GUC Berlin
German University in Cairo
الجامعة الألمانية بالقاهرة

German University in Cairo - GUC
Faculty of Information Engineering & Technology - IET
Department of Communication Engineering
Dr.-Ing. Heiko Schwarz

COMM901 – Source Coding and Compression
Winter Semester 2013/2014

Final Exam

Bar Code

Instructions: **Read Carefully Before Proceeding.**

1- Non-programmable calculators are allowed
2- Write your solutions in the space provided
3- The exam consists of **4 questions and 1 bonus question**
4- This exam booklet contains **18 pages** including this page
5- Total time allowed for this exam is **(180) minutes**
6- When you are told that time is up, stop working on the test

Good Luck!

Question	1	2	3	4	Bonus	∑
Possible Marks	25	25	25	25	5	100 + 5
Final Marks						

Question 1: Lossless Coding (25 Marks)

A company produces scanners for black-and-white documents. The scanners scan each page of a document in raster-scan order (from the top-left to the bottom-right corner) and represent it as a 3000×2000 matrix of samples. Each sample has either the value 0 (for "black") or the value 1 (for "white").

With the current system, the samples are directly written to a file, so that each page of a document is represented by 750 000 byte (6 000 000 bit). For a new generation of scanners, the company wants to reduce the memory usage. Therefore, the engineers investigate improved lossless coding algorithms. They analyzed the statistical dependencies for a large number of documents. The analysis yielded the following results:

- On average, 20% of the samples are black;
- If the previous sample S_{n-1} in scanning order is black,
  then in 60% of the cases the current sample S_n is also black.

(a) State the marginal probability density function $p(a) = P(S_n = \text{a})$ and write the probability masses directly into the following table. Assume that the number of analysed documents is so large that the measured statistics represent an accurate estimation of the statistical properties of the source. **[1 Mark]**

Solution:

a	$p(a) = P(S_n = a)$
0 (black)	0.2
1 (white)	0.8

[1]

The probability mass $p(0) = 0.2$ is given. For the probability mass $p(1)$, we have

$$p(1) = 1 - p(0) = 1 - 0.2 = 0.8.$$

(b) Determine the conditional pdfs $p(a|0) = P(S_n = a \mid S_{n-1} = 0)$ and $p(a|1) = P(S_n = a \mid S_{n-1} = 1)$ and write the corresponding probability masses directly into the following table. Again, assume that the number of analysed documents is so large that the measured statistics represent an accurate estimation of the statistical properties of the source. **[3 Marks]**

Solution:

a	$p(a\|0) = P(S_n = a \mid S_{n-1} = 0)$	$p(a\|1) = P(S_n = a \mid S_{n-1} = 1)$	
0 (black)	0.6	0.1	
1 (white)	0.4	0.9	**[3]**

The probability mass $p(0|0) = 0.6$ is given. For the probability mass $p(1|0)$, we have

$$p(1|0) = 1 - p(0|0) = 1 - 0.6 = 0.4.$$

The probability mass $p(0|1)$ can be determined as follows:

$$p(0) = p(0|0) \cdot p(0) + p(0|1) \cdot p(1)$$

$$p(0|1) = p(0) \cdot \frac{1 - p(0|0)}{p(1)} = 0.2 \cdot \frac{1 - 0.6}{0.8} = \frac{0.4}{4} = 0.1.$$

Finally, the probability mass $p(1|1)$ is then given by $p(1|1) = 1 - p(0|1) = 1 - 0.1 = 0.9$.

(c) Assume that the source is stationary and derive the joint pmf $p(a, b) = P(S_n = a, S_{n+1} = b)$. Write the probability masses into the following table. **[2 Marks]**

Solution:

a, b	$p(a, b) = P(S_n = a, S_{n+1} = b)$	
0, 0	$0.2 \cdot 0.6 = 0.12$	
0, 1	$0.2 \cdot 0.4 = 0.08$	
1, 0	$0.8 \cdot 0.1 = 0.08$	
1, 1	$0.8 \cdot 0.9 = 0.72$	**[2]**

The joint probability masses are given by

$$p(a, b) = p(a) \cdot p(b|a).$$

(d) One of the engineers suggested to use a block Huffman code, where a codeword is assigned to a group of two successive samples. Derive such a block Huffman code based on the joint pdf $p(a,b)$, which you derived above. Insert the codewords into the following table. **[3 Marks]**

Solution:

a,b	$p(a,b)$	codewords
0, 0	0.12	01
0, 1	0.08	000
1, 0	0.08	001
1, 1	0.72	1

[3]

(e) How many bytes are required, on average, for representing a page of a scanned document using the developed block Huffman code? **[5 Marks]**

Solution:

The average codeword length per symbol is given by

$$\bar{\ell} = \frac{1}{2} \cdot \sum_{a,b} p(a,b) \cdot \ell(a,b) \quad [\mathbf{1}]$$

$$= \frac{1}{2} \cdot \big(1 \cdot 0.72 + 2 \cdot 0.12 + 3 \cdot (0.08 + 0.08)\big) \quad [\mathbf{1}]$$

$$= \frac{1}{2} \cdot (0.72 + 0.24 + 0.48) = \frac{1}{2} \cdot 1.44 = 0.72 \text{ bit per symbol} \quad [\mathbf{1}]$$

Since a document page consists of 3000×2000 samples, we require

$$\bar{n} = 3000 \cdot 2000 \cdot \bar{\ell} \quad [\mathbf{1}]$$

$$= 3000 \cdot 2000 \cdot 0.72 = 4\,320\,000 \text{ bit}$$

$$= 540\,000 \text{ byte} \quad [\mathbf{1}]$$

for representing a document page on average.

(f) Another engineer suggested to apply arithmetic coding using the marginal symbol probabilities $p(a)$. Estimate the number of bytes that are on average required for representing a document page using marginal arithmetic coding. Assume that the precision of the arithmetic coding is so high that the impact of using fixed-precision arithmetic can be ignored. **[4 Marks]**

Solution:

If we code very long messages (which is the case, since a document page consists of 6 million samples) and ignore the impact of fixed-precision arithmetic, the average codeword length per symbol for marginal arithmetic coding is approximately equal to the marginal entropy of the source. Hence, we have

$$\begin{aligned} \bar{\ell} &\approx H(S_n) \quad [\mathbf{1}] \\ &= -p(0) \cdot \log_2 p(0) - p(1) \cdot \log_2 p(1) \quad [\mathbf{1}] \\ &= -0.2 \cdot \log_2 0.2 - 0.8 \cdot \log_2 0.8 \\ &\approx 0.7219280949 \text{ bit per symbol} \quad [\mathbf{1}] \end{aligned}$$

Since a document page consists of 3000×2000 samples, we require

$$\begin{aligned} \bar{n} &= 3000 \cdot 2000 \cdot \bar{\ell} \\ &\approx 3000 \cdot 2000 \cdot 0.7219280949 \\ &\approx 4\,331\,569 \text{ bit} \\ &= 541\,446 \text{ byte} \quad [\mathbf{1}] \end{aligned}$$

for representing a document page on average.

(g) Disappointed by the results of block Huffman coding and marginal arithmetic coding, the engineers analyzed the efficiency of conditional arithmetic coding, i.e., arithmetic coding in which the conditional symbol probabilities $p(a|b) = P(S_n = a|S_{n-1} = b)$ are used. Estimate the number of bytes that are on average required for representing a document page using such a conditional arithmetic coding. Again, assume that the precision of the arithmetic coding is so high that the impact of using fixed-precision arithmetic can be ignored. **[7 Marks]**

Solution:

If we code very long messages (which is the case, since a document page consists of 6 million samples) and ignore the impact of fixed-precision arithmetic, the average codeword length per symbol for conditional arithmetic coding is approximately equal to the conditional entropy $H(S_n|S_{n-1})$ of the source.

For the conditional entropy $H(S_n|S_{n-1} = 0)$, we have

$$H(S_n|S_{n-1} = 0) = -p(0|0) \cdot \log_2 p(0|0) - p(1|0) \cdot log_2\, p(1|0) \quad [\mathbf{1}]$$

$$= -0.6 \cdot \log_2 0.6 - 0.4 \cdot \log_2 0.4$$

$$\approx 0.97095059 \;\text{ bit per symbol} \quad [\mathbf{1}]$$

Similarly, for the conditional entropy $H(S_n|S_{n-1} = 1)$, we have

$$H(S_n|S_{n-1} = 1) = -p(0|1) \cdot \log_2 p(0|1) - p(1|1) \cdot \log_2 p(1|1)$$

$$= -0.1 \cdot \log_2 0.1 - 0.9 \cdot \log_2 0.9$$

$$\approx 0.46899559 \;\text{ bit per symbol} \quad [\mathbf{1}]$$

Hence, the average codeword length per symbol is given by

$$\begin{aligned} \bar{\ell} &\approx H(S_n|S_{n-1}) \quad [\mathbf{1}] \\ &= p(0) \cdot H(S_n|S_{n-1} = 0) + p(1) \cdot H(S_n|S_{n-1} = 1) \quad [\mathbf{1}] \\ &\approx 0.1 \cdot 0.97095059 + 0.9 \cdot 0.46899559 \\ &= 0.51919109 \quad \text{bit per symbol} \quad [\mathbf{1}] \end{aligned}$$

Since a document page consists of 3000×2000 samples, we require

$$\begin{aligned} \bar{n} &= 3000 \cdot 2000 \cdot \bar{\ell} \approx 3000 \cdot 2000 \cdot 0.51919109 \\ &\approx 3\,115\,147 \quad \text{bit} \\ &= 389\,393 \quad \text{byte} \quad [\mathbf{1}] \end{aligned}$$

for representing a document page on average.

Question 2: Quantization (25 Marks)

Consider an iid process $\{S_n\}$. The random variables S_n have the exponential probability density function

$$f_S(s) = \begin{cases} e^{-s} & : \quad s \geq 0 \\ 0 & : \quad s < 0 \end{cases}$$

We consider a scalar quantizer with two quantization intervals. The decision boundary between the two quantization intervals is denoted by u_1 (with $0 < u_1 < \infty$). The reconstruction level for the first quantization interval is denoted by s'_0 (with $0 < s'_0 < u_1$). The reconstruction level for the second quantization interval is denoted by s'_1 (with $u_1 < s'_1 < \infty$).

(a) Determine the decision threshold u_1 in a way that the probability mass function for the resulting quantization indexes is uniform (i.e., both quantization indexes have the same probability). **[5 Marks]**

Hint: $\int e^{-x}\,\mathrm{d}x = -e^{-x}$

Solution:

For obtaining a uniform pmf for the quantization indexes, we require

$$p_0 = \frac{1}{2} \quad [\mathbf{1}]$$

$$\int_{\infty}^{u_1} f_S(s)\ \mathrm{d}s = \frac{1}{2} \quad [\mathbf{1}]$$

$$\int_{0}^{u_1} e^{-s}\ \mathrm{d}s = e^{0} - e^{-u_1} = 1 - e^{-u_1} = \frac{1}{2} \quad [\mathbf{1}]$$

yielding

$$1 - e^{-u_1} = \frac{1}{2}$$

$$e^{-u_1} = \frac{1}{2} \quad [\mathbf{1}]$$

$$u_1 = -\ln\frac{1}{2}$$

$$u_1 = \ln 2 \approx 0.693147 \quad [\mathbf{1}]$$

The decision threshold u_1 that yields a uniform pmf for the quantization indexes is $u_1 = \ln 2$.

(b) Derive the reconstruction levels s_0' and s_1' that minimize the MSE distortion for the above derived decision threshold u_1. **[9 Marks]**

Hints: $\int e^{-x}\,\mathrm{d}x = -e^{-x}$

$\int x\,e^{-x}\,\mathrm{d}x = -e^{-x}(1+x)$ $\qquad$ $\lim_{x\to\infty} e^{-x}(1+x) = 0$

Solution:

The optimal reconstruction level s_0' is given by

$$s_0' = \frac{\int_{-\infty}^{u_1} s\cdot f_S(s)\;\mathrm{d}s}{\int_{-\infty}^{u_1} f_S(s)\;\mathrm{d}s} \quad [\mathbf{1}]$$

$$= \frac{\int_0^{u_1} s\cdot e^{-s}\;\mathrm{d}s}{\int_0^{u_1} e^{-s}\;\mathrm{d}s} \quad [\mathbf{1}]$$

$$= \frac{e^0(1+0) - e^{-u_1}(1+u_1)}{e^0 - e^{-u_1}} \quad [\mathbf{1}]$$

$$= \frac{1 - e^{-\ln 2}(1+\ln 2)}{1 - e^{-\ln 2}} = \frac{1 - \frac{1}{2}(1+\ln 2)}{1 - \frac{1}{2}} = 2\cdot\left(1 - \frac{1}{2} - \frac{1}{2}\ln 2\right) \quad [1]$$

$$= 1 - \ln 2 \approx 0.306853 \quad [\mathbf{1}]$$

Similarly, the optimal reconstruction level $s_1{}'$ is given by

$$s_1' = \frac{\int_{u_1}^{\infty} s\cdot f_S(s)\;\mathrm{d}s}{\int_{u_1}^{\infty} f_S(s)\;\mathrm{d}s} \quad [\mathbf{1}]$$

$$= \frac{\int_{u_1}^{\infty} s\cdot e^{-s}\;\mathrm{d}s}{\int_{u_1}^{\infty} e^{-s}\;\mathrm{d}s} \quad [\mathbf{1}]$$

$$= \frac{e^{-u_1}(1+u_1) - 0}{e^{-u_1} - 0} \quad [\mathbf{1}]$$

$$= 1 + u_1$$

$$= 1 + \ln 2 \approx 1.693147 \quad [\mathbf{1}]$$

For the decision threshold $u_1 = \ln 2$, the optimal reconstruction levels (for minimizing the MSE distortion) are given by $s_0' = 1 - \ln 2$ and $s_1' = 1 + \ln 2$.

(c) Assume we want to determine the reconstruction levels $s_0{}'$ and s_1' and the decision threshold u_1 in a way that the MSE distortion for the quantizer is minimized (i.e., we want to develop a Lloyd quantizer). What conditions would have to be fulfilled by the reconstruction levels and the decision threshold? **[2 Marks]**

Solution:

The following two conditions have to be fulfilled:

- **Centroid condition**: The reconstruction levels s_k' have to be selected in a way that the distortion inside the corresponding quantization interval is minimized. The optimal reconstruction level s_k' for the k-th quantization interval is given by

$$s_k' = \frac{\int_{u_k}^{u_{k+1}} s \cdot f_S(s)\ \mathrm{d}s}{\int_{u_k}^{u_{k+1}} f_S(s)\,\mathrm{d}s} \quad [1]$$

- **Nearest neighbour condition**: The decision thresholds have to be selected in a way that they lie in the centre between the neighbouring reconstruction levels:

$$u_k = \frac{s_{k-1}' + s_k'}{2} \quad [1]$$

(d) Does the developed quantizer (i.e., the quantizer given by the determined reconstruction levels s_0' and s_1' and the determined decision threshold u_1) represent a Lloyd quantizer for the MSE distortion measure? Explain your statement. **[2 Marks]**

Solution:

The centroid condition is fulfilled, since we used it for deriving the reconstruction levels $s_0{}'$ and $s_1{}'$.

The nearest neighbour conditions is not fulfilled, since we have

$$\frac{s_0' + s_1'}{2} = \frac{1 - \ln 2 + 1 + \ln 2}{2} = 1 \neq u_1 = \ln 2 \quad [1]$$

Hence, the developed quantizer does not represent a Lloyd quantizer. **[1]**

(e) The best scalar quantizers are so-called entropy-constrained quantizers, which can be designed by minimizing a Lagrange function of distortion D and rate R (estimated by the entropy). For high bit rates, the operational distortion rate function of entropy-constrained quantizers is given by

$$D(R) = \frac{1}{12} \cdot 2^{2h(S)} \cdot 2^{-2R}$$

Derive the high-rate distortion rate function for entropy-constrained scalar quantizers for the given iid process with the given exponential pdf. Simplify the resulting expression as much as possible. **[7 Marks]**

Hints:	$\int x\, e^{-x} \, \mathrm{d}x = -e^{-x}(1+x)$	$\lim_{x\to\infty} e^{-x}(1+x) = 0$	$\frac{1}{\ln 2} = \log_2 e$

Solution:

We first calculate the differential entropy, which is given by

$$h(S) = -\int_{-\infty}^{\infty} f(s) \cdot \log_2 f(s) \, \mathrm{d}s \quad [\mathbf{1}]$$

$$= -\int_{0}^{\infty} e^{-s} \cdot \log_2 e^{-s} \, \mathrm{d}s \quad [\mathbf{1}]$$

$$= \frac{1}{\ln 2} \int_{0}^{\infty} s \cdot e^{-s} \, \mathrm{d}s \quad [\mathbf{1}]$$

$$= \frac{1}{\ln 2} \cdot \left(e^{0}(1+0) - 0\right) = \frac{1}{\ln 2} \quad [\mathbf{1}]$$

$$= \log_2 e \quad [\mathbf{1}]$$

Inserting the obtained expression into the formulation of the operational distortion-rate function yields

$$D(R) = \frac{1}{12} \cdot 2^{2h(S)} \cdot 2^{-2R}$$

$$= \frac{1}{12} \cdot 2^{2\log_2 e} \cdot 2^{-2R} \quad [\mathbf{1}]$$

$$= \frac{1}{12} \cdot \left(2^{\log_2 e}\right)^2 \cdot 2^{-2R}$$

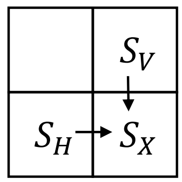
$$= \frac{e^2}{12} \cdot 2^{-2R} \quad [\mathbf{1}]$$

Question 3: Predictive Coding (25 Marks)

We want to predict the samples of a gray-level picture. Therefore, a current sample S_X is predicted by a linear combination of the sample S_H to the left and the sample S_V above the current sample. We consider the class of linear predictors that are described by a single prediction coefficient h and are given by

$$\hat{S}_X = h \cdot (S_H + S_V)$$

The gray-level pictures are realizations of a zero-mean stationary random process with the variance σ_S^2. The correlation coefficient between two horizontally or two vertically adjacent samples is $\varrho_H = \varrho_V = 0.95$. The correlation coefficient between two diagonally adjacent samples is $\varrho_D = 0.92$.



(a) Determine the variance $\sigma_U^2 = E\{U^2\}$ of the prediction error signal $U = S_X - \hat{S}_X$? Note that the mean of the input signal and the mean of the prediction error signal are both equal to zero. Formulate the prediction error variance σ_U^2 as function of the signal variance σ_S^2 and the prediction coefficient h. **[10 Marks]**

Solution:

The prediction error variance is given by

$$\sigma_U^2 = E\{U^2\} = E\left\{\left(S_X - \hat{S}_X\right)^2\right\} \quad [\mathbf{1}]$$

$$= E\{(S_X - h \cdot S_H - h \cdot S_V)^2\} \quad [\mathbf{1}]$$

$$= E\{S_X^2 - hS_XS_H - hS_XS_V - hS_XS_H + h^2S_H^2 + h^2\, S_HS_V - h\, S_XS_V + h^2\, S_HS_V + h^2S_V^2\}$$

$$= E\{S_X^2 - 2h\, S_XS_H - 2h\, S_XS_V + 2h^2\, S_HS_V + h^2S_H^2 + h^2S_V^2\} \quad [\mathbf{1}]$$

$$= E\{S_X^2\} - 2h\,(E\{S_XS_H\} + E\{S_XS_V\}) + h^2\,(2\,E\{S_HS_V\} + E\{S_H^2\} + E\{S_V^2\}) \quad [\mathbf{1}]$$

$$= \sigma_S^2 - 2h\,(\varrho_H \cdot \sigma_S^2 + \varrho_V \cdot \sigma_S^2) + h^2(2\,\varrho_D \cdot \sigma_S^2 + \sigma_S^2 + \sigma_S^2) \quad [\mathbf{4}]$$

$$= \sigma_S^2 \cdot (1 - 2h\,(\varrho_H + \varrho_V) + h^2\,(2\,\varrho_D + 2))$$

$$= \sigma_S^2 \cdot (1 - 2h \cdot (0.95 + 0.95) + h^2 \cdot (2 \cdot 0.92 + 2)) \quad [\mathbf{1}]$$

$$= \sigma_S^2 \cdot (1 - 3.8 \cdot h + 3.84 \cdot h^2) \quad [\mathbf{1}]$$

(b) Consider the simple predictor given by $h = 0.5$, where the current sample is predicted by the average of the sample to the left and the sample above the current sample. Calculate the variance $\sigma_U^2 = E\{U^2\}$ of the prediction error signal $U = S_X - \hat{S}_X$ for this simple predictor? Represent the prediction error variance σ_U^2 as function of the signal variance σ_S^2. **[2 Marks]**

Solution:

For $h = 0.5$, the prediction error variance is given by

$$\begin{aligned}\sigma_U^2 &= \sigma_S^2 \cdot (1 - 3.8 \cdot h + 3.84 \cdot h^2) \\ &= \sigma_S^2 \cdot \left(1 - \frac{3.8}{2} + \frac{3.84}{4}\right) \quad [\mathbf{1}] \\ &= \sigma_S^2 \cdot (1 - 1.9 + 0.96) \\ &= 0.06 \cdot \sigma_S^2 \quad [\mathbf{1}]\end{aligned}$$

(c) The quality of a predictor is typically expressed using the prediction gain. The prediction gain in dB is defined by

$$G_P = 10 \cdot \log_{10} \frac{\sigma_S^2}{\sigma_U^2}$$

Calculate the prediction gain (in dB) for the simple predictor given by $h = 0.5$. State the result with a precision of 4 digits after the decimal point. **[2 Marks]**

Solution:

The prediction gain for the predictor with $h = 0.5$ is given by

$$\begin{aligned}G_P &= 10 \cdot \log_{10} \frac{\sigma_S^2}{\sigma_U^2} = 10 \cdot \log_{10} \frac{\sigma_S^2}{0.06 \cdot \sigma_S^2} = -10 \cdot \log_{10} 0.06 \quad [\mathbf{1}] \\ &= 12.2185 \text{ dB} \quad [\mathbf{1}]\end{aligned}$$

(d) Derive the optimal predictor $\hat{S}_X = h \cdot (S_H + S_V)$. Determine the prediction coefficient h that minimizes the prediction error variance $\sigma_U^2 = E\left\{\left(S_x - \hat{S}_X\right)^2\right\}$. State the result with a precision of at least 8 digits after the decimal point. **[5 Marks]**

Solution:

For minimizing the prediction error variance, we can set the first derivative of the prediction error variance with respect to the prediction coefficient h equal to zero, yielding

$$\frac{\partial}{\partial h}\sigma_U^2 = 0 \quad [\mathbf{1}]$$

$$\frac{\partial}{\partial h}\left(\sigma_S^2 \cdot (1 - 3.8 \cdot h + 3.84 \cdot h^2)\right) = 0 \quad [\mathbf{1}]$$

$$\sigma_S^2 \cdot (0 - 3.8 + 3.84 \cdot 2 \cdot h) = 0 \quad [\mathbf{1}]$$

$$-3.8 + 7.68 \cdot h = 0 \quad [\mathbf{1}]$$

$$h = \frac{3.8}{7.68} = \frac{95}{192}$$

$$h \approx 0.49479167 \quad [\mathbf{1}]$$

The optimal prediction coefficient is $h \approx 0.49479167$.

(e) Calculate the prediction error variance $\sigma_U^2 = E\{U^2\}$ for the derived optimal predictor. State the result as function of the signal variance σ_S^2 and use a precision of 8 digits after the decimal point for the parameters the resulting function. **[2 Marks]**

Solution:

For the optimal predictor, the prediction error variance is given by

$$\sigma_U^2 = \sigma_S^2 \cdot (1 - 3.8 \cdot h + 3.84 \cdot h^2)$$

$$= \sigma_S^2 \cdot \left(1 - 3.8 \cdot \frac{95}{192} + 3.84 \cdot \frac{95^2}{192^2}\right) \quad [\mathbf{1}]$$

$$= 0.05989583 \cdot \sigma_S^2 \quad [\mathbf{1}]$$

(f) By what amount (in dB) is the prediction gain (for the optimal predictor) increased relative to the prediction gain of the simple predictor with $h = 0.5$? State the result in dB, with a precision of 4 digits after the decimal point.

Why are in image coding application often simple predictors (average of two neighboring samples is used as prediction for a current sample) are used, instead of optimal predictors. **[4 Marks]**

Solution:

The difference of the prediction gains is given by

$$\Delta G_P = 10 \cdot \log_{10} \frac{\sigma_S^2}{\sigma_{U,opt}^2} - 10 \cdot \log_{10} \frac{\sigma_S^2}{\sigma_{U,h=0.5}^2} \quad [\mathbf{1}]$$

$$= 10 \cdot \log_{10} \frac{\sigma_{U,h=0.5}^2}{\sigma_{U,opt}^2}$$

$$= 10 \cdot \log_{10} \frac{0.06 \cdot \sigma_S^2}{0.05989583 \cdot \sigma_S^2} \quad [\mathbf{1}]$$

$$= 10 \cdot \log_{10} \frac{0.06}{0.05989583}$$

$$= 0.0075 \ \text{dB} \quad [\mathbf{1}]$$

The optimal predictor increases the prediction gain relative to the simple predictor by 0.0075 dB.

The reasons why typically simple predictors are used in image coding application include:

- The simple predictor is easier to implement;
- The simple predictor is not signal dependent;
- For large correlation coefficients, as typically found in natural images, the prediction gains of the simple and optimal predictor are nearly the same. **[1]**

Question 4: Transform Coding (25 Marks)

We consider transform coding of a zero-mean stationary Gauss-Markov process with the variance σ_S^2 and the correlation coefficient ϱ, with $0 < \varrho < 1$. The input signal is partitioned into vectors $\boldsymbol{s_k} = (s_{2k}, s_{2k+1})^{\mathrm{T}}$, which consist of two successive signal samples s_{2k} and s_{2k+1}. The transform is given by

$$\boldsymbol{u_k} = \begin{bmatrix} u_{k,0} \\ u_{k,1} \end{bmatrix} = \frac{1}{2} \cdot \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} \cdot \begin{bmatrix} s_{2k} \\ s_{2k+1} \end{bmatrix} = \boldsymbol{A} \cdot \boldsymbol{s_k}$$

Each vector $\boldsymbol{s_k}$ of two input samples is mapped to a vector $\boldsymbol{u_k}$ of two transform coefficients. The transform is orthogonal (the basis vectors of $\boldsymbol{A}$ are orthogonal to each other and have a unit norm). As distortion measure, we use the mean squared error (MSE). We consider high-rate coding, for which the operational distortion-rate function of the used entropy-constrained scalar quantizers (for Gaussian random variables) is given by

$$D_i(R_i) = \frac{\pi e}{6} \cdot \sigma_i^2 \cdot 2^{-2R}$$

where σ_i^2 represents the variance of the corresponding transform coefficients (with i being equal to 0 or 1). For reconstructing the signal samples, the reconstructed transform coefficients are transformed using the inverse transform matrix $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}$.

(a) Let $U_{k,0}$ and $U_{k,1}$ represent the random variables for the transform coefficients $u_{k,0}$ and $u_{k,1}$, respectively. Determine the covariance $\sigma_{01}^2 = E\{U_{k,0}U_{k,1}\}$ between the transform coefficients $U_{k,0}$ and $U_{k,1}$. State the result as function of the signal variance σ_S^2 and the correlation coefficient ϱ. **[7 Marks]**

Solution:

For the covariance, we have

$$\begin{aligned}
\sigma_{01}^2 &= E\{U_{k,0}U_{k,1}\} \\
&= E\left\{\left(\frac{1}{2} \cdot S_{2k} + \frac{\sqrt{3}}{2} \cdot S_{2k+1}\right)\left(\frac{\sqrt{3}}{2} \cdot S_{2k} - \frac{1}{2} \cdot S_{2k+1}\right)\right\} \quad \textbf{[2]} \\
&= E\left\{\frac{\sqrt{3}}{4} \cdot S_{2k}^2 + \left(\frac{3}{4} - \frac{1}{4}\right) \cdot S_{2k} \cdot S_{2k+1} - \frac{\sqrt{3}}{4} \cdot S_{2k+1}^2\right\} \\
&= \frac{\sqrt{3}}{4} \cdot E\{S_{2k}^2\} + \frac{1}{2} \cdot E\{S_{2k}S_{2k+1}\} - \frac{\sqrt{3}}{4} \cdot E\{S_{2k+1}^2\} \quad \textbf{[1]} \\
&= \frac{\sqrt{3}}{4} \cdot \sigma_S^2 + \frac{1}{2} \cdot \varrho \cdot \sigma_S^2 - \frac{\sqrt{3}}{4} \cdot \sigma_S^2 \quad \textbf{[3]} \\
&= \sigma_S^2 \cdot \frac{\varrho}{2} \quad \textbf{[1]}
\end{aligned}$$

(b) Does the transform represent a Karhunen Loéve transform (KLT) for the considered source? Explain your statement. **[2 Marks]**

Solution:

The transform does not represent a KLT, since the transform coefficients are correlated. **[2]**

This has been shown in the previous sub-question: $E\{U_{k,0}U_{k,1}\} = \sigma_S^2 \cdot \frac{\varrho}{2} > 0 \quad \text{for} \quad \varrho > 0$

(c) Determine the transform coefficient variances $\sigma_0^2 = E\{U_{k,0}^2\}$ and $\sigma_1^2 = E\{U_{k,1}^2\}$, where $U_{k,0}$ and $U_{k,1}$ represent the random variables for the transform coefficients $u_{k,0}$ and $u_{k,1}$, respectively. State the variances σ_0^2 and σ_1^2 as function of the signal variance σ_S^2 and the correlation coefficient ϱ. **[10 Marks]**

Solution:

For the first transform coefficient, we have

$$\begin{aligned}
\sigma_0^2 = E\{U_{k,0}^2\} &= E\left\{\left(\frac{1}{2}\cdot S_{2k} + \frac{\sqrt{3}}{2}\cdot S_{2k+1}\right)^2\right\} \quad [1] \\
&= E\left\{\frac{1}{4}\cdot S_{2k}^2 + 2\cdot\frac{\sqrt{3}}{4}\cdot S_{2k}\cdot S_{2k+1} + \frac{3}{4}\cdot S_{2k+1}^2\right\} \\
&= \frac{1}{4}\cdot E\{S_{2k}^2\} + \frac{\sqrt{3}}{2}\cdot E\{S_{2k}S_{2k+1}\} + \frac{3}{4}\cdot E\{S_{2k+1}^2\} \quad [1] \\
&= \frac{1}{4}\cdot\sigma_S^2 + \frac{\sqrt{3}}{2}\cdot\varrho\cdot\sigma_S^2 + \frac{3}{4}\cdot\sigma_S^2 \quad [2] \\
&= \sigma_S^2\cdot\left(1 + \frac{\sqrt{3}}{2}\cdot\varrho\right) \quad [1]
\end{aligned}$$

Similarly, for the second transform coefficient, we obtain

$$\begin{aligned}
\sigma_1^2 = E\{U_{k,1}^2\} &= E\left\{\left(\frac{\sqrt{3}}{2}\cdot S_{2k} - \frac{1}{2}\cdot S_{2k+1}\right)^2\right\} \quad [1] \\
&= E\left\{\frac{3}{4}\cdot S_{2k}^2 - 2\cdot\frac{\sqrt{3}}{4}\cdot S_{2k}\cdot S_{2k+1} + \frac{1}{4}\cdot S_{2k+1}^2\right\} \\
&= \frac{3}{4}\cdot E\{S_{2k}^2\} - \frac{\sqrt{3}}{2}\cdot E\{S_{2k}S_{2k+1}\} + \frac{1}{4}\cdot E\{S_{2k+1}^2\} \quad [1] \\
&= \frac{3}{4}\cdot\sigma_S^2 - \frac{\sqrt{3}}{2}\cdot\varrho\cdot\sigma_S^2 + \frac{1}{4}\cdot\sigma_S^2 \quad [2] \\
&= \sigma_S^2\cdot\left(1 - \frac{\sqrt{3}}{2}\cdot\varrho\right) \quad [1]
\end{aligned}$$

(d) Calculate the geometric mean $\tilde{\sigma}^2$ of the transform coefficient variances σ_0^2 and σ_1^2. Formulate the geometric mean $\tilde{\sigma}^2$ as function of the signal variance σ_S^2 and the correlation coefficient ϱ. **[2 Marks]**

Solution:

The geometric mean $\tilde{\sigma}^2$ of the transform coefficient variances σ_0^2 and σ_1^2 is given by

$$\tilde{\sigma}^2 = \sqrt{\sigma_0^2 \cdot \sigma_1^2} \quad [\mathbf{1}]$$

$$= \sigma_S^2 \cdot \sqrt{\left(1+\frac{\sqrt{3}}{2}\right)\left(1-\frac{\sqrt{3}}{2}\right)} = \sigma_S^2 \cdot \sqrt{1-\frac{3}{4}\varrho^2} \quad [\mathbf{1}]$$

(e) How large is the transform coding gain (in dB) for high rates and optimal bit allocation? Formulate the transform coding gain as function of the correlation coefficient ϱ.

Calculate the transform coding gain (in dB) for a correlation coefficient $\varrho = \frac{\sqrt{3}}{2}$. **[4 Marks]**

Solution:

The high-rate transform coding gain for Gaussian random variables with optimum bit allocation is given by

$$G_T = 10 \cdot \log_{10} \frac{D_{SQ}(R)}{D_{TC}(R)} = 10 \cdot \log_{10} \frac{\frac{\pi e}{6} \cdot \sigma_S^2 \cdot 2^{-2R}}{\frac{\pi e}{6} \cdot \tilde{\sigma}_S \cdot 2^{-2R}} = 10 \cdot \log_{10} \frac{\sigma_S^2}{\tilde{\sigma}^2} \quad [\mathbf{2}]$$

By inserting the derived formula for the geometric mean of the transform coefficient variances, we obtain

$$G_T = 10 \cdot \log_{10} \frac{1}{\sqrt{1-\frac{3}{4}\varrho^2}} = -5 \cdot \log_{10}\left(1-\frac{3}{4}\varrho^2\right) = -5 \cdot \log_{10} \frac{4-3\varrho^2}{4} \quad [\mathbf{1}]$$

For the special case $\varrho = \frac{\sqrt{3}}{4}$, we have

$$G_T = -5 \cdot \log_{10} \frac{4-3\cdot\frac{3}{4}}{4}$$

$$= -5 \cdot \log_{10} \frac{16-9}{16} = -5 \cdot \log_{10} \frac{7}{16}$$

$$= 1.7951 \ \ \text{dB} \quad [\mathbf{1}]$$

<u>Bonus Question:</u> Miscellaneous Topics (5 Bonus Marks)

Determine for the following statements whether they are correct or wrong. An explanation is not required. Simply write “correct” or “wrong” below the statements. **[1 Mark for each statement]**

(a) Let S be a discrete random variable that is defined on an alphabet of M values. The entropy $H(S)$ always obeys the condition

$$H(S) \leq \log_2 M$$

<u>Solution:</u>

The statement is correct. **[1]**

(b) At high rates, the Shannon lower bound for the MSE distortion measure is always 1.53 dB worse than the corresponding information distortion-rate function at the same rate.

<u>Solution:</u>

The statement is wrong. At high rates the Shannon lower bound approaches the information distortion rate function. **[1]**

(c) Vector quantization improves the performance relative to scalar quantization only for sources with memory. For iid sources, the best scalar quantizer has the same efficiency as the best vector quantizer.

<u>Solution:</u>

The statement is wrong. Vector quantization can always improve the performance (space filling gain). **[1]**

(d) For AR(m) processes (autoregressive processes of order m), the linear predictor that minimizes the prediction error variances has exactly m non-zero prediction coefficients.

<u>Solution:</u>

The statement is correct. **[1]**

(e) Consider transform coding with orthogonal transforms. At high rates, the optimal bit allocation among the transform coefficients is obtained if all components have the same distortion.

<u>Solution:</u>

The statement is correct. **[1]**