 10, respectively. The light gray region is the Voronoi cell $V_1(X)$ and consists of 3 disconnected components. The darker gray region represents the cell $V_2(X)$, and it is connected but neither convex nor star-shaped. The remaining white regions represent the remaining Voronoi cells.

### C. Related Work and Applications

Index-dependent distortion measures have appeared in the context of a variety of applications. In particular, the corresponding weighted Voronoi diagrams appear in a diverse range of disciplines including crystallography [6], urban planning [7], logistics districting [8], or mesh generation [9]. Minimization of the average distortion in (1) for arbitrary index-dependent distortion measures have been considered in [10] for multi-robot deployment, and a gradient-descent based numerical solution has been provided. The same problem has been studied in [11] for the purpose of locational optimization of facilities having different service capabilities. The Lagrangian formulation of the entropy constrained quantizer design problem [12], [13] also induces an index-dependent distortion measure provided that the codeword lengths are considered fixed. In this case, one typically works with the distortion measure $\widetilde{h}_i(q,x) = \|q-x\|^r + \lambda \ell_i$, where $\lambda$ is a Lagrange multiplier, and $\ell_i$ is the length of the binary codeword that is assigned to quantizer index $i \in \mathcal{I}$. Note that the distortion measure $\widetilde{h}_i(q,x)$ weighs different quantization indices with different additive weights. This is in contrast to the multiplicatively-weighted distortion measure $h_i(q,x) = \omega_i \|q-x\|^r$ that we consider.

Minimizing the average distortion in (2) has been the subject of several works [14]–[18] on heterogeneous sensor networks. The setup consists of $n$ sensors with locations given by $X$. If the sensors are physically identical, one can then quantify the inaccuracy/distortion of sensing an event at point $q$ by Sensor $i$ located at point $x_i$ as $\|q-x_i\|^r$. The overall performance of the sensor network can be modeled via the average distortion of the optimal quantizer given $X$. The aforementioned works [14]–[18] have modeled heterogeneous sensor networks through assigning different multiplicative weights to the $r$th power quality functions of different sensors. In other words, one models the inaccuracy of sensing an event at point $q$ by Sensor $i$ located at point $x_i$ as $\omega_i \|q-x_i\|^r$, where $\omega_i > 0$ depends on the sensor index. In such a formulation, the indices corresponding to the more capable sensors are assigned a lower weight. For example, we can interpret Fig. 1 as a network of 8 sensors located on the sensing field $[0,1]^2$. There are 3 types of sensors with the black disks corresponding to the most capable sensors. The Voronoi cell $V_i(X)$ is the region where Sensor $i$ can provide the minimum sensing distortion among all other sensors. Note that the optimal sensing regions (Voronoi cells) for each sensor may be disconnected.

To gain practical insight on the possibly-disconnected nature of the optimal sensing regions, we may imagine that the 2 large black disks in Fig. 1 correspond to aerial sensors, which have a much longer range of accurate event detection compared to the 6 ground-based sensors represented by crosses and small squares. Clearly, in terms of sensing accuracy, a ground sensor will outperform the aerial sensors in its immediate vicinity. However, as the event location moves further away from a ground sensor, the aerial sensor will begin to outperform, even though the event location remains closer to the ground sensor than the aerial sensor. This results in possibly-disconnected cells for the aerial sensors.

Independent of the particular geometry of the optimal sensing regions, given any set of weights $\{\omega_i : i \in \mathcal{I}_n\}$, the goal of the above heterogeneous sensor network formulation is to find the optimal sensor locations such that the overall sensing distortion $D(X, W(X))$ is minimized. In this context, we note that for practical sensor network applications, the heterogeneity parameters ($\omega_i$s) for each sensor can be determined empirically through measurements, as done in [19], [20]. Experimental results concerning the weighted $r$th power sensing quality model are also available [15], [17]–[20]. It is also worth mentioning that, by utilizing an appropriate distortion measure, many other practical sensor network deployment problems (e.g. to achieve high network energy efficiency [21] or network connectivity [22]) can be formulated and solved by using quantization theory.

The cost function in (2) also appears in the context of cellular networks. In this case, the formulation is as follows: We consider $n$ base stations whose locations are given by $X$, and suppose that a mobile terminal at location $q$ transmits with signal-to-noise ratio (SNR) $\Upsilon$. Under the well-known log-

distance wireless path loss model [23, Ch. 1.3], the received SNR at Base Station $i$ is given by $\|q - x_i\|^{-r}\Upsilon$, where $r$ is called the path loss exponent. In general, Base Station $i$ can successfully decode the message(s) of the mobile terminal if its received SNR is greater than a certain threshold $\omega_i$. The threshold $\omega_i$ depends on the specific capabilities of Base Station $i$. For example, when the base station has multiple antennas, its threshold will be lower. Regardless, for successful data reception at Base Station $i$, the transmit SNR of the mobile terminal should thus satisfy $\Upsilon \geq \omega_i\|q - x_i\|^r$. Hence, the minimum transmit SNR that will guarantee successful reception at at least one of the base stations is $\min_{i\in\mathcal{I}_n} \omega_i\|q - x_i\|^r$. The cost function in (2) then corresponds to the average minimum transmit SNR that will ensure successful data transmission at every possible location of the mobile terminal.

For a concrete choice of weights for the cellular network scenario, suppose that the mobile terminal is equipped with only one antenna, and Base Station $i$ is equipped with $N_i \geq 1$ antennas. The capacity of the channel between the mobile terminal and Base Station $i$ is then $\log_2(1 + N_i\Upsilon\|x_i - q\|^{-r})$ bits/sec/Hz [24]. To support a fixed target rate $\xi > 0$ of data transmission from the mobile terminal, we need $\log_2(1 + N_i\Upsilon\|x_i - q\|^{-r}) \geq \xi$, or equivalently, $\Upsilon \geq (2^\xi - 1)/N_i\|x_i - q\|^r$. For this scenario, we shall thus set $\omega_i = (2^\xi - 1)/N_i$.

The design and analysis of quantizers for index-dependent distortion measures is thus a fundamental problem with potentially many applications, especially to heterogeneous sensor networks [14]–[18] and cellular networks as discussed above. On the other hand, to the best of our knowledge, all previous work in this context rely on methods such as the Lloyd algorithm [25] or its appropriate generalizations. One of the main goals of the present work is to obtain analytical expressions that characterize the performance of optimal quantizers.

In the case of the $r$th power distortion measure, there is a vast literature on the asymptotic performance of optimal quantizers [26]. In particular, under certain technical conditions on $f$, it is known that in the high-resolution regime $n \to \infty$, the best possible average distortion is [27]–[29]

$$\kappa_{rd} n^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}} + o(n^{-\frac{r}{d}}), \tag{6}$$

where

$$\|f\|_\gamma \triangleq \left(\int_{\mathbb{R}^d} f^\gamma(q)\mathrm{d}q\right)^{\frac{1}{\gamma}}, \ \gamma \in \mathbb{R}, \tag{7}$$

and $\kappa_{rd} > 0$ is a positive real number that depends only on the power $r$ and the dimension $d$. Regarding $\kappa_{rd}$, the widely-accepted Gersho's conjecture [30] states that at high resolution, the cells of an optimal quantizer should be congruent to each other, in which case $\kappa_{rd}$ becomes the normalized $r$th moment of the best tessellating polytope in $\mathbb{R}^d$ (In fact, it is further believed [31] that the best tessellation is generated by a lattice, which means that the best quantizer should locally look like a lattice quantizer). This is so far known to hold for dimensions 1 and 2, in which case the optimal cell shapes are intervals and regular hexagons, respectively. For the example case of squared-error distortion measure ($r = 2$), we can thus calculate $\kappa_{21} = \frac{1}{12}$, and $\kappa_{22} = \frac{5}{18\sqrt{3}}$. For $d \geq 3$, a useful lower bound is the ball-packing bound $\kappa_{rd} \geq m_{rd}$, where

$$m_{rd} \triangleq \frac{\int_{|q|\leq 1} |q|^r \mathrm{d}q}{(\int_{|q|\leq 1} \mathrm{d}q)^{\frac{d+r}{d}}} = \frac{d}{\pi^{\frac{r}{2}}(d + r)} \left[\Gamma\left(1 + \frac{d}{2}\right)\right]^{\frac{r}{d}} \tag{8}$$

is the normalized $r$th moment of the ball in $\mathbb{R}^d$, and $\Gamma(\cdot)$ is the Gamma function.

### D. Main Contributions

We generalize the existing formula (6) for the $r$th power distortion measure to the $r$th power distortion measure with index-dependent weights. To introduce this result, let $\mathcal{T} = \{1, \ldots, |\mathcal{T}|\}$ be a non-empty finite index set and suppose that all the weights $\omega_i$, $i \in \mathcal{I}_n$ are chosen among the finite set $\{\Omega_t : t \in \mathcal{T}\}$ of positive real numbers. Also suppose that for each $t \in \mathcal{T}$, the fraction of indices with weight $\Omega_t$ converges to some frequency $p_t$ as $n \to \infty$. Let $\boldsymbol{\Omega} = [\Omega_1 \cdots \Omega_{|\mathcal{T}|}]$ and $\mathbf{p} = [p_1 \cdots p_{|\mathcal{T}|}]$ denote the vector of weights and frequencies, respectively. We show that the minimum average distortion in this case is given by

$$\ell_{rd}(\boldsymbol{p}, \boldsymbol{\Omega}) n^{-\frac{r}{d}} \left(\sum_{t\in\mathcal{T}} p_t \Omega_t^{-\frac{d}{r}}\right)^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}} + o(n^{-\frac{r}{d}}), \tag{9}$$

where $\ell_{rd}(\boldsymbol{p}, \boldsymbol{\Omega})$ is a constant that satisfies $m_{rd} \leq \ell_{rd}(\boldsymbol{p}, \boldsymbol{\Omega}) \leq \kappa_{rd}$ and is independent of $f$. In particular, we have $\ell_{rd}(1, 1) = \kappa_{rd}$ for any $(r, d)$ for the special case of an index-independent $r$th power distortion measure $\boldsymbol{p} = \boldsymbol{\Omega} = 1$.

For scalar quantization ($d = 1$), since $\kappa_{r1} = m_{r1}, \forall r$, we have $\ell_{r1}(\boldsymbol{p}, \boldsymbol{\Omega}) = m_{r1}$ for any $(\boldsymbol{\Omega}, \mathbf{p}, r)$, and thus (9) provides a precise characterization of the minimum achievable average distortion. On the other hand, as $m_{rd} < \kappa_{rd}$ for vector quantization, there is a multiplicative uncertainty regarding the best average distortion for $d > 1$. In this context, we note that for the ordinary index-independent scenario with $\mathbf{p} = 1$ and $\boldsymbol{\Omega} = \omega$ for some constant $\omega > 0$, we have $\ell_{rd}(1, c) = \kappa_{rd}$. In general, for certain values of $\boldsymbol{p}$ and $\boldsymbol{\Omega}$, we show that the strict inequality $\ell_{rd}(\boldsymbol{p}, \boldsymbol{\Omega}) < \kappa_{rd}$ holds. In particular, $\ell_{rd}(\boldsymbol{p}, \boldsymbol{\Omega})$ can be arbitrarily close to $m_{rd}$ when one allows a sufficiently large number of distinct weights $|\mathcal{T}|$. An intuitive explanation as to why this happens is the following: Cells corresponding to larger weights should be smaller than cells with smaller weights so that each cell still contributes the same amount to the average distortion. This gives the opportunity of "squeezing in" quantization cells of large-weighted indices in between those of smaller-weighted indices. As a result, the shape of each quantization cell can approach to that of a ball, and this allows one to approach the ball packing bound $m_{rd}$.

The focus of this paper will be only on fixed-rate quantization as a fixed-rate formulation is sufficient to address the practical deployment problems in heterogeneous networks. Extensions of our results to variable-rate quantization will remain as an interesting future research direction.

We note that part of this work has been presented in [32]. Compared to [32], the current paper in addition proves the existence of optimal quantizers for index-dependent distortion measures. It also provides a formal proof of the high-resolution formula in (9), which was proved heuristically in [32]. Also,

the results of [32] were for the special case of squared-error distortion $r = 2$. In this work, we consider any $r \geq 1$.

### E. Comparison to Perceptual Distortion Measures

Before proceeding any further, it is instructive to compare our formulation and results with those with perceptual distortion measures [33]–[37], which can similarly weigh the squared distance between the source sample and the reproduction point by different factors. These distortion measures can be defined as $h(q, y) = (q-y)^T B(q, y)(q-y)$, where $B(q, y)$ is a positive-definite weighting matrix that may depend only on the source sample $q$ and the reproduction point $y$. For the special case $B(q, y) = b(q, y)\mathbf{I}_{d\times d}$ for some function $b : \mathbb{R}^2 \to \mathbb{R}$, we have $h(q, y) = b(q, y)\|q-y\|^2$. This resembles, for $r = 2$, the distortion measure $\omega_i\|q - x_i\|^2$ for the $i$th quantization index with our index-dependent formulation. On the other hand, with a perceptual distortion measure, the multiplicative weight involved in reproducing the source sample $q$ via a reproduction point $y$ is the fixed number $b(q, y)$. With an index-dependent distortion measure, the multiplicative weight of reproducing $q$ via $y$ may be any weight among the set of weights $\{\omega_1, \ldots, \omega_n\}$. In other words, when designing an optimal quantizer for an index-dependent distortion measure, one has the extra degree of freedom of choosing among many different weights for a $(q, y)$-pair instead of being limited to one pre-determined weight.

### F. Organization of the Paper

The rest of the paper is organized as follows. In Section II, we study the design and performance of optimal scalar quantizers for our multiplicatively-weighted index-dependent distortion measures and a uniform distribution. We analyze the general case of vector quantizers with non-uniform distributions in Section III. In Section IV, we present several numerical design examples for our index-dependent distortion measure. Finally, in Section V, we draw our main conclusions. Some of the technical proofs are provided in the appendices.

## II. Scalar Quantization of a Uniform Source

We begin with the design and analysis of optimal scalar quantizers with an index-dependent distortion measure. For simplicity, we first consider the case of a uniform distribution $f(x) = \mathbf{1}(x \in [0, 1))$, where $\mathbf{1}(\cdot)$ is the indicator function.

We need the following standard definitions and results that will be useful in estimating the average distortion of optimal quantizers: Let $\mu(A)$ denote the $d$-dimensional Lebesgue measure (volume) of a measurable set $A \subset \mathbb{R}^d$. The quantity

$$M_{rd}(A) \triangleq \frac{\int_A \|q\|^r \mathrm{d}q}{(\mu(A))^{\frac{d+r}{d}}} \tag{10}$$

is usually referred to as the normalized $r$th moment of $A$. In particular, for the unit ball in $\mathbb{R}^d$ at the origin, we have $m_{rd} = M_{rd}(\{q \in \mathbb{R}^d : \|q\| \leq 1\})$.

Consider now the quantizer $(X, U)$. Throughout the paper, we will assume $\mu(S_i) > 0, \forall i \in \mathcal{I}_n$ (Quantization cells with measure zero have no contribution towards the average distortion and thus can be ignored.). The following theorem then provides a lower bound on the average distortion of any quantizer with our index-dependent distortion measure. The proof of the theorem can be found in Appendix A.

**Theorem 1.** *Let $f(q) = \mathbf{1}(q \in [0, 1)^d)$. Then, for any quantizer $(X, U)$, we have*

$$D(X, U) \geq m_{rd} \left( \sum_{i\in\mathcal{I}_n} \omega_i^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \tag{11}$$

*with equality if and only if for every $i \in \mathcal{I}_n$, the quantization region $S_i$ is a ball centered at $x_i$ with*

$$\mu(S_i) = \frac{\omega_i^{-\frac{d}{r}}}{\sum_{j\in\mathcal{I}_n} \omega_j^{-\frac{d}{r}}}. \tag{12}$$

For the special case of $d = 1$, the theorem immediately leads to the following corollary.

**Corollary 1.** *Let $d = 1$ with the uniform distribution $f(q) = \mathbf{1}(q \in [0, 1))$. Then, the minimum possible average distortion given weights $\omega_i$, $i \in \mathcal{I}_n$ is provided by*

$$m_{r1} \left( \sum_{i\in\mathcal{I}_n} \omega_i^{-\frac{1}{r}} \right)^{-r}, \tag{13}$$

*where $m_{r1} = \frac{1}{(r+1)2^r}$. Such a performance is achieved by any quantizer $(X, U)$ of the form*

$$S_{c_i} = \frac{1}{\sum_{j\in\mathcal{I}_n} \omega_j^{-\frac{1}{r}}} \left[ \sum_{k=1}^{i-1} \omega_{c_k}^{-\frac{1}{r}}, \sum_{k=1}^{i} \omega_{c_k}^{-\frac{1}{r}} \right) \tag{14}$$

*with*

$$x_{c_i} = \frac{1}{\sum_{j\in\mathcal{I}_n} \omega_j^{-\frac{1}{r}}} \left( \sum_{k=1}^{i-1} \omega_{c_k}^{-\frac{1}{r}} + \frac{1}{2}\omega_{c_i}^{-\frac{1}{r}} \right), \tag{15}$$

*where $c_k$, $k \in \mathcal{I}_n$ are arbitrary indices with $\{c_k : k \in \mathcal{I}_n\} = \mathcal{I}_n$.*

In particular, for the equally-weighted case $\omega_i = 1$, $\forall i \in \mathcal{I}_n$ and the squared-error distortion measure $r = 2$, we obtain the well-known minimum average distortion formula $\frac{1}{12n^2}$. This performance is achieved by the unique $n$-level uniform quantizer on $[0, 1)$. In general, the corollary shows that the length of the quantizer cell corresponding to a weight-$\omega$ index is proportional to $\omega^{-\frac{1}{r}}$. Hence, optimal quantizer cells are intervals of different lengths in general, and thus there are many optimal quantizers corresponding to the different orderings of these cells. As an example, for $r = 2$, $\mathcal{I}_8 = \{1, \ldots, 8\}$ with weights $\omega_1 = \omega_2 = 16$, $\omega_3 = \omega_4 = 64$ and $\omega_5 = \cdots = \omega_8 = 256$, the minimum possible average distortion can be calculated to be $\frac{1}{12}$ by Corollary 1. According to Corollary 1, both quantizers in Fig. 2 achieve this performance.

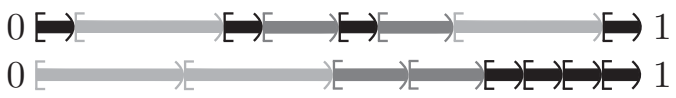


Fig. 2. Two optimal quantizers for the unit interval $[0, 1)$ when $r = 2$. Each quantization cell is a half-open interval. Length-$\frac{1}{4}$ light gray intervals, length-$\frac{1}{8}$ dark gray intervals and length-$\frac{1}{16}$ black intervals are the quantization cells for reproduction points whose indices have weights 4, 16, and 64, respectively. The reproduction point for any given interval is located at its center.

## III. Quantization of Non-Uniform Vector Sources

We now study the performance of optimal quantizers for the general case of a non-uniform vector source. We begin by deriving some of the elementary properties of optimal quantizers for index-dependent distortion measures (such as existence) that hold at every resolution $n$. Some of these results will be useful in our subsequent analysis of the high-resolution ($n \to \infty$) analysis of the quantizers.

### A. Elementary Properties

Let us first show that, as claimed in Section I-B, the average distortion $D(X, U)$ in (2) is indeed minimized by choosing the encoding cells $U$ to be the Voronoi cells $W(X)$ given $X$.

**Proposition 1.** *For any $X$ and $U$, we have $D(X, U) \geq D(X, W(X))$.*

The proof of the proposition can be found in Appendix B. As a result of the optimality of utilizing the collection of cells $W(X)$ given the codebook $X$, we may define

$$D_n(X) \triangleq D(X, W(X)) = \mathrm{E}\Big[\min_{i \in \mathcal{I}_n} \omega_i \|Q - x_i\|^r\Big] \tag{16}$$

to be the minimum possible average distortion given $X$. The expectation is over the source random variable $Q$. We also let

$$D_n^\star = \inf_{X \in \mathbb{R}^{dn}} D_n(X). \tag{17}$$

denote the infimum of all achievable average distortions. We write $P(A)$ as a shorthand notation for $\int_A f(q)\mathrm{d}q$.

The following proposition shows that if $D_n(X) = D_n^\star$ and thus if $X$ is an optimal collection of reproduction points, then the corresponding Voronoi regions should all have non-zero probability, and in particular, $x_i$, $i \in \mathcal{I}_n$ should all be distinct. The proof is omitted as it is a straightforward extension of the proof of the same result for index-independent distortion measures [39, Theorem 4.1].

**Proposition 2.** *Suppose $D_n(X) = D_n^\star$ for some $X \in \mathbb{R}^{dn}$. For any $i \in \mathcal{I}_n$, we have $P(V_i(X)) > 0$. Consequently, for any $i, j \in \mathcal{I}_n$ with $i \neq j$, we have $x_i \neq x_j$.*

On the other hand, the existence of an optimal collection of reproduction points is guaranteed by the following theorem. Meanwhile, the theorem shows that an additional reproduction point will provide a strictly lower average distortion regardless of its weight. The proof of the theorem can be found in Appendix C.

**Theorem 2.** *The following holds.*
*(a) There exists $X \in \mathbb{R}^{dn}$ such that $D_n(X) = D_n^\star$.*
*(b) For any $n \geq 2$, we have $D_n^\star < D_{n-1}^\star$.*

Note that the conclusions of Proposition 2 and Theorem 2 already follow from Corollary 1 for the special case of scalar quantization of a uniform source.

### B. High-Resolution Average Distortion of Optimal Quantizers

We now analyze the average distortion of optimal quantizers for index-dependent distortion measures. The existence of such quantizers is guaranteed by Theorem 2.

Even if one considers an ordinary index-independent quantization scenario, determining the best $n$-level quantizer for every $n$ is a very difficult task. Hence, most of the existing work have focused on the $n \to \infty$ regime, which is also known as the high-resolution regime or the asymptotic regime; see [31, Ch. 5.6] for an introduction to high-resolution quantization theory. In this case, the minimum asymptotic average distortion for the $r$th power index-independent distortion measure is given by the following theorem. For sequences of positive real numbers $g(n), h(n)$, we write $g(n) \sim h(n)$ whenever $\lim_{n\to\infty}(g(n)/h(n))$ exists and is equal to 1.

**Theorem 3.** *Suppose $\mathrm{E}[\|Q\|^{r+\epsilon}] < \infty$ for some $\epsilon > 0$, and $\omega_i = 1$, $\forall i \in \mathcal{I}_n$. Then,*

$$D_n^\star \sim \kappa_{rd} n^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}}, \tag{18}$$

*for some constant $\kappa_{rd} > 0$ that is independent of $f$.*

A proof can be found in [39, Theorem 6.2]. Under different technical conditions, the theorem is originally due to Bennett [27] for the special case of $d = 1$ and Zador [28] for a general $d$. The condition $\mathrm{E}[\|Q\|^{r+\epsilon}] < \infty$ ensures that $\|f\|_{\frac{d}{d+r}}$ is finite.

We consider a similar high-resolution analysis for $r$th power distortion measure with index-dependent weights. Let $\mathcal{T}$ be a finite index set, and $\{\Omega_t : t \in \mathcal{T}\}$ be a set of positive real numbers. Suppose that for every $i \in \mathcal{I}_n$, we have

$$\omega_i \in \{\Omega_t : t \in \mathcal{T}\}. \tag{19}$$

Moreover, suppose that for every $t \in \mathcal{T}$, there exists a non-negative frequency $p_t \geq 0$ such that

$$|\{i \in \mathcal{I}_n : \omega_i = \Omega_t\}| \sim p_t n, \tag{20}$$

In other words, according to (19), we assume that the weight $\omega_i$ of the quantization index $i \in \mathcal{I}_n$ is to be chosen among the finite set $\{\Omega_t : t \in \mathcal{T}\}$ of weights for every $i \in \mathcal{I}_n$. Also, according to (20), we assume that the weight for some $p_t$-fraction of the indices is $\Omega_t$ for every $t \in \mathcal{T}$, asymptotically as $n \to \infty$. For example, for $\mathcal{T} = \{1, 2\}$, $\Omega_1 = 2$, $\Omega_2 = 3$, $p_1 = \frac{3}{4}$, $p_2 = \frac{1}{4}$, we would consider a scenario where, asymptotically as $n \to \infty$, three-quarters of the indices have weight 2, and one quarter of the indices have weight 3. In practice, the example may also correspond to a sensor network with two types of sensors: The sensing distortion for three-quarters of the sensors is weighted by 2, and the sensing distortion for a quarter of the sensors is weighted by 3.

Note that in (19) and (20), we do not need to explicitly specify the exact weight that is associated to a given index. This is because, by definition, for any given sequence of

weights $\omega_i'$, $i \in \mathcal{I}_n$, the optimal quantizer performance for $\omega_i = \omega_i'$, $i \in \mathcal{I}_n$ is the same as the optimal quantizer performance for $\omega_i = \omega_{j_i}'$, $i \in \mathcal{I}_n$, where $(j_1, \ldots, j_n)$ is an arbitrary permutation of $(1, \ldots, n)$. For example, for $n = 3$, the minimum average distortion for $\omega_1 = 3$, $\omega_2 = 4$, $\omega_3 = 2$ is the same as the minimum average distortion for $\omega_1 = 4$, $\omega_2 = 2$, $\omega_3 = 3$.

We can now state and prove our main result on the minimum asymptotic distortion of optimal vector quantizers with weighted $r$th power distortion measures. Let $\mathbf{\Omega} = [\Omega_1 \cdots \Omega_{|\mathcal{T}|}]$ and $\mathbf{p} = [p_1 \cdots p_{|\mathcal{T}|}]$.

**Theorem 4.** *Suppose $E[\|Q\|^{r+\epsilon}] < \infty$ for some $\epsilon > 0$, and $\omega_i$, $i \in \mathcal{I}_n$ are such that (19) and (20) hold. Then,*

$$D_n^\star \sim \ell_{rd}(\boldsymbol{p}, \mathbf{\Omega}) n^{-\frac{r}{d}} \left( \sum_{t \in \mathcal{T}} p_t \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}}, \tag{21}$$

*for some constant $\ell_{rd}(\boldsymbol{p}, \mathbf{\Omega})$ that satisfies*

$$m_{rd} \leq \ell_{rd}(\boldsymbol{p}, \mathbf{\Omega}) \leq \kappa_{rd} = \ell_{rd}(1, 1) \tag{22}$$

*and is independent of $f$.*

*Proof.* The proof is provided in Appendix D. □

Let us now discuss the implications of the theorem, and in particular the dependence of the constant $\ell_{rd}(\boldsymbol{p}, \mathbf{\Omega})$ on $(\mathbf{p}, \mathbf{\Omega})$.

For scalar quantization ($d = 1$), we have

$$m_{r1} = \kappa_{r1} = \frac{1}{(1+r)2^r}, \tag{23}$$

and thus $\ell_{r1}(\boldsymbol{p}, \mathbf{\Omega}) = \frac{1}{(1+r)2^r}$ is independent of $(\mathbf{p}, \mathbf{\Omega})$. On the other hand, for vector quantization ($d > 1$), we have $m_{rd} < \kappa_{rd}$, and it is no more trivial to determine the exact value of $\ell_{rd}(\boldsymbol{p}, \mathbf{\Omega})$. In this context, the equality $\ell_{rd}(\boldsymbol{p}, \mathbf{\Omega}) = k_{rd}$ would always have hold if the optimal cell shapes for index-dependent distortion measures had matched those of index-independent distortion measures.[2] Unfortunately, this is not the case and $\ell_{rd}(\boldsymbol{p}, \mathbf{\Omega})$ varies with $(\mathbf{p}, \mathbf{\Omega})$ and may be strictly less than $\kappa_{rd}$.

A first hint towards proving this phenomenon is the observation that one can cover the unit cube $[0, 1)^d$ (up to sets of measure zero) using infinitely many open balls $\mathbf{B} \triangleq \{B_i : i \in \mathbb{N}\}$ with radii $\rho_i$, $i \in \mathbb{N}$, centers $\mathbf{y} \triangleq \{y_i : i \in \mathbb{N}\}$, and $B_i \subset [0, 1)^d$ (via, e.g., an Apollonian ball packing [38]). Then, given weights $\omega_i = \rho_i^{-r}$, $i \in \mathbb{N}$, the quantizer $(\mathbf{y}, \mathbf{B})$ achieves the ball packing bound in Theorem 1, i.e. $D(\mathbf{y}, \mathbf{B}) = m_{rd} (\sum_{i \in \mathcal{I}_n} {\omega_i}^{-\frac{d}{r}})^{-\frac{r}{d}}$. Of course, such a quantizer is not very relevant for practical applications as it has infinitely many quantization indices with infinitely many different weights. On the other hand, it provides the following valuable intuition: Given weights $\mathbf{\Omega}$ with frequencies $\mathbf{p}$, dense packings of balls with radii $\Omega_t^{-1/r}$, $t \in \mathcal{T}$ and frequencies $\mathbf{p}$ should correspond to good quantizers for index-dependent distortion measures. We verify this intuition via the following example scenario in a two-dimensional space.

**Example 1.** *Let $r = 2$. Consider weights $\Omega_1 = 1$ and $\Omega_2 = (3 + 2\sqrt{3})^2$ appearing according to frequencies $p_1 = \frac{1}{3}$ and $p_2 = \frac{2}{3}$, respectively. Correspondingly, we seek dense packings of disks of radius $\rho_1 = {\Omega_1}^{-\frac{1}{2}} = 1$ and $\rho_2 = {\Omega_2}^{-\frac{1}{2}} = 1/(3 + 2\sqrt{3})$ with frequencies $\frac{1}{3}$ and $\frac{2}{3}$, respectively. One packing is illustrated in Fig. 3(a) and has density $\pi/\sqrt{12}$. The corresponding quantizer structure is as shown in Fig. 3(b). For the uniform distribution on $[0, 1)^2$, it achieves an average distortion of $\frac{\kappa_{22}}{n}(\frac{p_1}{\Omega_1} + \frac{p_2}{\Omega_2})^{-1} + o(n^{-1})$. This is also the minimum average distortion if one assumes the quantization cells are all congruent to hexagons. On the other hand, one of the best packings is illustrated in Fig. 3(c) and achieves the density $\frac{\pi}{2\sqrt{3}}(1 + \frac{1}{(3+2\sqrt{3})^2})$ (The smaller circles fit exactly in between the larger circles.). The induced quantizer is as shown in Fig. 3(d). The corresponding average distortion can be calculated to be $\frac{\ell}{n}(\frac{p_1}{\Omega_1} + \frac{p_2}{\Omega_2})^{-1} + o(\frac{1}{n})$, where $\ell = 0.159496\cdots$ with $m_{22} = 0.159154\cdots < \ell < \kappa_{22} = 0.160375\cdots$. We omit the details of the tedious but straightforward calculation of $\ell$. Note that there are two different Voronoi cell shapes in Fig. 3(d). The cell shape for indices with weight $\Omega_2$ is a truncated hexagon and is shown in Fig. 3(e). On the other hand, the cell shape for indices with weight $\Omega_1$ resembles a Reuleaux triangle and is shown in Fig. 3(f). The relative size of the cells in Figs. 3(e) and 3(f) are not to scale.*

The fact that $\ell < \kappa_{22}$ in Example 1 suggests that a natural extension of Gersho's conjecture [30] does not hold for our index-dependent distortion measure. For the $r$th power distortion measure and a uniform distribution on $[0, 1]^d$, Gersho conjectures the existence of a tessellation $T$ with congruent cells such that $T$ provides an asymptotically optimal quantizer with appropriate normalization. As detailed in [32], for the setup in Example 1, forcing a similar quantizer cell congruency results in $\ell = \kappa_{22}$, contradicting $\ell < \kappa_2$. Hence, for certain weights/frequencies, a tessellation of non-congruent quantization cells can outperform tessellations of congruent cells.

Despite the fact that the quantizer in Fig. 3(d) is a non-lattice quantizer with non-congruent cells, it is nevertheless highly-structured. In fact, it can be viewed as a two-level tree-structured lattice vector quantizer (TSLVQ) [40]: The first level is the hexagonal lattice quantizer induced by the indices with weight $\Omega_1$, and the second level consists of one index with weight $\Omega_1$, and six "surrounding" indices with weight $\Omega_2$. Investigating the optimality of multi-level TSLVQs for index-dependent distortion measures in general will remain as an interesting direction for future research.

We now discuss whether the lower bound $\ell_{rd} > m_{rd}$ on $\ell_{rd}$ is tight. In this context, Example 1 shows one particular instance where $\ell_{rd} < \kappa_{rd}$. In fact, $\ell_{rd}$ can be arbitrarily close to $m_{rd}$. This can be proved by considering ball packings with high enough density.

[2] To see this, first note that the normalized $r$th moment of the cells of an optimal quantizer would then all be equal to $\kappa_{rd}$. As a result, for the case of a uniform distribution, the minimum average distortion would be

$$D_n^\star = \kappa_{rd} \sum_{i \in \mathcal{I}_n} \omega_i (\mu(S_i))^{\frac{d+r}{d}} \geq \kappa_{rd} \left( \sum_{i \in \mathcal{I}_n} {\omega_i}^{-\frac{d}{r}} \right)^{-\frac{r}{d}}, \tag{24}$$

where the lower bound follows from Lemma 2-(ii) and is achievable by choosing $\mu(S_i) = c {\omega_i}^{-d/r}$ for an appropriate constant $c$. This would imply $\ell_{rd}(\boldsymbol{p}, \mathbf{\Omega}) \geq k_{rd}$. On the other hand, by Theorem 4, we have $\ell_{rd}(\boldsymbol{p}, \mathbf{\Omega}) \leq k_{rd}$. Combining the two bounds, we would obtain $\ell_{rd}(\boldsymbol{p}, \mathbf{\Omega}) = k_{rd}$.

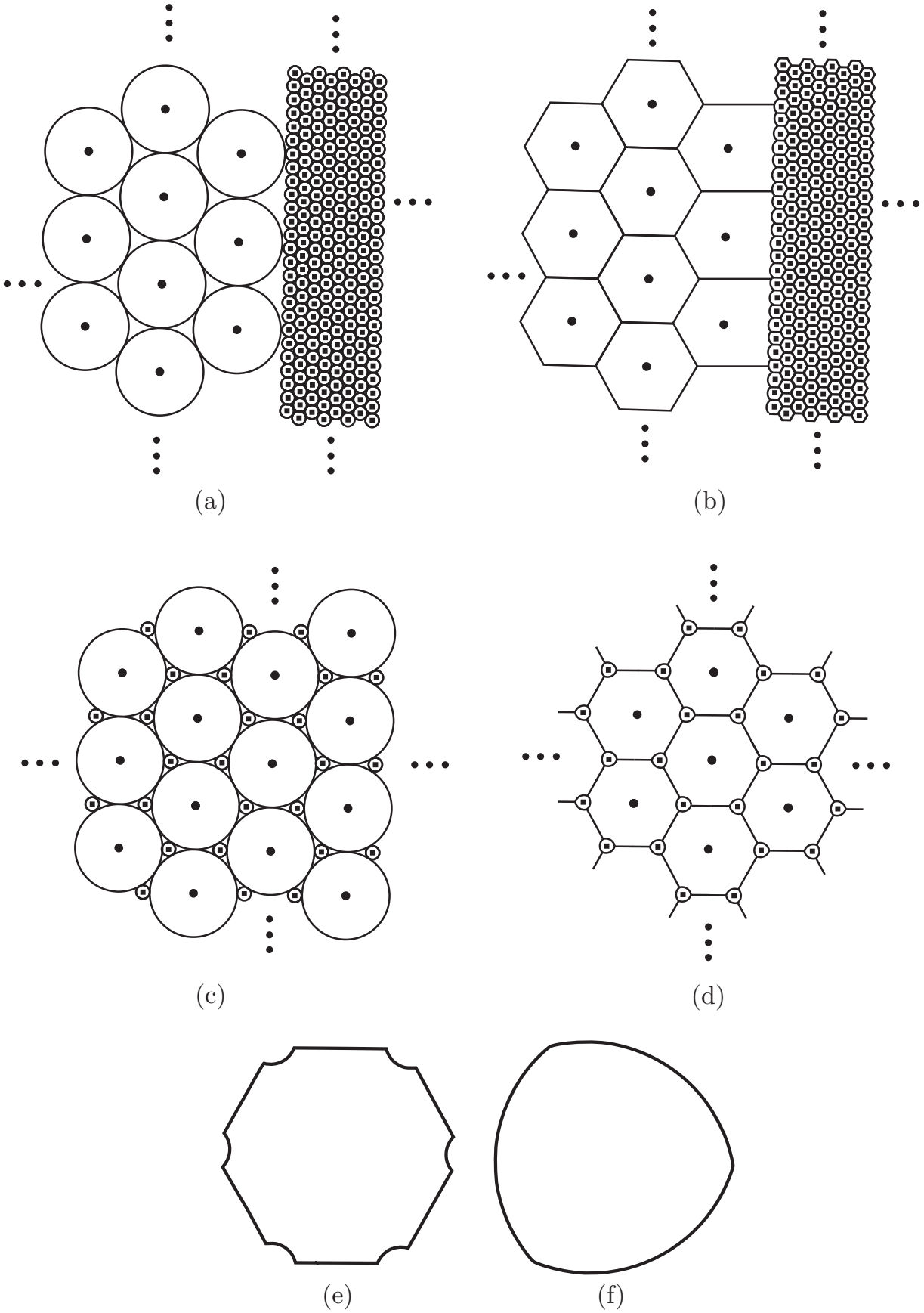


Fig. 3. Packings of non-identical disks and the corresponding quantizers for index-dependent distortion measures.

**Proposition 3.** *For a given set of weights $\omega_1, \ldots, \omega_N$, where $N \geq 1$, let $v_i = \omega_i^{-\frac{d}{r}} / \sum_{j=1}^{N} \omega_j^{-\frac{d}{r}}$, $i \in \{1, \ldots, N\}$. Suppose that there is a set of reproduction points $x_1, \ldots, x_N \in [0,1)^d$ such that the following holds for some $\delta \in (0,1)$ and $A > 1$.*

*(i) The open balls centered at $x_i$ with volume $(1-\delta)v_i$ are mutually disjoint and members of $[0,1)^d$.*

*(ii) For every $q \in [0,1)^d$, there is an index $i \in \{1, \ldots, N\}$ such that $\|q - x_i\| \leq A v_i^{\frac{1}{d}}$.*

*Then, for the optimal quantizer given $x_1, \ldots, x_N$ with weights $\omega_1, \ldots, \omega_N$, we have*

$$D(X, W(X)) \leq \left( m_{rd}(1-\delta)^{\frac{d+r}{d}} + \delta A^r \right) \left( \sum_{k=1}^{N} \omega_k^{-\frac{d}{r}} \right)^{-\frac{r}{d}}. \quad (25)$$

*In particular, for $\mathcal{T} = \{1, \ldots, N\}$ with $\Omega_i = \omega_i$, $p_i = \frac{1}{N}$, $i = 1, \ldots, N$ and $f(q) = \mathbf{1}(q \in [0,1)^d)$,*

$$\ell_{rd}(\mathbf{p}, \mathbf{\Omega}) \leq m_{rd}(1-\delta)^{\frac{d+r}{d}} + \delta A^r. \quad (26)$$

The proof of the proposition can be found in Appendix E. The proposition shows that a ball packing with density $1-\delta$ translates to a certain set of weights and frequencies where (26) holds. As $\delta \to 0$, the upper bound in (26) approaches $m_{rd}$, proving that $\ell_{rd}(\mathbf{p}, \mathbf{\Omega})$ can be arbitrarily close to $m_{rd}$.

## IV. Numerical Design Examples

We now study how the results of this paper can be applied to the design of heterogeneous networks through numerical examples. As a case study, we consider the heterogeneous sensor network application as described in Section I-C. We recall that, in such a scenario, one is given the task of optimally deploying $n$ heterogeneous sensors to some field of interest $S \subset \mathbb{R}^d$. Given an event distribution $f$ over the sensing field $S$, the inaccuracy of sensing an event that occurs at location $x \in S$ via Sensor $i$ is modeled by the distortion measure $\omega_i \|q - x_i\|^r$, where $\omega_i$ is the weight that indicates the sensing quality of Sensor $i$. The goal is to find an optimal deployment of sensors $X = [x_1 \cdots x_n]$ such that the overall sensing inaccuracy of the network $D(X, W(X)) = E[\min_{1 \leq i \leq n} \omega_i \|q - x_i\|^r]$ is minimized. As before, we use the notation $D_n^\star = \min_X D(X, W(X))$ to represent the corresponding minimum possible sensing inaccuracy.

We have used the Lloyd algorithm [25] to find a locally-optimal deployment of sensors. The performance of the locally-optimal deployment provides an estimate for the optimal performance $D_n^\star$. We have compared our numerical estimate on $D_n^\star$ with the analytical results in Corollary 1 and Theorem 4. In general, we consider $T \geq 1$ types of sensors for a total of $n$ sensors, where $p_t n$ of the sensors have weight $\Omega_t$ for $t \in \{1, \ldots, T\}$. For $\mathbf{\Omega} = [\Omega_1 \cdots \Omega_T]$ and $\mathbf{p} = [p_1 \cdots p_T]$, we have considered the following scenarios:

(S1) $f(q) = \mathbf{1}(q \in [0,1])$, $\mathbf{\Omega} = [\frac{1}{4}\ 1]$, $\mathbf{p} = [\frac{1}{2}\ \frac{1}{2}]$.
(S2) $f(q) = \mathbf{1}(q \in [0,2]) \frac{1}{2\sqrt{2q}}$, $\mathbf{\Omega} = [\frac{1}{4}\ 1]$, $\mathbf{p} = [\frac{1}{2}\ \frac{1}{2}]$.
(S3) $f(q) = \mathbf{1}(q \in [0,1])$, $\mathbf{\Omega} = [1\ 2\ 2\ 5]$, $\mathbf{p} = [\frac{1}{8}\ \frac{2}{8}\ \frac{2}{8}\ \frac{3}{8}]$.
(S4) $f(q) = \mathbf{1}(q \in [0,2]) \frac{1}{2\sqrt{2q}}$, $\mathbf{\Omega} = [1\ 2\ 2\ 5]$, $\mathbf{p} = [\frac{1}{8}\ \frac{2}{8}\ \frac{2}{8}\ \frac{3}{8}]$.
(S5) $f(q) = \mathbf{1}(q \in [0,1]^2)$, $\mathbf{\Omega} = [\frac{1}{10}\ \frac{1}{4}\ 1]$, $\mathbf{p} = [\frac{1}{8}\ \frac{1}{8}\ \frac{3}{4}]$.
(S6) $f(q) = \frac{1}{2\pi} e^{-\frac{\|q\|^2}{2}}$, $\mathbf{\Omega} = [\frac{1}{10}\ \frac{1}{4}\ 1]$, $\mathbf{p} = [\frac{1}{8}\ \frac{1}{8}\ \frac{3}{4}]$.
(S7) $f(q) = \mathbf{1}(q \in [0,1]^2)$, $\mathbf{\Omega} = \frac{1}{4}[1\ 2\ \cdots\ 16]$, $\mathbf{p} = [\frac{1}{16} \cdots \frac{1}{16}]$.
(S8) $f(q) = \frac{1}{2\pi} e^{-\frac{\|q\|^2}{2}}$, $\mathbf{\Omega} = \frac{1}{4}[1\ 2\ \cdots\ 16]$, $\mathbf{p} = [\frac{1}{16} \cdots \frac{1}{16}]$.

Scenarios (S1)-(S4) correspond to one-dimensional sensor networks with two different sensor weight distributions for both uniform and non-uniform event densities. The corresponding numerical results are illustrated in Fig. 4. The horizontal axis represents $n$ and the vertical axis represents the best possible sensing inaccuracy $D_n^\star$. "Simulation" refers to the performance of the deployments that are designed using the Lloyd algorithm. For the uniform event distribution scenarios (S1) and (S3), "analysis" refers to the exact minimum distortion formula (13) of Corollary 1. We can observe that for scenarios (S1) and (S3), the performance of the quantizers that is obtained using the Lloyd algorithm is an almost exact match to the analytical results. The very minor difference (which are almost unnoticeable to the naked eye) is a result of the finite number of source samples that are used in the Lloyd algorithm and Monte Carlo simulations.

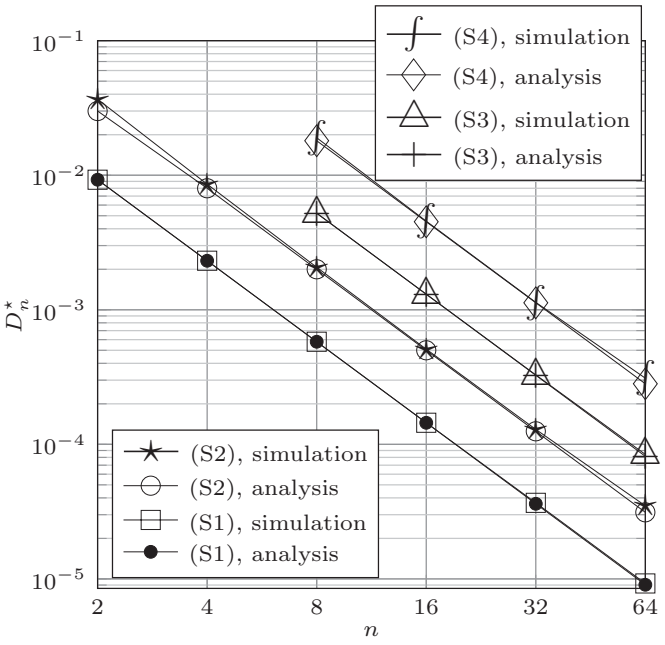


Fig. 4. Optimal sensing performance for one-dimensional sensor networks.

For the non-uniform scenarios (S2) and (S4), "analysis" refers to the formula (21) of Theorem 4. For our current purposes of numerically approximating $D_n^\star$, we thus consider (21) as an equality that holds for every $n$. Since (21) is an asymptotic equality as $n \to \infty$, we expect it to provide a very close approximation to $D_n^\star$ when $n$ is large. The results in Fig. 4 demonstrate that (21) is, in fact, a very good approximation to $D_n^\star$ even for values of $n$ that are as small as 2 or 4.

We present numerical results for the two-dimensional sensor network design scenarios (S5)-(S8) in Fig. 5. Similarly, "simulation" refers to the performance of the deployments that are designed using the Lloyd algorithm. "Analysis" refers to the superposition of the curves defined by (21) for every possible

value of $\ell_{rd}(\mathbf{p}, \mathbf{\Omega})$. In other words, it refers to the set of points

$$\left\{ \left( n, \frac{\ell}{n^{\frac{r}{d}}} \left( \sum_{t \in \mathcal{T}} p_t \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}} \right) : \ell \in [m_{22}, \kappa_{22}] \right\}, \quad (27)$$

which is as a "thickened curve." On the other hand, since $\frac{m_{22}}{\kappa_{22}} \simeq 0.99239\cdots$ is very close to 1, the thickened curves appear as regular non-thickened curves in Fig. 5.

Regarding the results of Fig. 5, by Theorem 4, we expect the analysis to coincide with the simulation as $n \to \infty$. In fact, especially for the uniform density scenarios (S5) and (S7), the analysis is almost an exact match to the simulation even for values of $n$ that are as small as 8 or 16. For scenarios (S6) and (S8) with a Gaussian density of events, the simulation results somewhat deviate from the analysis for small values of $n$. Nevertheless, they still converge to their respective analytical approximations as $n$ grows larger.

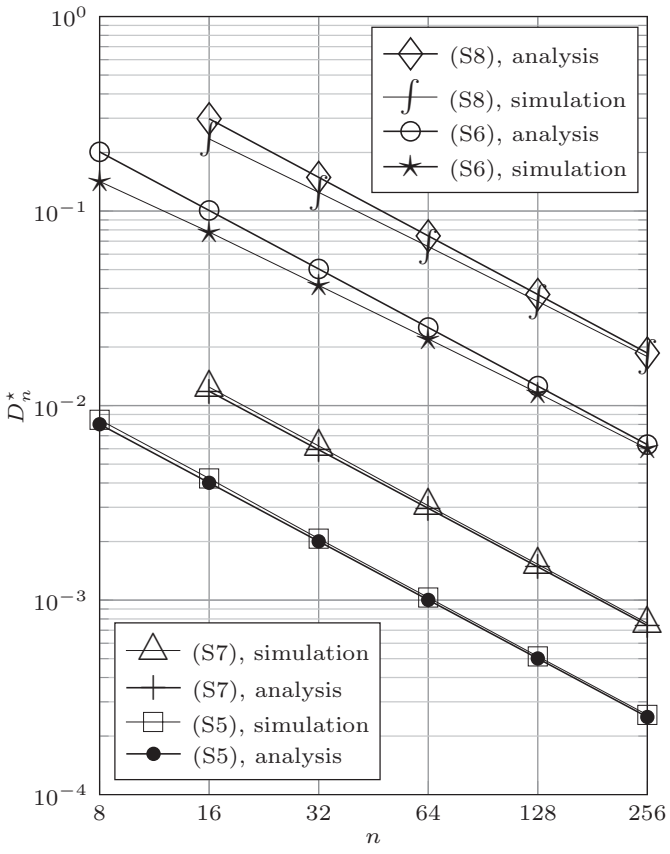


Fig. 5. Optimal sensing performance for two-dimensional sensor networks.

To also gain an intuitive understanding on the structure of optimal deployments, a locally-optimal deployment that is designed using the Lloyd algorithm for scenario (S5) is illustrated in Fig. 6 together with the corresponding optimal sensing regions (Voronoi cells). We can observe that the sensing regions for sensors of the same weight have similar sizes. In fact, as also shown in the proof of Theorem 4, the area of a given sensing region is roughly inversely proportional to the associated sensor's weight. We can also observe that bad sensors with large weights usually surround good sensors with small weights. In general, this allows each sensing region to approach a more disk-like shape, resulting in a lesser overall sensing distortion. Example 1 provides another manifestation of the same phenomenon.

In the examples above, we have focused on the optimal deployment of a fixed collection of sensors with specified weights and frequencies. One may further consider the optimization of the sensor collection given cost constraints. For example, the cost of deploying a given sensor may be modeled as a decreasing function of the sensor weight. In such a formulation, the goal would be to optimize the sensor weights and the number of sensors to be utilized for each weight subject to a constraint on the cost of deployment. Precise formulations and solutions of such design problems will remain as interesting directions for future research.

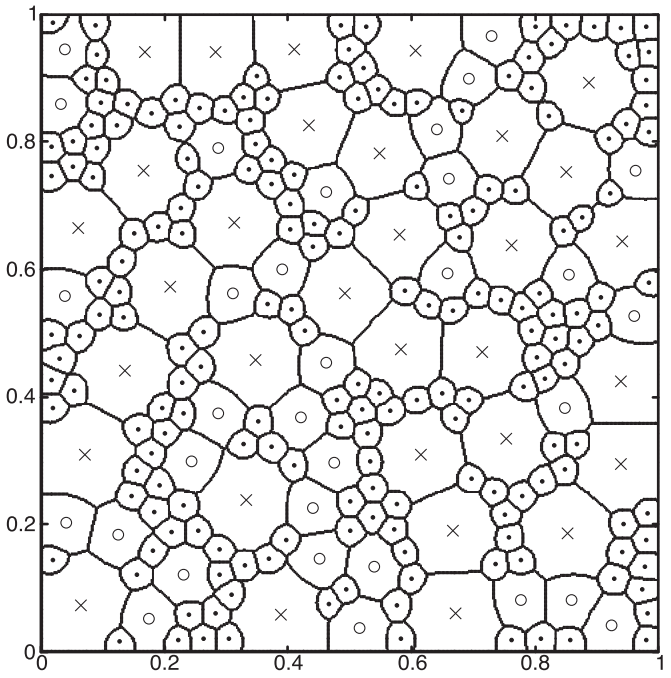


Fig. 6. Voronoi regions of a two-dimensional sensor network deployment that is designed using Lloyd's algorithm. Scenario (S5) is considered for $n = 256$. Sensors of weights 0.1, 0.25 and 1 are represented by crosses, hollow disks, and dots, respectively.

## V. Conclusions

We have studied quantization problems whose distortion measures exhibit dependency on the quantization index. Such problems arise in a variety of applications that involve locational optimization of unidentical agents; examples include sensor networks and cellular networks. We have studied the design and analysis of quantizers where different quantization indices induce squared-error distortions with different multiplicative weights. We have derived the exact minimum average distortion for the special case of scalar quantization over a uniform source, and performed a high-resolution analysis of the minimum average distortion for the general case of vector quantizers over non-uniform source distributions. Numerical simulations have confirmed our analytical findings.

## APPENDIX A
## PROOF OF THEOREM 1

We need the following two lemmas.

**Lemma 1.** *For any $A \subset \mathbb{R}^d$ and $x_0 \in \mathbb{R}^d$, we have*

$$\int_A \|q - x_0\|^r \mathrm{d}q \geq m_{rd} \left(\mu(A)\right)^{\frac{d+r}{d}} \tag{28}$$

*with equality if and only if $A$ is (up to measure zero sets) a ball centered at $x_0$.*

*Proof.* A ball at the origin has the lowest normalized moment among all other sets. In other words, for any $A \subset \mathbb{R}^d$, we have $M_{rd}(A) \geq m_{rd}$, and therefore

$$\begin{aligned}
\int_A \|q - x_0\|^r \mathrm{d}q &= \int_{A - x_0} \|q\|^r \mathrm{d}q && (29)\\
&= M_{rd}(A - x_0) \left(\mu(A - x_0)\right)^{\frac{d+r}{d}} && (30)\\
&\geq m_{rd} \left(\mu(A - x_0)\right)^{\frac{d+r}{d}} && (31)\\
&= m_{rd} \left(\mu(A)\right)^{\frac{d+r}{d}}, && (32)
\end{aligned}$$

with equality if and only if $A - x_0$ is a ball centered at the origin. □

**Lemma 2.** *Let $\gamma > 1$. The following holds:*

*(i) For any positive functions $f, g$, we have*

$$\|fg\|_1 \geq \|f\|_{\frac{1}{\gamma}} \|g\|_{-\frac{1}{\gamma-1}}. \tag{33}$$

*If further $\|fg\|_1 < \infty$, equality occurs if and only if $\exists c > 0$ for which $f = c g^{-\frac{\gamma}{\gamma-1}}$ almost everywhere.*

*(ii) For any two sequences of positive real numbers $a_i$, $i \in \mathbb{N}$ and $b_i$, $i \in \mathbb{N}$, we have*

$$\sum_{i \in \mathbb{N}} a_i b_i \geq \left( \sum_{i \in \mathbb{N}} a_i^{\frac{1}{\gamma}} \right)^{\gamma} \left( \sum_{i \in \mathbb{N}} b_i^{-\frac{1}{\gamma-1}} \right)^{-(\gamma-1)}. \tag{34}$$

*If further $\sum_{i \in \mathbb{N}} a_i b_i < \infty$, equality occurs if and only if $\exists c > 0$ for which $a_i = c b_i^{-\frac{\gamma}{\gamma-1}}$, $\forall i \in \mathbb{N}$.*

*Proof.* The two claims are special cases of the reverse Hölder's inequality. □

We are now ready to prove the theorem. For any quantizer $(X, U)$, we have

$$\begin{aligned}
D(X,U) &= \sum_{i \in \mathcal{I}_n} \omega_i \int_{S_i} \|q - x_i\|^r \mathrm{d}q && (35)\\
&\geq m_{rd} \sum_{i \in \mathcal{I}_n} \omega_i \left(\mu(S_i)\right)^{\frac{d+r}{d}} && (36)\\
&\geq m_{rd} \left( \sum_{i \in \mathcal{I}_n} \omega_i^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \left( \sum_{i \in \mathcal{I}_n} \mu(S_i) \right)^{\frac{d}{d+r}} && (37)\\
&= m_{rd} \left( \sum_{i \in \mathcal{I}_n} \omega_i^{-\frac{d}{r}} \right)^{-\frac{r}{d}}, && (38)
\end{aligned}$$

where the first inequality follows from Lemma 1, and the second inequality follows from Lemma 2-(ii). According to Lemma 1, the first inequality is an equality if and only if $S_i$ is a ball centered at $x_i$, while the second inequality is an equality if and only if there exists $\tau > 0$ such that $\mu(S_i) = \tau \omega_i^{-\frac{d}{r}}$, $\forall i \in \mathcal{I}_n$. Due to the obvious constraint $\sum_{i \in \mathcal{I}_n} \mu(S_i) = 1$, we have $\tau = (\sum_{i \in \mathcal{I}_n} \omega_i^{-\frac{d}{r}})^{-1}$.

## APPENDIX B
## PROOF OF PROPOSITION 1

Our proof relies on the same argument that is used for optimal entropy-constrained quantizers [12], [13]. For any $X$ and $U$, we have

$$\begin{aligned}
D(X,U) &= \sum_{i \in \mathcal{I}_n} \int_{S_i} \omega_i \|q - x_i\|^r f(q) \mathrm{d}q && (39)\\
&\geq \sum_{i \in \mathcal{I}_n} \int_{S_i} \min_{j \in \mathcal{I}_n} \omega_j \|q - x_j\|^r f(q) \mathrm{d}q && (40)\\
&= \int_{\mathbb{R}^d} \min_{j \in \mathcal{I}_n} \omega_j \|q - x_j\|^r f(q) \mathrm{d}q && (41)\\
&= \sum_{i \in \mathcal{I}_n} \int_{V_i(X)} \min_{j \in \mathcal{I}_n} \omega_j \|q - x_j\|^r f(q) \mathrm{d}q && (42)\\
&= \sum_{i \in \mathcal{I}_n} \int_{V_i(X)} \omega_i \|q - x_i\|^r f(q) \mathrm{d}q && (43)\\
&= D(X, W(X)), && (44)
\end{aligned}$$

where (41) follows since the cells $S_i$, $i \in \mathcal{I}_n$ is a partition of $\mathbb{R}^d$, and (42) likewise follows since $V_i(X)$, $i \in \mathcal{I}_n$ is a partition of $\mathbb{R}^d$. Equality (43) is by the definition of $V_i(X)$. This concludes the proof.

## APPENDIX C
## PROOF OF THEOREM 2

We generalize the argument of Pollard [41] for index-independent distortion measures to index-dependent distortion measures; also see [39, Theorem 4.8].

For $\boldsymbol{\omega}' = [\omega'_1 \cdots \omega'_n]$, let

$$\begin{aligned}
D_n(X, \boldsymbol{\omega}') &\triangleq \mathrm{E}\Big[ \min_{i \in \mathcal{I}_n} \omega'_i \|Q - x_i\|^2 \Big], && (45)\\
D_n^\star(\boldsymbol{\omega}') &\triangleq \inf_{X \in \mathbb{R}^{dn}} D_n(X, \boldsymbol{\omega}'). && (46)
\end{aligned}$$

In particular, if we vectorize the weight set as $\boldsymbol{\omega} \triangleq [\omega_1 \cdots \omega_n]$, we have $D_n(X) = D_n(X, \boldsymbol{\omega})$, and $D_n^\star = D_n^\star(\boldsymbol{\omega})$. For $n > 1$, we also let

$$D'_{n-1} \triangleq \min_{\boldsymbol{\omega}'' \in \mathcal{W}_{n-1}} D_{n-1}^\star(\boldsymbol{\omega}''), \tag{47}$$

where

$$\begin{aligned}
\mathcal{W}_{n-1} = \{ &[\omega_2 \cdots \omega_n], [\omega_1\, \omega_3 \cdots \omega_n], \ldots,\\
&[\omega_1\, \omega_2 \cdots \omega_{n-1}] \}. \qquad (48)
\end{aligned}$$

We need the following lemma.

**Lemma 3.** *Suppose $n > 1$. If $D_n^\star < D'_{n-1}$, the set $Z \triangleq \{X \in \mathbb{R}^{dn} : D_n(X) \leq z\}$ is non-empty and compact for every $z \in (D_n^\star, D'_{n-1})$.*

*Proof.* Clearly, $Z$ is non-empty. By the continuity of $X \mapsto \min_{i \in \mathcal{I}_n} \omega_i \|Q - x_i\|^2$ and Lebesgue's dominated convergence

theorem, the function $X \mapsto D_n(X)$ is continuous. By the continuity of $X \mapsto D_n(X)$, the set $Z$ is closed. To show that $Z$ is also bounded, let $\underline{\omega} = \min_{i\in\mathcal{I}_n} \omega_i$ and $\overline{\omega} = \max_{i\in\mathcal{I}_n} \omega_i$. Consider some $X \in Z$. Without loss of generality, suppose $\|x_1\| \le \cdots \le \|x_n\|$. Choose $0 < s < S$ such that

$$P(B(0,s)) > 0, \tag{49}$$

$$\underline{\omega}(S-s)^r P(B(0,s)) > z, \tag{50}$$

$$\overline{\omega} \int_{[B(0,2S)]^c} (2\|x\|)^r f(x)\mathrm{d}x < D'_{n-1} - z. \tag{51}$$

We shall prove that

$$\|x_n\| \le 5S(\overline{\omega}/\underline{\omega})^{\frac{1}{r}}, \tag{52}$$

which implies that $Z$ is bounded and thus concludes the proof. To prove (52), first note that $\|x_1\| \le S$, as otherwise

$$z \ge D_n(X) = \int_{B(0,s)} \min_{i\in\mathcal{I}_n} \omega_i \|q - x_i\|^r f(q)\mathrm{d}q \tag{53}$$

$$\ge \underline{\omega} \int_{B(0,s)} \min_{i\in\mathcal{I}_n} \|q - x_i\|^r f(q)\mathrm{d}q \tag{54}$$

$$\ge \underline{\omega}(S-s)^r P(B(0,s)), \tag{55}$$

a contradiction. Now, suppose (52) does not hold. Let $\|q\| \le 2S$. Since $\|x_1\| \le S$, we have

$$\begin{aligned}
&\omega_1\|q - x_1\|^r \\
&\le \omega_1(\|q\| + \|x_1\|)^r && (56)\\
&\le \omega_1(3S)^r && (57)\\
&\le \omega_1(\|x_n\|(\overline{\omega}/\underline{\omega})^{-\frac{1}{r}} - \|q\|)^r && (58)\\
&\le \omega_1(\|x_n - q\|(\overline{\omega}/\underline{\omega})^{-\frac{1}{r}} + \|q\|(\overline{\omega}/\underline{\omega})^{-\frac{1}{r}} - \|q\|)^r && (59)\\
&\le \omega_1(\underline{\omega}/\overline{\omega})\|x_n - q\|^r && (60)\\
&\le \underline{\omega}\|x_n - q\|^r && (61)\\
&\le \omega_n\|x_n - q\|^r. && (62)
\end{aligned}$$

On the other hand, for $\|q\| > 2S$, we have $\|q\| > 2S \ge 2\|x_1\|$, and thus

$$\begin{aligned}
\omega_1\|q - x_1\|^r &\le \omega_1(\|q\| + \|x_1\|)^r && (63)\\
&\le \omega_1(\|q\| + \tfrac{1}{2}\|q\|)^r && (64)\\
&\le \omega_1(2\|q\|)^r && (65)\\
&\le \overline{\omega}(2\|q\|)^r. && (66)
\end{aligned}$$

Therefore,

$$\begin{aligned}
D'_{n-1} &\le D^\star_{n-1}([\omega_1 \cdots \omega_{n-1}]) && (67)\\
&\le \sum_{i\in\mathcal{I}_n} \int_{\mathcal{V}_i(X)} \min_{i\in\mathcal{I}_n - \{n\}} \omega_i \|q - x_i\|^r f(q)\mathrm{d}q && (68)\\
&\le \sum_{i\in\mathcal{I}_n - \{n\}} \int_{\mathcal{V}_i(X)} \omega_i \|q - x_i\|^r f(q)\mathrm{d}q + \\
&\qquad \int_{\mathcal{V}_n(X)} \omega_1 \|q - x_1\|^r f(q)\mathrm{d}q && (69)\\
&\le \sum_{i\in\mathcal{I}_n - \{n\}} \int_{\mathcal{V}_i(X)} \omega_i \|q - x_i\|^r f(q)\mathrm{d}q + \\
&\qquad \int_{\mathcal{V}_n(X)\cap B(0,2S)} \omega_n \|q - x_n\|^r f(q)\mathrm{d}q + \\
&\qquad \int_{\mathcal{V}_n(X) - B(0,2S)} \overline{\omega}(2\|q\|)^r f(q)\mathrm{d}q && (70)\\
&\le \sum_{i\in\mathcal{I}_n} \int_{\mathcal{V}_i(X)} \omega_i \|q - x_i\|^r f(q)\mathrm{d}q \\
&\qquad + \int_{[B(0,2S)]^c} \overline{\omega}(2\|q\|)^r f(q)\mathrm{d}q && (71)\\
&< D_n(X) + D'_{n-1} - z && (72)\\
&\le D'_{n-1}, && (73)
\end{aligned}$$

a contradiction. □

We are now ready to prove the theorem. The claim (a) for $n = 1$ follows from the already-known existence result [41] of an optimal homogeneous quantizer. We thus prove (b). We have $D^\star_2 < D'_1$, as otherwise, according to the case $n = 1$ of (a), there exists $X \in \mathbb{R}^d$ and $\boldsymbol{\omega} \in \mathcal{W}_1$ such that $D_1(X, \boldsymbol{\omega}) = D^\star_1(\boldsymbol{\omega}) = D^\star_2$, which contradicts the case $n = 2$ of Proposition 2. This proves the case $n = 2$ of (b) as $D'_1 \le D^\star_1$ obviously holds. Moreover, since $D^\star_2 < D'_1$, the set $\{X : D_2(X) \le z\}$ is non-empty and compact for any $z \in (D^\star_2, D'_1)$ by Lemma 3. Therefore,

$$\{X : D_2(X) = D^\star_2\} = \bigcap_{D^\star_2 < z < D'_1} \{X : D_2(X) \le z\} \ne \emptyset, \tag{74}$$

proving the case $n = 2$ of (a). Hence, for any $i \ge 2$, we have $D^\star_i < D'_{i-1} \le D^\star_{i-1}$ by the case $n = i - 1$ of (a) and Proposition 2, which, in turn, proves the case $n = i$ of (b) and (a) via Lemma 3.

## Appendix D
## Proof of Theorem 4

We first provide a short heuristic proof of a slightly weaker claim using the method of point density functions. A formal proof will be presented subsequently.

### A. A Heuristic Proof of a Weaker Claim

Here, we assume that

$$D^\star_n \sim \ell n^{-\frac{r}{d}} \left( \sum_{t\in\mathcal{T}} p_t \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}}, \tag{75}$$

for some constant $\ell$. We will then provide a heuristic proof of the inequality $m_{rd} \le \ell \le \kappa_{rd}$. On the other hand, our argument will not be strong enough to prove that the constant $\ell$ exists and is independent of $f$ unless $d = 1$.

We first prove $\ell \ge m_{rd}$. When $n$ is large, approximately $p_t n$ of the indices have weight $\Omega_t$ for every $t \in \mathcal{T}$. Assume the existence of $|\mathcal{T}|$ point density functions $\lambda_t(q)$, $t \in \mathcal{T}$ such that for any $t \in \mathcal{T}$, the cube centered at $q$ with volume $\mathrm{d}q$ contains $n\lambda_t(q)\mathrm{d}q$ quantization points whose indices have weight $\Omega_t$, and $\int_{\mathbb{R}^d} \lambda_t(q)\mathrm{d}q = p_t$, $\forall t \in \mathcal{T}$. By Theorem 1, the best possible conditional average distortion on the cube at $q$ with volume $\mathrm{d}q$ is lower bounded by

$$m_{rd} n^{-\frac{r}{d}} \left( \sum_{t\in\mathcal{T}} \lambda_t(q) \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}}. \tag{76}$$

Averaging out the source, the minimum average distortion thus satisfies

$$D_n^\star \geq m_{rd} n^{-\frac{r}{d}} \int_{\mathbb{R}^d} \left( \sum_{t\in\mathcal{T}} \lambda_t(q) \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} f(q)\mathrm{d}q. \tag{77}$$

Using Lemma 2-(i), we have

$$D_n^\star \geq m_{rd} n^{-\frac{r}{d}} \left( \int_{\mathbb{R}^d} \sum_{t\in\mathcal{T}} \lambda_t(q) \Omega_t^{-\frac{d}{r}} \mathrm{d}q \right)^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}} \tag{78}$$

$$= m_{rd} n^{-\frac{r}{d}} \left( \sum_{t\in\mathcal{T}} p_t \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}}, \tag{79}$$

which implies $\ell \geq m_{rd}$. To prove $\ell \leq \kappa_{rd}$, dissect $[0,1)^d$ to $|\mathcal{T}|$ disjoint regions $\mathcal{R}_t$, $t \in \mathcal{T}$ (e.g. to rectangles of the form $[a,b)\times[0,1)^{d-1}$) in such a way that

$$\frac{\int_{\mathcal{R}_t} f^{\frac{d}{d+r}}(q)\mathrm{d}q}{\int_{\mathbb{R}^d} f^{\frac{d}{d+r}}(q)\mathrm{d}q} = \frac{p_t \Omega_t^{-\frac{d}{r}}}{\sum_{u\in\mathcal{T}} p_u \Omega_u^{-\frac{d}{r}}}, \forall t \in \mathcal{T}. \tag{80}$$

Given $t \in \mathcal{T}$, we put all $np_t$ reproduction points whose indices have weight $\Omega_t$ optimally in $\mathcal{R}_t$. The average distortion with this strategy is

$$\sum_{t\in\mathcal{T}} \Omega_t \kappa_{rd} (np_t)^{-\frac{r}{d}} \left( \int_{\mathcal{R}_t} f^{\frac{d}{d+r}}(q)\mathrm{d}q \right)^{\frac{d+r}{d}} + |\mathcal{T}| o(n^{-\frac{r}{d}})$$

$$= \kappa_{rd} n^{-\frac{r}{d}} \left( \sum_{t\in\mathcal{T}} p_t \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}} + o(n^{-\frac{r}{d}}), \tag{81}$$

from which the bound $\ell \leq \kappa_{rd}$ follows. This concludes the heuristic proof.

### *B. The Formal Proof of Theorem 4*

Formally, first, note that, according to (19) and (20), we have $p_t > 0$ for some $t \in \mathcal{T}$. Hence, given $n \in \mathbb{N}$, $n \geq 1/\max_{t\in\mathcal{T}} p_t$, we may define $\Delta_n$ to be the minimum average distortion given $\lfloor p_t n \rfloor$ indices with weight $\Omega_t$ for every $t \in \mathcal{T}$. The condition $n \geq 1/\max_{t\in\mathcal{T}} p_t$ ensures that the quantizer has at least one reproduction point so that $\Delta_n$ is well-defined.

We will be working with the quantity $\Delta_n$ instead of $D_n^\star$. For notational convenience, we define

$$L \triangleq \left( \sum_{t\in\mathcal{T}} p_t \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}}. \tag{82}$$

We have the following analogue to Theorem 4.

**Theorem 5.** *We have $\Delta_n \sim \ell L n^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}}$, where $\ell > 0$ satisfies $m_{rd} \leq \ell \leq \kappa_{rd}$ and is independent of $f$.*

The proof of Theorem 5 will be provided later on. First, we note that $\Delta_n$ and $D_n^\star$ have the same asymptotic behavior, as shown by the following lemma. For a sequence $g(n)$, we use the notations $\lim g(n) \triangleq \lim_{n\to\infty} g(n)$, $\liminf g(n) \triangleq \liminf_{n\to\infty} g(n)$, $\limsup g(n) \triangleq \limsup_{n\to\infty} g(n)$.

**Lemma 4.** *We have $\Delta_n \sim D_n^\star$.*

*Proof.* Regarding $D_n^\star$, for every $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we have at least $\lfloor (1-\epsilon)p_t n \rfloor$ points of weight $\Omega_t$ for every $t \in \mathcal{T}$. Thus,

$$\limsup n^{\frac{r}{d}} D_n^\star \leq \limsup n^{\frac{r}{d}} \Delta_{(1-\epsilon)n} \tag{83}$$

$$= (1-\epsilon)^{-\frac{r}{d}} \limsup n^{\frac{r}{d}} \Delta_n, \tag{84}$$

where the equality follows from Theorem 5. Since $\epsilon > 0$ can be chosen arbitrarily, we obtain $\limsup n^{\frac{r}{d}} D_n^\star \leq \limsup n^{\frac{r}{d}} \Delta_n$. On the other hand, since (for large $n$) we have less than $\lfloor (1+\epsilon)p_t n \rfloor$ indices of weight $\Omega_t$ for every $t \in \mathcal{T}$, we can obtain

$$\liminf n^{\frac{r}{d}} D_n^\star \geq \liminf n^{\frac{r}{d}} \Delta_n \tag{85}$$

using a similar argument. Therefore, $\lim n^{\frac{r}{d}} D_n^\star$ exists and equals $\lim n^{\frac{r}{d}} \Delta_n$, implying $D_n^\star \sim \Delta_n$ by Theorem 5. □

Note that Lemma 4 and Theorem 5 together imply Theorem 4. Therefore, we now present a proof of Theorem 5. We begin with the case of a uniform distribution over a cube and then gradually proceed to a general $f$.

Given $a > 0$, let $u_a(q) \triangleq a^{-d}\mathbf{1}(q \in [0,a]^d)$ denote the uniform distribution on $[0,a]^d$. We sometimes write $\Delta_n(f)$ or $\Delta_n(\mathbf{p}, \mathbf{\Omega}, f)$ instead of $\Delta_n$ to signify the dependence of $\Delta_n$ to one or all of $\mathbf{p}$, $\mathbf{\Omega}$, and $f$. We need the following two lemmas.

**Lemma 5.** *For any $a > 0$, we have $\Delta_n(u_a) = a^r \Delta_n(u_1)$.*

*Proof.* Trivial. □

**Lemma 6.** *We have $\liminf n^{\frac{r}{d}} \Delta_n(u_1) \geq m_{rd} L$, and $\limsup n^{\frac{r}{d}} \Delta_n(u_1) \leq \kappa_{rd} L$.*

*Proof.* According to Theorem 1, we have

$$\Delta_n(u_1) \geq m_{rd} \left( \sum_{t\in\mathcal{T}} \lfloor p_t n \rfloor \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}}. \tag{86}$$

Using the inequality $\lfloor p_t n \rfloor \leq p_t n$, we obtain the lower bound on $\liminf n^{\frac{r}{d}} \Delta_n(u_1)$.

For the upper bound on $\limsup n^{\frac{r}{d}} \Delta_n$, dissect $[0,1]^d$ to $T$ Borel measurable regions $\mathcal{R}_t$, $t \in \mathcal{T}$ (e.g. to rectangles of the form $[a,b]\times[0,1]^{d-1}$) in such a way that $t \neq u \implies P(\mathcal{R}_t \cap \mathcal{R}_u) = 0$ and

$$P(\mathcal{R}_t) = \frac{p_t \Omega_t^{-\frac{d}{r}}}{\sum_{u\in\mathcal{T}} p_u \Omega_u^{-\frac{d}{r}}}, \forall t \in \mathcal{T}. \tag{87}$$

Given $t \in \mathcal{T}$, consider putting all the $\lfloor np_t \rfloor$ reproduction points whose indices have weight $\Omega_t$ optimally in $\mathcal{R}_t$. Let $\Delta'_{tn}$ denote the corresponding conditional average distortion given $X \in \mathcal{R}_t$ and

$$\Delta'_n \triangleq \sum_{t\in\mathcal{T}} P(\mathcal{R}_T) \Delta'_{tn} \tag{88}$$

denote the (unconditional) average distortion.

By Theorem 3, and the fact that $\lfloor np_t \rfloor \sim np_t$, we have

$$\Delta'_{tn} \sim \kappa_{rd} \Omega_t (np_t)^{-\frac{r}{d}} (P(\mathcal{R}_t))^{\frac{r}{d}}. \tag{89}$$

Therefore, according to (88),

$$\Delta'_n \sim \sum_{t\in\mathcal{T}} \kappa_{rd} \Omega_t (np_t)^{-\frac{r}{d}} (P(\mathcal{R}_t))^{1+\frac{r}{d}} \tag{90}$$

$$\sim \kappa_{rd} L n^{-\frac{r}{d}}, \tag{91}$$

where the last asymptotic equality follows once we substitute the value of $P(\mathcal{R}_t)$ in (87). The inequality $\limsup n^{\frac{r}{d}} \Delta_n(u_1) \leq \kappa_{rd} L$ follows since $\Delta_n(u_1) \leq \Delta'_n$, $\forall n$. □

We now prove Theorem 5 for the special case of a uniform distribution over a cube.

**Proposition 4.** *Let $f = u_1$. We have $\Delta_n \sim \ell_1 L n^{-\frac{r}{d}}$ for some constant $\ell_1$ that satisfies $m_{rd} \leq \ell_1 \leq \kappa_{rd}$.*

*Proof.* The proof follows from the self-similarity argument of Zador: Let $m, n \in \mathbb{N}$, $m < n$, and $k \triangleq \lfloor (n/m)^{\frac{1}{d}} \rfloor$. Consider a tessellation of $[0,1]^d$ to $k^d$ subcubes of side-length $\frac{1}{k}$. At each subcube, we use the optimal quantizer with $\lfloor p_t m \rfloor$ indices of weight $\Omega_t$. By Lemma 5, the conditional average distortion at each subcube is $k^{-r}\Delta_m$. Therefore, the resulting quantizer of $[0,1]^d$ achieves an average distortion of $k^{-r}\Delta_m$. Moreover, since it has

$$\lfloor p_t m \rfloor k^d = \lfloor p_t m \rfloor \lfloor (n/m)^{\frac{1}{d}} \rfloor^d \leq \lfloor p_t n \rfloor \tag{92}$$

indices with weight $\Omega_t$, we have $\Delta_n \leq k^{-r}\Delta_m$. The definition of $k$ implies the bound $k+1 \geq (n/m)^{\frac{1}{d}}$, using which we can obtain

$$n^{\frac{r}{d}} \Delta_n \leq \left( \frac{k+1}{k} \right)^r m^{\frac{r}{d}} \Delta_m. \tag{93}$$

Keeping $m$ fixed and letting $n \to \infty$, we have $k \to \infty$ so that $\limsup n^{\frac{r}{d}} \Delta_n \leq m^{\frac{r}{d}} \Delta_m$ for every $m \in \mathbb{N}_{>0}$. Therefore, $\lim n^{\frac{r}{d}} \Delta_n$ exists. According to Lemma 6, we have $m_{rd} L \leq \lim n^{\frac{r}{d}} \Delta_n \leq \kappa_{rd} L$. □

From now on, let $\ell_1$ be as defined in Proposition 4. We will show that the constant $\ell$ in Theorem 5 can be chosen to be equal to $\ell_1$ regardless of the density function $f$. We first consider uniform densities over rectangular regions.

**Proposition 5.** *Let $a_1, \ldots, a_d > 0$, $\mathcal{R} = \prod_{i=1}^d [0, a_i]$, and $f(q) = \mathbf{1}(q \in \mathcal{R})/\mu(\mathcal{R})$. We have*

$$\Delta_n \sim \ell_1 L n^{-\frac{r}{d}} (\mu(\mathcal{R}))^{\frac{r}{d}} = \ell_1 L n^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}}. \tag{94}$$

*Proof.* Let $M$ be a large real number. Consider $R = \prod_{i=1}^d \lceil M/a_i \rceil$ disjoint (up to sets of probability zero) translates $\mathcal{R}_1, \ldots, \mathcal{R}_R$ of $\mathcal{R}$ that can cover the cube $[0, M]^d$. At each translate, we use the optimal quantizer with $\lfloor p_t n \rfloor$ indices of weight $\Omega_t$. With this strategy, given the density function $u_M$, the conditional average distortion given $X \in \mathcal{R}_\rho$ is at most $\Delta_n$ for every $\rho \in \{1, \ldots, R\}$ with $\mathcal{R}_\rho \subset M$. Hence, for any $\rho$, the contribution of all the source samples in $\mathcal{R}_\rho$ to the average distortion is at most $P(\mathcal{R})\Delta_n = \mu(\mathcal{R}) M^{-d} \Delta_n(f)$. Since there are $R$ rectangles, the average distortion is at most $R\mu(\mathcal{R}) M^{-d} \Delta_n(f)$. On the other hand, since we have a total of $R\lfloor p_t n \rfloor = \lfloor p_t n R \rfloor$ indices with weight $\Omega_t$, we have

$$R\mu(\mathcal{R}) M^{-d} \Delta_n \geq \Delta_n(R\mathbf{p}, \mathbf{\Omega}, u_M) \tag{95}$$

$$\sim M^r (nR)^{-\frac{r}{d}} \ell_1 L, \tag{96}$$

where the asymptotic equality follows from Lemma 5 and Proposition 4. Thus,

$$\liminf n^{\frac{r}{d}} \Delta_n \geq M^{r+d} R^{-1-\frac{r}{d}} (\mu(\mathcal{R}))^{-1} \ell_1 L. \tag{97}$$

Since $R\mu(\mathcal{R}) \sim M^d$ as $M \to \infty$, we obtain

$$\liminf n^{\frac{r}{d}} \Delta_n \geq \ell_1 L (\mu(\mathcal{R}))^{\frac{r}{d}}. \tag{98}$$

We now consider the "reverse scenario" where $\epsilon > 0$ is small and we have $C = \prod_{i=1}^d \lceil a_i/\epsilon \rceil$ disjoint translates of $[0, \epsilon]^d$ that cover $\mathcal{R}$. At each translate, we use the optimal quantizer with $\lfloor p_t n / C \rfloor$ indices of weight $\Omega_t$. Using the same arguments as above, we obtain

$$C\epsilon^d (\mu(\mathcal{R}))^{-1} \Delta_n \left( \frac{\mathbf{p}}{C}, \mathbf{\Omega}, u_\epsilon \right) \geq \Delta_n, \tag{99}$$

which leads to

$$\limsup n^{\frac{r}{d}} \Delta_n \leq \ell_1 L (\mu(\mathcal{R}))^{\frac{r}{d}}. \tag{100}$$

This concludes the proof. □

Hence, Theorem 5 holds for uniform distributions over rectangles or cubes. We now prove Theorem 5 for the more general case of "piecewise-uniform" density functions. First, we need the following lemma. We write $\ell_1(\mathbf{p}, \mathbf{\Omega})$ to signify the dependence of $\ell_1$ on $\mathbf{p}$ and $\mathbf{\Omega}$.

**Lemma 7.** *Let $M \geq 1$. Suppose $p_{tm} \geq 0$ be given such that $\sum_{m=1}^M p_{tm} = p_t$ and $\forall m \in \{1, \ldots, M\}$, $\exists t \in \mathcal{T}$, $p_{tm} > 0$. Let $\mathbf{p}_m \triangleq [p_{1m} \cdots p_{Tm}]$. We have*

$$\left( \sum_{m=1}^M (\ell_1(\mathbf{p}_m, \mathbf{\Omega}))^{-\frac{d}{r}} \left( \sum_{t \in \mathcal{T}} p_{tm} \Omega_t^{-\frac{d}{r}} \right) \right)^{-\frac{r}{d}} \geq \ell_1 L. \tag{101}$$

*Proof.* Consider the rectangles

$$[0,1]^{d-1} \times \left[ \sum_{i=1}^{m-1} \eta_i, \sum_{i=1}^{m} \eta_i \right], \; m = 1, \ldots, M, \tag{102}$$

where

$$\eta_m \triangleq \frac{a_m}{\sum_{n=1}^M a_n}, \tag{103}$$

and

$$a_m \triangleq (\ell_1(\mathbf{p}_m, \mathbf{\Omega}))^{-\frac{d}{r}} \sum_{t \in \mathcal{T}} p_{tm} \Omega_t^{-\frac{d}{r}}. \tag{104}$$

Note that the rectangles cover $[0,1]^d$. At the $m$th rectangle, we use the optimal quantizer with $\lfloor p_{tm} n \rfloor$ indices of weight $\Omega_t$. We can compare the asymptotic average distortion with this strategy to that of the optimal quantizer on $[0,1]^d$ by using Propositions 4 and 5. This yields

$$\sum_{m=1}^M \eta_m^{\frac{r+d}{d}} a_m^{-\frac{r}{d}} \geq \ell_1 L. \tag{105}$$

Substituting the value of $\eta_m$, we obtain the statement of the lemma. □

Note that the lemma can also be used to estimate $\ell_1(\mathbf{p}, \mathbf{\Omega})$ given that $\ell_1(\mathbf{p}', \mathbf{\Omega})$, $\mathbf{p}' \in \mathcal{P}'$ are known, where $\mathcal{P}'$ is some collection of probabilities. Here, we utilize the lemma to prove the following proposition, which verifies Theorem 5 for piecewise-uniform density functions.

**Proposition 6.** *Let $\mathcal{C}_1, \ldots, \mathcal{C}_M$ be a sequence of closed cubes with side-length $c$ and $i \neq j \implies P(\mathcal{C}_i \cap \mathcal{C}_j) = 0$. Let*

$$f(q) = \sum_{m=1}^{M} s_m c^{-d} \mathbf{1}(q \in \mathcal{C}_m), \tag{106}$$

*where $s_1, \ldots, s_M \geq 0$ and $\sum_{m=1}^{M} s_m = 1$. Then,*

$$\Delta_n \sim \ell_1 L n^{-\frac{r}{d}} \|f\|_{\frac{d}{d+r}}. \tag{107}$$

*Proof.* Without loss of generality, assume $s_1, \ldots, s_M > 0$. Let

$$q_m = \frac{s_m^{\frac{d}{d+r}}}{\sum_{n=1}^{M} s_n^{\frac{d}{d+r}}}, \; m = 1, \ldots, M. \tag{108}$$

We put $\lfloor q_m p_t n \rfloor$ points with weight $\Omega_t$ optimally to $\mathcal{C}_m$. By Proposition 4, the average distortion with this strategy is asymptotically equal to

$$\sum_{m=1}^{M} s_m c^r n^{-\frac{r}{d}} \ell_1(q_m \mathbf{p}, \mathbf{\Omega}) \left( \sum_{t \in \mathcal{T}} q_m p_t \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \tag{109}$$

$$= n^{-\frac{r}{d}} \ell_1 L c^r \sum_{m=1}^{M} s_m q_m^{-\frac{r}{d}} \tag{110}$$

$$= n^{-\frac{r}{d}} \ell_1 L c^r \left( \sum_{m=1}^{M} s_m^{\frac{d}{d+r}} \right)^{\frac{d+r}{d}} \tag{111}$$

$$= n^{-\frac{r}{d}} \ell_1 L \|f\|_{\frac{d}{d+r}}, \tag{112}$$

where the first equality follows since $\ell_1(q\mathbf{p}, \mathbf{\Omega}) = \ell_1, \; \forall q > 0$, and the second equality follows upon substitution of the value of $q_m$. Since we have a total of

$$\sum_{m=1}^{M} \lfloor q_m p_t n \rfloor \leq \lfloor p_t n \rfloor \tag{113}$$

of weight $\Omega_t$, we obtain

$$\limsup n^{\frac{r}{d}} \Delta_n \leq \ell_1 L \|f\|_{\frac{d}{d+r}}. \tag{114}$$

Given the optimal quantizer, let $\mathcal{X}_{n,t}$ denote the set of reproduction points whose indices have weight $\Omega_t$, and $\mathcal{X}_{n,t,m} \triangleq \mathcal{X}_{n,t} \cap \operatorname{int}\mathcal{C}_m$, where $\operatorname{int}\mathcal{C}_m$ is the interior of $\mathcal{C}_m$, i.e. the open cube with the same mid-point and side-length as $\mathcal{C}_m$. Given $\epsilon \in (0, \frac{c}{2})$, let the closed cube $\mathcal{C}_{m,\epsilon}$ have the same midpoint as $\mathcal{C}_m$ but have side-lengths $c - 2\epsilon$. Choose a finite set $\mathcal{K}_{m,\epsilon} \subset \mathcal{C}_{m,\epsilon}$ that satisfies

$$\min_{d \in \mathcal{K}_{m,\epsilon}} \|q - y\| \leq \inf_{y \notin \mathcal{C}_m} \|q - y\|, \; \forall q \in \mathcal{C}_{m,\epsilon} \tag{115}$$

and has cardinality $k_{\epsilon,c} \triangleq |\mathcal{K}_{m,\epsilon}|$ that depends only on $\epsilon$ and $c$. We have

$$\Delta_n = \sum_{m=1}^{M} s_m \int_{\mathcal{C}_m} \min\{\omega_t \|q - y\|^r : t \in \mathcal{T}, \; y \in \mathcal{X}_{n,t}\} \frac{\mathrm{d}q}{c^d} \tag{116}$$

$$\geq \sum_{m=1}^{M} s_m \int_{\mathcal{C}_{m,\epsilon}} \min\{\omega_t \|q - y\|^r : t \in \mathcal{T}, \; y \in \mathcal{X}_{n,t}\} \frac{\mathrm{d}q}{c^d} \tag{117}$$

$$\geq \sum_{m=1}^{M} s_m \int_{\mathcal{C}_{m,\epsilon}} \min\{\omega_t \|q - y\|^r : t \in \mathcal{T}, \; y \in \mathcal{X}_{n,t} \cup \mathcal{K}_{m,\epsilon}\} \frac{\mathrm{d}q}{c^d} \tag{118}$$

$$= \sum_{m=1}^{M} s_m I_{m,n}, \tag{119}$$

where

$$I_{m,n} \triangleq \int_{\mathcal{C}_{m,\epsilon}} \min\{\omega_t \|q - y\|^r : t \in \mathcal{T}, \; y \in \mathcal{X}_{n,t,m} \cup \mathcal{K}_{m,\epsilon}\} \frac{\mathrm{d}q}{c^d}. \tag{120}$$

Note that given any $m \in \{1, \ldots, M\}$, the quantity $I_{m,n}$ is lower bounded by the average distortion of an optimal quantizer for $u_{c-2\epsilon}$ with $|\mathcal{X}_{n,t,m}| + |\mathcal{K}_{m,\epsilon}| = |\mathcal{X}_{n,t,m}| + k_{\epsilon,c}$ indices of weight $\Omega_t$. Hence, we choose a subsequence $i_n$, $n \in \mathbb{N}$ such that

$$\lim i_n^{\frac{r}{d}} \Delta_{i_n}(f) = \liminf n^{\frac{r}{d}} \Delta_n, \tag{121}$$

and

$$\lim i_n^{-1}(|\mathcal{X}_{i_n,t,m}| + k_{\epsilon,c}) \triangleq p_{tm}. \tag{122}$$

Let $\mathbf{p}_m \triangleq [p_{1m} \cdots p_{Tm}]$. Note that $\sum_{m=1}^{M} p_{tm} \leq p_t$. We have

$$\liminf n^{\frac{r}{d}} \Delta_n = \lim i_n^{\frac{r}{d}} \Delta_{i_n}(f) \tag{123}$$

$$\geq \sum_{m=1}^{M} s_m \liminf i_n^{\frac{r}{d}} I_{m,i_n}. \tag{124}$$

On the other hand, by Proposition 4, we obtain

$$\lim i_n^{\frac{r}{d}} I_{m,i_n} = \ell_1(\mathbf{p}_m, \mathbf{\Omega}) \left( \sum_{t \in \mathcal{T}} p_{tm} \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} (c - 2\epsilon)^r \tag{125}$$

provided that $\forall m \in \{1, \ldots, M\}$, $\exists t \in \mathcal{T}$, $p_{tm} > 0$. Assume the contrary. By considering the scenario $p_t \to 0, \; \forall t \in \mathcal{T}$ in Proposition 4, we obtain $\lim i_n^{\frac{r}{d}} I_{m,i_n} = \infty$ for some $m \in \{1, \ldots, M\}$. According to (123), we have $\lim n^{\frac{r}{d}} \Delta_n = \infty$, which contradicts the finiteness of $\limsup n^{\frac{r}{d}} \Delta_n$ that we have already proved. We can therefore substitute (125) to (123) and let $\epsilon \to 0$ to obtain

$$\liminf n^{\frac{r}{d}} \Delta_n \geq \sum_{m=1}^{M} s_m c^r \ell_1(\mathbf{p}_m, \mathbf{\Omega}) \left( \sum_{t \in \mathcal{T}} p_{tm} \Omega_t^{-\frac{d}{r}} \right)^{-\frac{r}{d}} \tag{126}$$

$$\geq c^r \left( \sum_{m=1}^{M} s_m^{\frac{d}{d+r}} \right)^{\frac{d+r}{d}} \times \left( \sum_{m=1}^{M} (\ell_1(\mathbf{p}_m, \mathbf{\Omega}))^{-\frac{d}{r}} \left( \sum_{t \in \mathcal{T}} p_{tm} \Omega_t^{-\frac{d}{r}} \right) \right)^{-\frac{r}{d}} \tag{127}$$

$$\geq c^r \left( \sum_{m=1}^{M} s_m^{\frac{d}{d+r}} \right)^{\frac{d+r}{d}} \ell_1 L \tag{128}$$

$$= \ell_1 L \|f\|_{\frac{d}{d+r}}, \tag{129}$$

where the second inequality follows from the reverse Hölder inequality, and the last inequality follows from Lemma 7. This concludes the proof. □

The rest of the proof of Theorem 5 relies on using a sequence of piecewise-uniform densities with increasing resolution (the average distortion of which could be determined through Proposition 6) for approximating an arbitrary density and its average distortion. Theorem 5 then follows with $\ell = \ell_1$. The details are straightforward extensions of those for index-independent $r$th power distortion measure; see e.g. Steps $3-6$ as in the proof of [39, Theorem 6.2]. Omitting the details, we consider the proof of Theorem 5 complete. As we have mentioned before, Theorem 5, together with Lemma 4, implies Theorem 4. This concludes the proof.

## Appendix E
## Proof of Proposition 3

We consider the following (suboptimal) quantization cells. Consider a source sample $q$ that belongs to the open ball with volume $(1-\delta)v_i$ centered at $x_i$ for some $i \in \{1,\ldots,N\}$. We quantize such a source sample to $x_i$ (with ties broken arbitrarily), and thus the contribution of $q$-like points to the average distortion of the quantizer is

$$\sum_{i=1}^{N} m_{rd}\omega_i((1-\delta)v_i)^{\frac{d+r}{d}} = m_{rd}(1-\delta)^{\frac{d+r}{d}}\left(\sum_{k=1}^{N}\omega_k^{-\frac{d}{r}}\right)^{-\frac{r}{d}} \quad (130)$$

by our assumption in (i). For any other source sample $q'$, by our assumption in (ii), there is an index $j \in \{1,\ldots,N\}$ such that $|q'-x_j| \le Av_j^{\frac{1}{d}}$, and we quantize the source sample $q'$ to $x_j$. The corresponding average distortion is

$$\omega_j\|q'-x_j\|^r \le A^r\omega_j v_j^{\frac{r}{d}} = A^r\left(\sum_{k=1}^{N}\omega_k^{-\frac{d}{r}}\right)^{-\frac{r}{d}} \quad (131)$$

and thus admits an upper bound that is independent of $q'$ and $j$. The contribution of $q'$-like source samples (which have volume $\delta$ by our assumption in (i)) to the average distortion is therefore at most

$$\delta A^r\left(\sum_{k=1}^{N}\omega_k^{-\frac{d}{r}}\right)^{-\frac{r}{d}}. \quad (132)$$

Combining (130) and (132), we obtain the statement of the proposition.

## References


[1] F. Aurenhammer amd H. Edelsbrunner, "An optimal algorithm for constructing the weighted Voronoi diagram in the plane," *Pattern Recognition*, vol. 17, no. 2, pp. 251–257, 1984.

[2] A. György and T. Linder, "On the structure of optimal entropy-constrained scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 416–427, Feb. 2002.

[3] W. A. Pearlman, "Polar quantization of a complex Gaussian random variable," *IEEE Trans. Commun.*, vol. COM-27, no. 6, pp. 892–899, June 1979.

[4] P. Nazari, B.-K. Chun, F. Tzeng, and P. Heydari, "Polar quantizer for wireless receivers: Theory, analysis, and CMOS implementation," *IEEE Trans. Circuits Syst.-I: Regular Papers*, vol. 61, no. 3, pp. 877–887, Mar. 2014.

[5] B. Matschkal and J. B. Huber, "Spherical logarithmic quantization," *IEEE Trans. Audio Speech Lang. Process.*, vol. 18, no. 1, pp. 126–140, Jan. 2010.

[6] B. F. Schaudt, R. L. S. Drysdale, "Multiplicatively weighted crystal growth Voronoi diagrams," *ACM $7^{th}$ Annual Symp. Comput. Geometry*, June 1991.

[7] D. Huff, "The delineation of a national system of planning regions on the basis of urban spheres of influence," *J. Regional Studies Assoc.*, vol. 7, no. 3, pp. 323–329, Sept. 1973.

[8] L. C. Galvãoa, A. G. N. Novaesb, J. E. Souza de Cursic, and J. C. Souzab, "A multiplicatively-weighted Voronoi diagram approach to logistics districting," *Elsevier Computers & Operations Research*, vol. 33, no. 1, pp. 93–114, Jan. 2006.

[9] Q. Du and D. Wang, "Anisotropic centroidal Voronoi tessellations and their applications," *SIAM J. Sci. Comput.*, vol. 26, no. 3, pp. 737–761, Jan. 2005.

[10] M. Schwager, D. Rus, and J. J. Slotine, "Unifying geometric, probabilistic, and potential field approaches to multirobot deployment," *Intl. J. Robot. Res.*, vol. 30, no. 3, pp. 371–383, Mar. 2011.

[11] K. R. Guruprasad and D. Ghose, "Heterogeneous locational optimisation using a generalized Voronoi partition," *Intl. J. Control*, vol. 86, no. 6, pp. 977–993, Apr. 2013.

[12] P. A. Chou, T. Lookabaough, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Trans. Acoustics, Speech and Signal Process.*, vol. 37, no. 1, pp. 31–42, Jan. 1989.

[13] R. M. Gray, T. Linder, and J. Li, "A Lagrangian formulation of Zador's entropy-constrained quantization theorem," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 695–707, Mar. 2002.

[14] A. Gusrialdi, S. Hirche, T. Hatanaka, and M. Fujita, "Voronoi based coverage control with anisotropic sensors," *American Control Conf.*, June 2008.

[15] A. Gusrialdi, S. Hirche, D. Asikin, T. Hatanaka, and M. Fujita, "Voronoi based coverage control with anisotropic sensors," *Intelligent Service Robotics*, vol. 2, no. 4, pp. 195–204, Oct. 2009.

[16] J.-S. Marier, C.-A. Rabbath, and N. Léchevin, "Optimizing the location of sensors subject to health degradation," *IEEE American Control Conference*, June-July 2011.

[17] ———, "Health-aware coverage control with application to a team of small UAVs," *IEEE Trans. Control Sys. Tech.*, vol. 21, no. 5, pp. 1719–1730, Sept. 2013.

[18] F. Sharifi, A. Chamseddine, H. Mahboubi, Y. Zhang, and A. G. Aghdam, "A distributed deployment strategy for a network of cooperative autonomous vehicles," *IEEE Trans. Control Systems Tech.*, vol. 23, no. 2, pp. 737–745. Mar. 2015.

[19] M. Schwager, D. Rus, and J. J. Slotine, "Optimal coverage for multiple hovering robots with downward facing cameras," *IEEE Intl. Conf. Robotics and Automation*, May 2009.

[20] A. Deshpande, S. Poduri, D. Rus, and G. S. Sukhatme, "Distributed coverage control for mobile sensors with location-dependent sensing models," *IEEE Intl. Conf. Robotics and Automation*, May 2009.

[21] A. Kwok and S. Martínez, "Energy-balancing cooperative strategies for sensor deployment," *IEEE Conf. Decision and Control*, Dec. 2007.

[22] J. Guo and H. Jafarkhani, "Sensor deployment with limited communication range in homogeneous and heterogeneous wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6771–6784, Oct. 2016.

[23] H. Jafarkhani, *Space-Time Coding: Theory and Practice*, Cambridge University Press, 2005.

[24] İ. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. Tleecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[25] P. S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[26] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[27] W. R. Bennett, "Spectra of quantized signals," *The Bell System Tech. J.*, vol. 27, no. 3, pp. 446–472, July 1948.

[28] P. L. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 139–148, Mar. 1982.

[29] J. Bucklew and G. Wise, "Multidimensional asymptotic quantization theory with $r$th power distortion measures," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 239–247, Mar. 1982.

[30] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inf. Theory*, vol. 25, no. 4, pp. 373–380, July 1979.

[31] A. Gersho and R. Gray, *Vector quantization and signal compression*. Kluwer Academic Publishers, 1992.

[32] E. Koyuncu and H. Jafarkhani, "On the minimum distortion of quantizers with heterogeneous reproduction points," *IEEE Data Comp. Conf.*, Mar. 2016.

[33] A. Buzo, A. H. Gray, Jr., R. M. Gray, and J. D. Markel, "Speech coding based upon vector quantization," *IEEE Trans. Acoustics, Speech, and Signal Process.*, vol. ASSP–28, no. 5, pp. 562–574, Oct. 1980.

[34] W. R. Gardner and B. D. Rao, "Theoretical analysis of the high-rate vector quantization of LPC parameters," *IEEE Trans. Speech and Audio Process.*, vol. 3, no. 5, pp. 367–381, Sept. 1995.

[35] J. Li, N. Chaddha, and R. M. Gray, "Asymptotic performance of vector quantizers with a perceptual distortion measure," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1082–1091, May 1999.

[36] T. Linder, R. Zamir, and K. Zeger, "High-resolution source coding for non-difference distortion measures: multidimensional companding," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 548–561, Mar. 1999.

[37] T. Linder and R. Zamir, "High-resolution source coding for non-difference distortion measures: The rate-distortion function," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 533–547, Mar. 1999.

[38] M. Borkovec, W. de Paris, and R. Peikert, "The fractal dimension of the Apollonian sphere packing," *Fractals*, vol. 2, no. 4, pp. 521–526, Dec. 1994.

[39] S. Graf and H. Luschgy, *Foundations of Quantization for Probability Distributions*, Springer-Verlag, 2000.

[40] V. Ricordel and C. Labit, "Tree-structured lattice vector quantization," *European Signal Process. Conf.*, Sept. 1996.

[41] D. Pollard, "Quantization and the method of $k$-means," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 199–205, Mar. 1982.