Exercises with solutions (Set E)

15. Consider a symmetric scalar quantizer with 3 intervals,

$$q(x) = \begin{cases} -b & : \quad x < -a \\ 0 & : \quad |x| \le a \\ b & : \quad x > a \end{cases}$$

and a quantizer input with a zero-mean Laplace pdf,

$$f(x) = \frac{1}{2m}\, e^{-\frac{|x|}{m}}$$

(a) Derive the optimal reconstruction value b as a function of the decision threshold a for MSE distortion.

Express the resulting distortion as function of a and the variance $\sigma^2 = 2m^2$.

<u>*Solution:*</u>

Due to the symmetry of the pdf and the quantizer design, the MSE distortion can be written as

$$\begin{aligned} D(a,b) &= 2\int_0^a x^2\, f(x)\,\mathrm{d}x + 2\int_a^\infty (x-b)^2\, f(x)\,\mathrm{d}x \\ &= 2\int_0^\infty x^2\, f(x)\,\mathrm{d}x - 4b\int_a^\infty x\, f(x)\,\mathrm{d}x + 2b^2\int_a^\infty f(x)\,\mathrm{d}x \end{aligned}$$

Setting the first partial derivative with respect to b equal to 0,

$$\frac{\partial}{\partial b} D(a,b) = 0 = -4\int_a^\infty x\, f(x)\,\mathrm{d}x + 4b\int_a^\infty f(x)\,\mathrm{d}x$$

yields the condition for the optimal reconstruction value b,

$$b = \frac{\int_a^\infty x\, f(x)\,\mathrm{d}x}{\int_a^\infty f(x)\,\mathrm{d}x},$$

which is known as the centroid condition for MSE distortion.

Inserting this expression into the equation yields the following formula for the distortion (when the centroid condition is fulfilled),

$$D^*(a) = 2\int_0^\infty x^2\, f(x)\,\mathrm{d}x - 2\,\frac{\left(\int_a^\infty x\, f(x)\,\mathrm{d}x\right)^2}{\int_a^\infty f(x)\,\mathrm{d}x}$$

By defining

$$\begin{aligned}
A &= 2\int_0^\infty x^2\, f(x)\,\mathrm{d}x = \frac{1}{m}\int_0^\infty x^2\, e^{-\frac{x}{m}}\,\mathrm{d}x \\
B(a) &= 2\int_a^\infty x\, f(x)\,\mathrm{d}x = \frac{1}{m}\int_a^\infty x\, e^{-\frac{x}{m}}\,\mathrm{d}x \\
C(a) &= 2\int_a^\infty f(x)\,\mathrm{d}x = \frac{1}{m}\int_a^\infty e^{-\frac{x}{m}}\,\mathrm{d}x
\end{aligned}$$

we can also write

$$b(a) = \frac{B(a)}{C(a)} \qquad \text{and} \qquad D^*(a) = A - \frac{B(a)^2}{C(a)}$$

For the integrals A, $B(a)$ and $C(a)$, we obtain

$$\begin{aligned}
A &= \frac{1}{m}\int_0^\infty x^2\, e^{-\frac{x}{m}}\,\mathrm{d}x = \left[-e^{-\frac{x}{m}}\left(x^2 + 2m\,x + 2m^2\right)\right]_0^\infty = 2\,m^2 \\
B(a) &= \frac{1}{m}\int_a^\infty x\, e^{-\frac{x}{m}}\,\mathrm{d}x = \left[-e^{-\frac{x}{m}}\left(x + m\right)\right]_a^\infty = (a+m)\,e^{-\frac{a}{m}} \\
C(a) &= \frac{1}{m}\int_a^\infty e^{-\frac{x}{m}}\,\mathrm{d}x = \Big[-e^{-\frac{x}{m}}\Big]_a^\infty = e^{-\frac{a}{m}}
\end{aligned}$$

Then, we obtain for the optimal reconstruction value

$$b(a) = \frac{B(a)}{C(a)} = \frac{(a+m)\,e^{-\frac{a}{m}}}{e^{-\frac{a}{m}}} = a + m = a + \frac{1}{2}\sqrt{2\sigma^2}$$

And the distortion for a centroidal quantizer is given by

$$D^*(a) = A - \frac{B(a)^2}{C(a)} = 2m^2 - \frac{(a+m)^2 e^{-\frac{2a}{m}}}{e^{-\frac{a}{m}}} = 2m^2 - (a+m)^2 e^{-\frac{a}{m}}$$

By using the variance $\sigma^2 = 2m^2$, we obtain

$$D^*(a) = \sigma^2 - \left(a + \frac{1}{2}\sqrt{2\sigma^2}\right)^2 e^{-\frac{2a}{\sqrt{2\sigma^2}}}$$

(b) Determine the decision threshold a in a way that a Lloyd quantizer for MSE distortion is obtained.

Determine the distortion and rate for the Lloyd quantizer by assuming fixed-length coding ($R = \log_2 N$) and compare the obtained R-D point with the Shannon lower bound.

<u>*Solution:*</u>

A Lloyd quantizer is the optimal quantizer (minimum distortion) for a given number of reconstruction levels. It fulfills two conditions: the centroid condition considered above and the so-called nearest neighbor condition:

$$u_k = \frac{1}{2}\left(s'_{k-1} + s'_k\right)$$

For our given 3-interval quantizer, we have

$$a = \frac{1}{2}\left(0 + b\right) = \frac{b}{2}$$

Hence, we obtain

$$a = \frac{a+m}{2} \qquad \Longrightarrow \qquad \begin{array}{rcccl} a^* &=& m &=& \frac{1}{2}\sqrt{2\sigma^2} \\ b^* &=& 2m &=& \sqrt{2\sigma^2} \end{array}$$

It should be noted that these parameters are also obtained by setting the derivative of the distortions $D^*(a)$ with respect to a equal to 0.

For the distortion, we obtain

$$\begin{aligned} D^* &= \sigma^2 - \left(a^* + \frac{1}{2}\sqrt{2\sigma^2}\right)^2 e^{-\frac{2a^*}{\sqrt{2\sigma^2}}} \\ &= \sigma^2 - \left(\frac{1}{2}\sqrt{2\sigma^2} + \frac{1}{2}\sqrt{2\sigma^2}\right)^2 e^{-\frac{2}{\sqrt{2\sigma^2}}\frac{\sqrt{2\sigma^2}}{2}} \\ &= \sigma^2 - \frac{2}{e}\,\sigma^2 = \sigma^2\left(\frac{e-2}{e}\right) \approx 0.264241 \cdot \sigma^2 \end{aligned}$$

The nominal rate R for both Lloyd quantizers is

$$R = \log_2 3$$

For MSE distortion, the Shannon lower bound is given by

$$D^{\text{SLB}}(R) = \frac{e}{\pi}\,\sigma^2\,2^{-2R} \qquad \text{or} \qquad R^{\text{SLB}}(D) = \frac{1}{2}\log_2\left(\frac{\sigma^2}{D}\cdot\frac{e}{\pi}\right)$$

Hence, the distortion for the Shannon lower bound D^{SLB} at rate R is

$$\begin{aligned} D^{\text{SLB}}(R) &= \frac{e}{\pi}\,\sigma^2\,2^{-2\log_2 3} = \frac{e}{9\pi}\,\sigma^2 = \left(\frac{e-2}{e}\,\sigma^2\right)\cdot\frac{e^2}{9\pi(e-2)} \\ &= \frac{e^2}{9\pi(e-2)}\,D^* \approx 0.363833 \cdot D^* \end{aligned}$$

which means that the distortion for the Lloyd quantizer is approximately factor 2.75 (or 4.39 dB) larger than the Shannon lower bound. The rate for the Shannon lower bound R^{SLB} at distortion D^* is

$$\begin{aligned} R^{\text{SLB}}(D^*) &= \frac{1}{2}\log_2\left(\frac{e^2}{\pi\,(e-2)}\right) = \log_2 3 - \frac{1}{2}\log_2\left(\frac{9\pi\,(e-2)}{e^2}\right) \\ &\approx R - 0.729326 \end{aligned}$$

which means that the rate for the Lloyd quantizer is approximately 0.729 bits per symbol (or 85.2%) larger than the Shannon lower bound.

(c) Can the derived optimal quantizer for fixed-length coding be improved by adding entropy coding (without changing the decision thresholds and reconstruction levels)?

<u>*Solution:*</u>

For the developed Lloyd quantizer, we obtain reconstruction symbols with the following pmf $\{p_0, p_1, p_2\}$:

$$\begin{aligned} p_1 &= 2\int_0^a f(x)\,\mathrm{d}x = \frac{1}{m}\int_0^m e^{-\frac{x}{m}}\,\mathrm{d}x = \left[-e^{-\frac{x}{m}}\right]_0^m = 1 - \frac{1}{e} \\ p_0 = p_2 &= \frac{1-p_1}{2} = \frac{1}{2e} \end{aligned}$$

Since the probability masses are not the same, the performance of the quantizer can be improved by entropy coding.
The minimum rate is given by

$$\begin{aligned} R^* &= H = -2\cdot\frac{1}{2e}\cdot\log_2\left(\frac{1}{2e}\right) - 1\cdot\left(1-\frac{1}{e}\right)\cdot\log_2\left(1-\frac{1}{e}\right) \\ &= \frac{1}{e}\log_2(2e) - \left(1-\frac{1}{e}\right)\cdot\log_2\left(1-\frac{1}{e}\right) \approx 1.316909 \end{aligned}$$

Hence, by entropy coding, the rate (for same distortion) could be reduced by approximately 0.27 bits per sample or 16.9%.

16. Given is a Centroidal quantizer (not necessarily a Lloyd quantizer) for MSE distortion and a source X. The quantizer has 5 reconstruction levels $\{-3, -1, 0, 1, 3\}$ which are chosen with probabilities $\{0.05, 0.1, 0.4, 0.3, 0.15\}$ and achieves an MSE of 1.05.

   (a) Determine the mean μ and variance σ^2 of the source X.

   Solution:
   We know that the quantizer obeys the centroid condition for MSE. Hence, the reconstruction levels can be written as

$$s'_k = \frac{\int_{u_k}^{u_{k+1}} x\, f(x)\, \mathrm{d}x}{\int_{u_k}^{u_{k+1}} f(x)\, \mathrm{d}x} = \frac{1}{p_k} \int_{u_k}^{u_{k+1}} x\, f(x)\, \mathrm{d}x,$$

   where p_k is the probability that the reconstruction level s'_k is chosen, i.e., the probability that the value of X falls inside the k-th quantization interval $[u_k, u_{k+1})$.
   Then, the mean value μ of X can be written as

$$\mu = \int_{-\infty}^{\infty} x\, f(x)\, \mathrm{d}x = \sum_k \int_{u_k}^{u_{k+1}} x\, f(x)\, \mathrm{d}x = \sum_k p_k\, s'_k$$

   It should be noted that this is the definition of the mean of the quantizer output $q(X)$. Hence, a centroidal quantizer for MSE distortion does not modify the mean of the source X, we have

$$E\{X\} = E\{Q(X)\}$$

   For the mean of the quantization error $e(X) = X - Q(X)$, we obtain

$$E\{e(X)\} = E\{X - Q(X)\} = E\{X\} - E\{Q(X)\} = 0$$

   The quantization error of a centroidal quantizer for MSE distortion has always zero mean.
   For the given quantizer, we obtain

$$\mu = -0.05 \cdot 3 - 0.1 \cdot 1 + 0.3 \cdot 1 + 0.15 \cdot 3 = 0.5$$

   Another given value is the MSE distortion, which can be written as

$$\begin{aligned} D &= \sum_k \int_{u_k}^{u_{k+1}} (x - s'_k)^2\, f(x)\, \mathrm{d}x \\ &= \sum_k \left(\int_{u_k}^{u_{k+1}} x^2\, f(x)\, \mathrm{d}x - 2s'_k \int_{u_k}^{u_{k+1}} x\, f(x)\, \mathrm{d}x + s'^2_k \int_{u_k}^{u_{k+1}} f(x)\, \mathrm{d}x \right) \\ &= \int_{-\infty}^{\infty} x^2\, f(x)\, \mathrm{d}x - \sum_k \left(2s'_k \cdot (s'_k\, p_k) - s'^2_k \cdot p_k \right) \\ &= E\{X^2\} - \sum_k s'^2_k\, p_k \end{aligned}$$

The last term on the right side is the second moment $E\{q(X)^2\}$ of the quantizer output and the distortion is the second moment $E\{e(X)\} = E\{(X - q(X))^2\}$ of the quantization error. Hence, for a centroidal quantizer for MSE distortion, the second moment of the input X is equal to the sum of the second moments of the quantizer output $q(X)$ and the quantization error $e(X) = X - q(X)$,

$$E\{X^2\} = E\{q(X)^2\} + E\{e(X)^2\}$$

Since the mean of X and $q(X)$ is the same and the mean of $e(X)$ is zero, we also have

$$\begin{aligned} E\{X^2\} &= E\{q(X)^2\} + E\{e(X)^2\} \\ \sigma_X^2 + \mu_X^2 &= \sigma_{q(X)}^2 + \mu_{q(X)}^2 + \sigma_{e(X)}^2 \\ \sigma_X^2 &= \sigma_{q(X)}^2 + \sigma_{e(X)}^2 \end{aligned}$$

The variance of the source X is equal to the sum of the variances of the quantizer output and the quantization error.
The variance can then be expressed according to

$$\begin{aligned} \sigma^2 &= E\{(X-\mu)^2\} = E\{X^2\} - 2\mu\, E\{X\} + \mu^2 = E\{X^2\} - \mu^2 \\ &= E\{q(X)^2\} + E\{(X - q(X))^2\} - \mu^2 \\ &= \sum_k s_k'^2\, p_k + D - \mu^2 \end{aligned}$$

For the given quantizer, we obtain

$$\sigma^2 = 0.05 \cdot 9 + 0.1 \cdot 1 + 0.3 \cdot 1 + 0.15 \cdot 9 + 1.05 - 0.5^2 = 3$$

(b) With $q(X)$ being the quantizer output and $e(X) = X - q(X)$ being the quantization error, determine the correlations $E\{X\, q(X)\}$, $E\{X\, e(X)\}$, and $E\{q(X)\, e(X)\}$.

<u>*Solution:*</u>
The correlation $E\{X\, q(X)\}$ can be written as

$$\begin{aligned} E\{X\, q(X)\} &= \int_{-\infty}^{\infty} x\, q(x)\, f_{Xq(X)}(x, q(x))\, \mathrm{d}x \\ &= \sum_k \int_{u_k}^{u_{k+1}} x\, s_k'\, f_X(x)\, g_{q(X)|X}(q(x)|x)\, \mathrm{d}x \end{aligned}$$

where $f_{Xq(X)}(x, q(x))$ denotes the joint pdf of X and $q(X)$, $f_X(x)$ denotes the marginal pdf of X, and $g_{q(X)|X}(y|x)$ denotes the conditional pmf of $q(X)$ given X.
Since $q(X)$ is a deterministic function of X, the conditional pmf $g_{q(X)|X}(y|x)$ is given by

$$g_{q(X)|X}(y|x) = \begin{cases} 1 & : \; y = s_k' \quad \text{with} \quad k:\; x \in [u_k, u_{k+1}) \\ 0 & : \; \text{otherwise} \end{cases}$$

Hence, we obtain

$$\begin{aligned}
E\{X\,q(X)\} &= \sum_k \int_{u_k}^{u_{k+1}} x\, s'_k\, f_X(x)\, g_{q(X)|X}(q(x)|x)\,\mathrm{d}x \\
&= \sum_k s'_k \int_{u_k}^{u_{k+1}} x\, f(x)\,\mathrm{d}x = \sum_k s'^2_k\, p_k \\
&= E\{q(X)^2\}
\end{aligned}$$

For centroidal quantizers for MSE distortion, the correlation between the input and the output signal is equal to the second moment of the quantizer output.

For our example quantizer, we obtain

$$E\{X\,q(X)\} = 0.05 \cdot 9 + 0.1 \cdot 1 + 0.3 \cdot 1 + 0.15 \cdot 9 = 2.2$$

For the correlation between the input signal X and the quantization error, we obtain

$$\begin{aligned}
E\{X\,e(X)\} &= E\{X\,(X - q(X))\} = E\{X^2\} - E\{X\,q(X)\} \\
&= E\{e(X)^2\} + E\{q(X)^2\} - E\{q(X)^2\} \\
&= E\{e(X)^2\} = D
\end{aligned}$$

For centroidal quantizers for MSE distortion, the correlation between the input and quantization error is equal to the second moment of the quantization error, i.e., it is equal to the MSE distortion. Except the quantizer yields a distortion of zero, i.e., it does not apply any quantization, the input signal and the quantization error are always correlated.

For the given quantizer, the correlation $E\{X\,e(X)\}$ is equal to 1.05.

Finally, for the correlation between quantizer output and quantization error, we obtain

$$\begin{aligned}
E\{q(X)\,e(X)\} &= E\{q(X)\,(X - q(X))\} = E\{q(X)\,X\} - E\{q(X)^2\} \\
&= E\{q(X)^2\} - E\{q(X)^2\} = 0
\end{aligned}$$

For centroidal quantizers for MSE distortion, the quantizer output and the quantization error are always uncorrelated.

17. Consider a discrete Markov process $\mathbf{X} = \{X_n\}$ with the symbol alphabet $\mathcal{A}_X = \{0, 2, 4, 6\}$ and the conditional pmf

$$p_{X_n|X_{n-1}}(x_n|x_{n-1}) = \begin{cases} a & : \quad x_n = x_{n-1} \\ \frac{1}{3}(1-a) & : \quad x_n \neq x_{n-1} \end{cases},$$

for $x_n, x_{n-1} \in \mathcal{A}_X$. The parameter a, with $0 < a < 1$, is a variable that specifies the probability that the current symbol is equal to the previous symbol. For $a = 1/4$, our source $\mathbf{X}$ would be i.i.d.

Given is a quantizer of size 2 with the reconstruction levels $s'_0 = 1$ and $s'_1 = 5$ and the decision threshold $u_1 = 3$.

(a) Assume optimal entropy coding using the marginal probabilities of the quantization indices and determine the rate-distortion point of the quantizer.

<u>*Solution:*</u>
First, we determine the marginal pmf $p_X(x)$. By reasons of symmetry, it can easily be seen that the marginal pmf is given by

$$p_X(x) = \frac{1}{4}$$

for all symbols $x \in \mathcal{A}_X$. In a more rigorous way, this can also be derived by

$$\begin{aligned}
p_X(x) &= \sum_{\forall y \in \mathcal{A}_X} p_{X_n|X_{n-1}}(x|y) \cdot p_X(y) \\
&= a \cdot p_X(x) + \frac{1-a}{3} \sum_{\forall y \in \mathcal{A}_X :\, y \neq x} p_X(y) \\
&= a \cdot p_X(x) + \frac{1-a}{3} (1 - p_X(x)) \\
&= \left(\frac{4a}{3} - \frac{1}{3} \right) \cdot p_X(x) + \frac{1}{3} - \frac{a}{3} \\
3\, p_X(x) &= (4a - 1)\, p_X(x) + (1 - a) \\
(4 - 4a)\, p_X(x) &= 1 - a \\
p_X(x) &= \frac{1-a}{4-4a} = \frac{1}{4}
\end{aligned}$$

Let $\mathbf{Y} = \{Y_n\}$ denote the random sequence of quantization indices, with $Y_n \in \{0, 1\}$. With $\mathcal{C}_k$ denoting the quantization cell with the reconstruction value s'_k, the marginal pmf $p_Y(y)$ can be written as

$$p_Y(k) = \sum_{x \in \mathcal{C}_k} p_X(x) = 2 \cdot \frac{1}{4} = \frac{1}{2}$$

For the distortion D, we then obtain

$$D = \sum_{k=0}^{1} \sum_{\forall x \in \mathcal{C}_k} (x - s'_k)^2\, p_X(x) = 4 \cdot 1^2 \cdot \frac{1}{4} = 1$$

And the rate for optimal entropy coding is given by

$$R = -\sum_{k=0}^{1} p_Y(k)\, \log_2 p_Y(k) = -2 \cdot \frac{1}{2} \cdot \log_2\left(\frac{1}{2}\right) = 1$$

An optimal entropy coding can be realized by a simple code that assigns a bit equal to 0 to the quantization index 0 and a bit equal to 1 to the quantization index 1 (or vice versa).

The rate-distortion point for the quantizer with scalar entropy coding is given by $R = 1$ and $D = 1$.

(b) Can the overall quantizer performance be improved by applying conditional entropy coding (e.g., using arithmetic coding with conditional probabilities)? How does it depend on the parameter a?

<u>*Solution:*</u>

The distortion is only dependent on the decision threshold and reconstruction levels. Hence, it does not change by modifying the entropy coding for the quantization indices and is $D = 1$.

For conditional entropy coding, the minimum achievable rate is given by the conditional entropy

$$R = -\sum_{i=0}^{1}\sum_{j=0}^{1} p_{Y_n Y_{n-1}}(i,j)\, \log_2 p_{Y_n|Y_{n-1}}(i|j)$$

For determining the rate, we first have to calculate the joint and conditional pmf for the quantization indices. The joint pmf is given by

$$\begin{aligned} p_{Y_n Y_{n-1}}(i,j) &= \sum_{\forall x \in \mathcal{C}_i} \sum_{\forall y \in \mathcal{C}_j} p_{X_n X_{n-1}}(x,y) \\ &= \sum_{\forall x \in \mathcal{C}_i} \sum_{\forall y \in \mathcal{C}_j} p_{X_n|X_{n-1}}(x|y)\, p_X(y) \end{aligned}$$

For the case $i \neq j$, we obtain

$$p_{Y_n Y_{n-1}}(i,j) = 2 \cdot 2 \cdot \frac{1-a}{3} \cdot \frac{1}{4} = \frac{1-a}{3},$$

and for the case $i = j$, we have

$$p_{Y_n Y_{n-1}}(i,i) = 2 \cdot \left(a + \frac{1-a}{3}\right) \cdot \frac{1}{4} = \frac{1}{6}\,(3a + 1 - a) = \frac{1+2a}{6}$$

Hence, the joint pmf is

$$p_{Y_n Y_{n-1}}(i,j) = \begin{cases} \frac{1}{6}(1+2a) & : \; i = j \\ \frac{1}{3}(1-a) & : \; i \neq j \end{cases}$$

And for the conditional pmf, we obtain

$$p_{Y_n|Y_{n-1}}(i|j) = \frac{p_{Y_n Y_{n-1}}(i,j)}{p_Y(j)} = \begin{cases} \frac{1}{3}(1+2a) & : \; i = j \\ \frac{2}{3}(1-a) & : \; i \neq j \end{cases}$$

Consequently, the minimum rate achievable by conditional entropy coding is

$$
\begin{aligned}
R &= -\sum_{i=0}^{1}\sum_{j=0}^{1} p_{Y_n Y_{n-1}}(i,j)\, \log_2 p_{Y_n|Y_{n-1}}(i|j) \\
&= -2\left(\frac{1+2a}{6}\,\log_2\left(\frac{1+2a}{3}\right) + \frac{1-a}{3}\,\log_2\left(\frac{2-2a}{3}\right)\right) \\
&= -\left(\frac{1+2a}{3}\right)\log_2\left(\frac{1+2a}{3}\right) - \left(\frac{2-2a}{3}\right)\log_2\left(\frac{2-2a}{3}\right) \\
&= -\left(\frac{1+2a}{3}\right)\log_2\left(\frac{1+2a}{3}\right) - \left(1-\frac{1+2a}{3}\right)\log_2\left(1-\frac{1+2a}{3}\right) \\
&= H_b\left(\frac{1+2a}{3}\right),
\end{aligned}
$$

where $H_b(p)$ represents the binary entropy function.

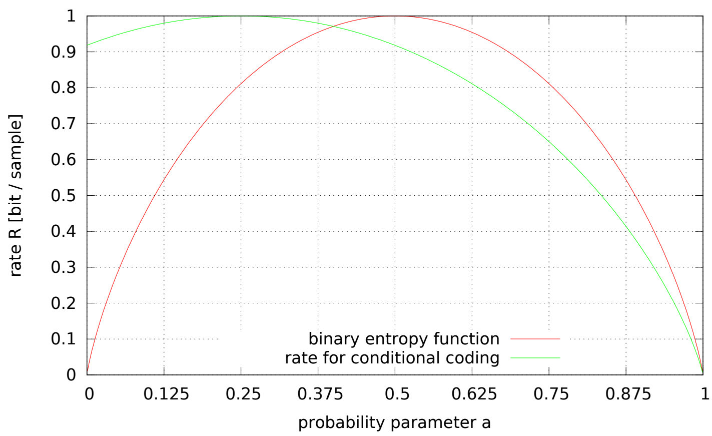
The rate R is maximized (equal to 1) if the argument of the binary entropy function is equal to $1/2$, i.e., if

$$\frac{1}{2} = \frac{1+2a}{3} \qquad \Longrightarrow \qquad a = \frac{1}{4}$$

In this case, the source samples are independent.

The larger the absolute difference $\left|a-\frac{1}{4}\right|$, the more dependent successive source samples are and the lower the rate becomes. If a approaches 1 (i.e., if the difference $\left|a-\frac{1}{4}\right|$ approaches the maximum of $\frac{3}{4}$), the rate approaches 0.

In the following diagram, the rate for optimal entropy coding using the conditional pmf of the quantization indices is plotted over the probability parameter a.



In summary, we note that the performance of scalar quantizers for sources with memory can be improved if we apply entropy coding techniques that employ conditional (or joint) probabilities.

18. Calculate the gain of optimal 2-dimensional vector quantization relative to optimal scalar quantization for high rates on the example of a uniform pdf.

Hint: For high rates, border effects can be neglected. It can be assumed that the signal space for which the pdf is non-zero is completely filled with regular quantization cells.

Solution:

For a uniform pdf the optimal scalar quantizer is a uniform threshold quantizer with reconstruction levels at the center of the quantization intervals. This quantizer represents both, a Lloyd quantizer and an entropy-constrained Lloyd quantizer.

Without loss of generality, we write the uniform pdf according to

$$f(x) = \begin{cases} \frac{1}{2a} & : \; |x| \le a \\ 0 & : \; |x| > a \end{cases},$$

where a determines the with of the distribution.

Let N denote the number of quantization cells. The width of the quantization cells is given by

$$\Delta = \frac{2a}{N}$$

and the rate for the quantizer (all cells have the same probability) is given by

$$R = \log_2 N$$

For the distortion D_k inside one interval, we obtain

$$\begin{aligned} D_k &= \int_{s'_k-\frac{\Delta}{2}}^{s'_k+\frac{\Delta}{2}} (x - s'_k)^2 \, f(x)\,\mathrm{d}x = \frac{1}{2a} \int_{s'_k-\frac{\Delta}{2}}^{s'_k+\frac{\Delta}{2}} (x - s'_k)^2 \,\mathrm{d}x = \frac{1}{2a} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} t^2 \,\mathrm{d}t \\ &= \frac{1}{2a} \left[\frac{t^3}{3} \right]_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} = \frac{1}{6a} \left[\frac{\Delta}{8} + \frac{\Delta}{8} \right] = \frac{1}{24a}\,\Delta^3 = \frac{1}{24a} \left(\frac{2a}{N} \right)^3 \\ &= \frac{a^2}{3\,N^3} \end{aligned}$$

For the overall distortion, we then obtain

$$D = \sum_{k=0}^{N-1} D_k = N \cdot D_k = N \cdot \frac{a^2}{3\,N^3} = \frac{a^2}{3\,N^2}$$

By using $R = \log_2 N$ (and thus $N = 2^R$), finally obtain the operational rate-distortion function for scalar quantization

$$D_1(R) = \frac{a^2}{3} \cdot 2^{-2R} = \sigma^2 \cdot 2^{-2R}$$

We now consider vector quantization with 2 dimensions. The joint pdf in 2 dimensions is given by

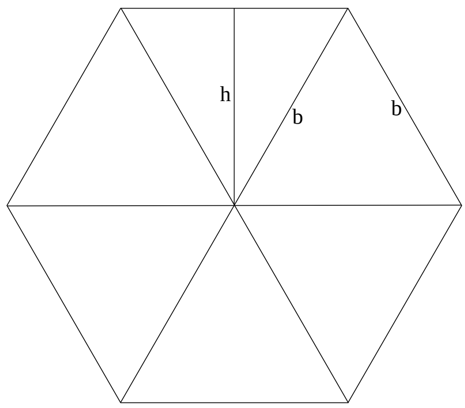
$$f(x,y) = \begin{cases} \frac{1}{4a^2} & : \quad |x| \le a \wedge |y| \le a \\ 0 & : \quad |x| > a \vee |y| > a \end{cases}$$

If we ignore border effects, the optimal quantization cells in two dimensions are regular hexagons, as they provide the densest packing. At the borders of the non-zero range $[-a \cdots a] \times [-a \cdots a]$ of our joint pdf $f(x,y)$ the shapes would be different, but these effects can be ignored if we consider high rates (i.e., a large number of quantization cells). For the uniform pdf, a quantizer with hexagonal cells and reconstruction values inside the center is a Lloyd quantizer as well as an entropy-constrained Lloyd quantizer.

In the high rate case, the number of quantization cells can be approximated by

$$N = \frac{4a^2}{A_{\text{hexagon}}},$$

where A_{hexagon} represents the area of a hexagonal quantization cell.



For determining the area of a hexagonal cell, we can divide it into 6 equilateral triangles as shown in the figure above. Let b denote the length of a side of the hexagon and let h denote the height of the triangles. Then, we have

$$\begin{aligned} A_{\text{hexagon}} &= 6 \cdot A_{\text{triangle}} = 6 \cdot \frac{1}{2} \cdot h \cdot b = 3 \cdot (b \cdot \cos(30^\circ)) \cdot b = 3 \cdot (b \cdot \frac{\sqrt{3}}{2}) \cdot b \\ &= \frac{3\sqrt{3}}{2} b^2 \end{aligned}$$

The number of quantization cells becomes

$$N = \frac{4a^2}{\frac{3\sqrt{3}}{2} b^2} = \frac{8\sqrt{3}}{9} \frac{a^2}{b^2}$$

For calculating the distortion D_k inside a quantization cell, we further divide one of the triangles (the top center one) into two right triangles and calculate the distortion of one of these triangles (the one to the right). The line at the right side of the triangle we consider is given by

$$y = \tan(60°) \cdot x = \sqrt{3}x$$

The distortion for a quantization cell is

$$\begin{aligned}
D_k &= 12 \cdot D_{\text{triangle}} = 12 \int\limits_{x=0}^{b/2} \int\limits_{y=\sqrt{3}x}^{h} r^2 \cdot f(x,y) \,\mathrm{d}y \,\mathrm{d}x \\
&= \frac{12}{4a^2} \int\limits_{x=0}^{b/2} \int\limits_{y=\sqrt{3}x}^{h} (x^2 + y^2) \,\mathrm{d}y \,\mathrm{d}x \\
&= \frac{12}{4a^2} \left(\int\limits_{x=0}^{b/2} \left(\int\limits_{y=\sqrt{3}x}^{h} y^2 \,\mathrm{d}y \right) \mathrm{d}x + \int\limits_{x=0}^{b/2} x^2 \left(\int\limits_{y=\sqrt{3}x}^{h} \mathrm{d}y \right) \mathrm{d}x \right) \\
&= \frac{12}{4a^2} \left(\int\limits_{x=0}^{b/2} \left[\frac{h^3}{3} - \frac{3\sqrt{3}x^3}{3} \right] \mathrm{d}x + \int\limits_{x=0}^{b/2} x^2 \left[h - \sqrt{3}x \right] \mathrm{d}x \right) \\
&= \frac{12}{4a^2} \left(\frac{h^3}{3} \int\limits_{x=0}^{b/2} \mathrm{d}x - \sqrt{3} \int\limits_{x=0}^{b/2} x^3 \,\mathrm{d}x + h \int\limits_{x=0}^{b/2} x^2 \,\mathrm{d}x - \sqrt{3} \int\limits_{x=0}^{b/2} x^3 \,\mathrm{d}x \right) \\
&= \frac{12}{4a^2} \left(\frac{h^3}{3} \cdot \left(\frac{b}{2} \right) - \sqrt{3} \cdot \left(\frac{b^4}{4 \cdot 16} \right) + h \cdot \left(\frac{b^3}{3 \cdot 8} \right) - \sqrt{3} \cdot \left(\frac{b^4}{4 \cdot 16} \right) \right) \\
&= \frac{12}{4a^2} \left(\frac{b\,h^3}{6} - \frac{\sqrt{3}\,b^4}{64} + \frac{b^3\,h}{24} - \frac{\sqrt{3}\,b^4}{64} \right)
\end{aligned}$$

With $h = b\cos(30°) = \frac{\sqrt{3}}{2}\,b$, we obtain

$$\begin{aligned}
D_k &= \frac{12}{4a^2} \left(\frac{3\sqrt{3}\,b^4}{6 \cdot 8} - \frac{\sqrt{3}\,b^4}{64} + \frac{\sqrt{3}\,b^4}{24 \cdot 2} - \frac{\sqrt{3}\,b^4}{64} \right) \\
&= \frac{12\sqrt{3}\,b^4}{4\,a^2} \left(\frac{3}{48} - \frac{1}{64} + \frac{1}{48} - \frac{1}{64} \right) = 3\sqrt{3}\,\frac{b^4}{a^2} \left(\frac{4}{48} - \frac{2}{64} \right) \\
&= 3\sqrt{3}\,\frac{b^4}{a^2} \left(\frac{1}{12} - \frac{1}{32} \right) = 3\sqrt{3}\,\frac{b^4}{a^2} \cdot \frac{8-3}{96} = \frac{5\sqrt{3}}{32}\,\frac{b^4}{a^2}
\end{aligned}$$

Using the previously derived relation

$$N = \frac{8\sqrt{3}}{9}\,\frac{a^2}{b^2} \qquad \Longrightarrow \qquad b^4 = \frac{64 \cdot 3}{81}\,\frac{a^4}{N^2} = \frac{64}{27}\,\frac{a^4}{N^2},$$

we get

$$D_k = \frac{5\sqrt{3}}{32}\,\frac{1}{a^2}\cdot\frac{64}{27}\,\frac{a^4}{N^2} = \frac{10\sqrt{3}}{27}\,\frac{a^2}{N^2}$$

The overall distortion for 2 samples is then

$$D^{(2)} = \sum_{k=0}^{N-1} D_k = N\cdot D_k = \frac{10\sqrt{3}}{27}\,\frac{a^2}{N}$$

And for the distortion D per sample, we obtain

$$D = \frac{D^{(2)}}{2} = \frac{5\sqrt{3}}{27}\,\frac{a^2}{N}$$

The rate for 2 samples is given by

$$R^{(2)} = \log_2 N$$

The rate per sample is then

$$R = \frac{R^{(2)}}{2} = \frac{1}{2}\log_2 N$$

yielding the following expression for the number of samples

$$N = 2^{2R}$$

Inserting this expression into the expression for the distortion per sample yields the operational rate-distortion function for 2-dimensional vector quantization

$$D_2(R) = \frac{5\sqrt{3}\,a^2}{27}\cdot 2^{-2R}$$

For the ratio of the distortions for 2-d vector quantization and scalar quantization at the same rate, we obtain

$$\frac{D_2(R)}{D_1(R)} = \frac{\frac{5\sqrt{3}\,a^2}{27}\cdot 2^{-2R}}{\frac{a^2}{3}\cdot 2^{-2R}} = \frac{5\cdot 3\cdot\sqrt{3}}{27} = \frac{5\sqrt{3}}{9} \approx 0.962250$$

The signal-to-noise ratio for high rates is improved by

$$\Delta\rho = -10\log_{10}\left(\frac{D_2(R)}{D_1(R)}\right) = 10\log_{10}\left(\frac{3\sqrt{3}}{5}\right) \approx 0.167119 \text{ dB}$$

The increase in SNR that is obtained by increasing the quantizer dimension from 1 to 2 at high rates (and using optimal entropy-constrained quantizers) is approximately 0.17 dB. Hence, the difference to the rate-distortion curve has been reduced from 1.533 dB to 1.366 dB.

By further increasing the quantizer dimension, the distance to the rate-distortion function can be further reduced. Asymptotically, the rate-distortion function is achieved if the quantizer dimension approaches infinity.

19. Consider scalar quantization of a Laplacian source at high rates:

$$f(x) = \frac{\lambda}{2} \cdot e^{-\lambda\,|x|} \qquad \text{with} \qquad \sigma_S^2 = \frac{2}{\lambda^2}$$

In a given system, the used quantizer is a Lloyd quantizer with fixed-length entropy coding (the number of quantization intervals represents a power of 2).

How many bits per sample can be saved if the quantizer is replaced by an entropy-constrained quantizer with optimal entropy coding?

Note: The operation points of the quantizers can be accurately described by high rate approximations.

<u>*Solution:*</u>

The high-rate approximation of the distortion-rate function for the Lloyd quantizer with fixed-length codes is given by

$$D_F(R) = \frac{1}{12} \left(\int_{-\infty}^{\infty} \sqrt[3]{f(x)}\, \mathrm{d}x \right)^3 \cdot 2^{-2R}$$

Inserting the Laplace pdf yields

$$\begin{aligned}
D_F(R) &= \frac{1}{12} \left(\int_{-\infty}^{\infty} \sqrt[3]{f(x)}\, \mathrm{d}x \right)^3 \cdot 2^{-2R} \\
&= \frac{1}{12} \left(\int_{-\infty}^{\infty} \sqrt[3]{\frac{\lambda}{2}\, e^{-\lambda|x|}}\, \mathrm{d}x \right)^3 \cdot 2^{-2R} \\
&= \frac{\lambda}{24} \left(\int_{-\infty}^{\infty} e^{-\frac{\lambda}{3}|x|}\, \mathrm{d}x \right)^3 \cdot 2^{-2R} \\
&= \frac{\lambda}{3} \left(\int_{0}^{\infty} e^{-\frac{\lambda}{3}\,x}\, \mathrm{d}x \right)^3 \cdot 2^{-2R} \\
&= \frac{\lambda}{3} \left(-\frac{3}{\lambda} \int_{0}^{-\infty} e^{t}\, \mathrm{d}t \right)^3 \cdot 2^{-2R} \\
&= \frac{9}{\lambda^2} \left(e^0 - e^{-\infty} \right)^3 \cdot 2^{-2R} \\
&= \frac{9}{\lambda^2} \cdot 2^{-2R} \\
&= \frac{9}{2} \cdot \sigma^2 \cdot 2^{-2R}
\end{aligned}$$

The high rate approximation of the distortion-rate function for the entropy-constrained Lloyd quantizer with optimal entropy coding is given by

$$D_V(R) = \frac{1}{12} \cdot 2^{2\,h(X)} \cdot 2^{-2R}$$

where $h(X)$ denotes the differential entropy.

For the differential entropy, we obtain

$$\begin{aligned}
h(X) &= -\int_{-\infty}^{\infty} f(x)\,\log_2 f(x)\,\mathrm{d}x \\
&= -\frac{\lambda}{2}\int_{-\infty}^{\infty} e^{-\lambda\,|x|}\left(\log_2\frac{\lambda}{2} - \frac{\lambda}{\ln 2}|x|\right)\,\mathrm{d}x \\
&= -\lambda\int_{0}^{\infty} e^{-\lambda\,x}\left(\log_2\frac{\lambda}{2} - \frac{\lambda}{\ln 2}x\right)\,\mathrm{d}x \\
&= \frac{\lambda^2}{\ln 2}\int_{0}^{\infty} x\,e^{-\lambda x}\,\mathrm{d}x - \lambda\,\log_2\frac{\lambda}{2}\int_{0}^{\infty} e^{-\lambda x}\,\mathrm{d}x
\end{aligned}$$

By setting $t = -\lambda x$ and using the integration rule $\int uv' = uv - \int u'v$ with $u = x$ and $v' = e^t$, we obtain

$$\begin{aligned}
h(X) &= \frac{1}{\ln 2}\int_{0}^{-\infty} t\,e^t\,\mathrm{d}t + \log_2\frac{\lambda}{2}\int_{0}^{-\infty} e^t\,\mathrm{d}t \\
&= \frac{1}{\ln 2}\left(\left[t\,e^t\right]_0^{-\infty} - \int_{0}^{-\infty} e^t\,\mathrm{d}t\right) + \log_2\frac{\lambda}{2}\int_{0}^{-\infty} e^t\,\mathrm{d}t \\
&= \frac{\lambda}{\ln 2}\Big[0 - 0\Big] + \left(\log_2\frac{\lambda}{2} - \frac{1}{\ln 2}\right)\int_{0}^{-\infty} e^t\,\mathrm{d}t \\
&= \left(\log_2\frac{\lambda}{2} - \log_2 e\right)\left[e^t\right]_0^{-\infty} = \log_2\frac{\lambda}{2e}\Big[0 - 1\Big] \\
&= \log_2\frac{2e}{\lambda} = \frac{1}{2}\log_2\left(\frac{4e^2}{\lambda^2}\right) = \frac{1}{2}\log_2\left(2e^2\,\sigma^2\right)
\end{aligned}$$

Inserting this expression yields the operational distortion-rate function

$$\begin{aligned}
D_V(R) &= \frac{1}{12} \cdot 2^{2\cdot\frac{1}{2}\log_2\left(2\,e^2\,\sigma^2\right)} \cdot 2^{-2R} \\
&= \frac{1}{12} \cdot \left(2\,e^2\,\sigma^2\right) \cdot 2^{-2R} \\
&= \frac{e^2}{6} \cdot \sigma^2 \cdot 2^{-2R}
\end{aligned}$$

Setting the distortions equal to zero yields the following rate difference

$\Delta R = R_F - R_V$:

$$\begin{aligned} D_V(R_V) &= D_F(R_F) \\ \frac{e^2}{6} \cdot \sigma^2 \cdot 2^{-2R_V} &= \frac{9}{2} \cdot \sigma^2 \cdot 2^{-2R_F} \\ 2^{2\Delta R} = 2^{2(R_F - R_V)} &= \frac{27}{e^2} \\ \Delta R &= \frac{1}{2} \log_2 \left(\frac{27}{e^2} \right) \approx 0.9347 \end{aligned}$$

By replacing the Lloyd quantizer with fixed-length coding by an entropy-constrained quantizer with optimal variables-length coding approximately 0.93 bit per sample can be saved at high rates (i.e., at rates for which the high rate approximations are sufficiently accurate).