ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 14 Advanced Digital Communications
Homework 6 (*Graded, Due Nov. 7, 2016*) Oct. 31, 2016

PROBLEM 1. (15 pts) Recall that if two random variables (X_1, X_2) satisfy the bivariate normal distribution, they have the probability density function (pdf) in the following form:

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

where we used the vector notations

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \end{bmatrix}, \qquad \text{and} \qquad \Sigma = \begin{bmatrix} \sigma_{X_1}^2 & \rho\sigma_{X_1}\sigma_{X_2} \\ \rho\sigma_{X_1}\sigma_{X_2} & \sigma_{X_2}^2 \end{bmatrix}.$$

If we form a new (complex) random variable Z as $Z := X_1 + jX_2$, we can find the pdf of Z form the pdf of (X_1, X_2). In this case it is very easy to see that $f_Z(z) = f_{X_1,X_2}(\Re(z), \Im(z))$. Under certain conditions (one of them being $\sigma_{X_1}^2 = \sigma_{X_2}^2 = N_0/2$), the pdf of Z takes a particularly nice form:

$$f_Z(z) = \frac{1}{\pi N_0} e^{-\frac{1}{N_0}|z|^2}.$$

This subclass of Gaussian random variables is called *circularly symmetric complex Gaussian random variable* and has desirable properties, as we will explore in the following questions.

(a) Find the other conditions on $\boldsymbol{\mu}$ and ρ such that Z is a circularly symmetric complex Gaussian random variable.

(b) Show that if Z is a circularly symmetric complex Gaussian random variable then $V = e^{j\theta}Z$, where $\theta \in \mathbb{R}$ is also a circularly symmetric complex Gaussian random variable.

(c) Find the pdf of $W = aZ$ where a is a complex constant and Z is a circularly symmetric complex Gaussian random variable. Express the pdf as compactly as possible.

PROBLEM 2. (12 pts) Consider the following discrete time channel:

$$Y[k] = X[k] + X[k-1] + Z[k]$$

where $Z[k]$ is i.i.d. Gaussian with variance σ^2 and is independent of $X[k]$. Moreover suppose that $X[k]$ are independently and uniformly chosen from $\{A, -A\}$.

(a) Find the optimal estimator of $X[k]$ given only the current channel output symbol, $\hat{X}(Y[k])$, and compute its error probability.
Hint. Consider the cases $X[k-1] = X[k]$ and $X[k-1] \neq X[k]$ separately.

(b) Find the optimal estimator $\hat{X}^{\text{DFE}}(Y[k], \hat{X}^{\text{DFE}}[k-1])$ that uses the previously estimated symbol to detect the last sent symbol, compute its error probability and compare it to the previous estimator.
Hint. Consider the cases $\hat{X}^{\text{DFE}}[k-1] = X[k-1]$ and $\hat{X}^{\text{DFE}}[k-1] \neq X[k-1]$ separately.

Problem 3. (12 pts) Consider the general L-tap ISI channel:

$$Y[n] = \sum_{l=0}^{L-1} h_l X[n-l] + W[n],$$

where $X[n]$ is a complex-valued signal and $W[n]$ is circularly-symmetric complex white Gaussian noise with mean zero and variance N_0. We want to use the OFDM approach discussed in class to communicate across this channel.

(a) In the class, we analyzed the implementation of OFDM where we transmit N information symbols with $N+1$ channel uses (which is also the length of one OFDM symbol) for $L=2$. Now assume $L=5$ in the ISI channel and we want to design a system such that each OFDM symbol contains $N=5$ information symbols. What is the length of an OFDM symbol in this design? Give the matrix presentation of the system, by analogy to Equation (4.42) in the lecture notes. Repeat the exercise and answer both questions again for $N=7$.

(b) In general, the equivalent channel matrix of the OFDM system is a circulant matrix which looks like

$$H = \begin{bmatrix} h_0 & h_{N-1} & \ddots & h_2 & h_1 \\ h_1 & h_0 & h_{N-1} & \ddots & h_2 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ h_{N-2} & \ddots & h_1 & h_0 & h_{N-1} \\ h_{N-1} & h_{N-2} & \ddots & h_1 & h_0 \end{bmatrix}$$

where for all $n > L$, $h_n = 0$. The key of OFDM is to *diagonalize* H. Now we will verify this step by showing that the circulant matrices are diagonalized by the Fourier matrix F_N. For $m \in \{0, 1, \ldots, N-1\}$ define

$$\boldsymbol{u}_m = \begin{bmatrix} 1 & \alpha_m^{-1} & \alpha_m^{-2} & \cdots & \alpha_m^{-(N-1)} \end{bmatrix}^T, \qquad \text{where } \alpha_m = e^{-\jmath 2\pi m/N}.$$

Note that $\boldsymbol{u}_m$ is a column of F_N^{-1}. Show that $\boldsymbol{u}_m$ is an eigenvector of H with the corresponding eigenvalue $\lambda_m = \sum_{l=0}^{N-1} h_l \alpha_m^l$. Conclude that $F_N H F_N^{-1} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$.

Problem 4. (18 pts) One of the main drawbacks of the OFDM approach is its "Peak-to-Average Power Ratio" (PAPR in the literature). In this problem, you explore this issue. As we saw in class, in an OFDM system, we consider a block of N information symbols,

$$\boldsymbol{X} = (X_0, X_1, X_2, \ldots, X_{N-1})^T.$$

Assume that each X_i can only assume the values $\{-\sqrt{\mathcal{E}}, \sqrt{\mathcal{E}}\}$ with equal probability, and that all X_i are independent of each other. We take an (inverse) FFT of this block in order to obtain the signal values that will be transmitted over the channel, denoted by $\boldsymbol{x} = (x_0, x_1, x_2, \ldots, x_{N-1})^T$.

(a) The average power is defined as

$$P_{average} = \mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{N-1}|x_k|^2\right] = \frac{1}{N}\mathbb{E}[\|\boldsymbol{x}\|^2],$$

where the expectation is over all possible constellation points of the original signal $X_0, \ldots, X_{N-1}$. Calculate $P_{average}$.

(b) Let us define the peak power as

$$P_{peak} = \max_{X_0, X_1, \ldots, X_{N-1}} \max\{|x_0|^2, |x_1|^2, |x_2|^2, \ldots, |x_{N-1}|^2\},$$

where the outer maximum is over all possible constellation points of the original signal $X_0, \ldots, X_{N-1}$. Calculate P_{peak}.

(c) Using your answers from parts (a) and (b) compute the PAPR, namely $P_{peak}/P_{average}$. Explain what issues can arise from such a PAPR when the OFDM system is to be implemented in practice.

PROBLEM 5. (25 pts) In the class, we study resource allocation for parallel channels. In this problem, we explore a slightly more general problem where the total system capacity is given by

$$C = \sum_{i=1}^{K} \alpha_i \log_2(1 + \beta_i P_i)$$

for some fixed positive numbers α_i and β_i, and we need to maximize this over all possible power allocations that satisfy $P_1 + P_2 + \cdots + P_K = P$. Clearly, we must have $P_i \geq 0$.

(a) Write out the general optimization problem (for the case of general K) in standard form. An optimization problem is in standard form if it is expressed as[1]

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{x}) \\ \text{subject to} \quad & f_i(\boldsymbol{x}) \leq 0, \ \text{for } i = 1, 2, \ldots, m, \\ & h_j(\boldsymbol{x}) = 0, \ \text{for } j = 1, 2, \ldots, p. \end{aligned}$$

Here, $\boldsymbol{x}$ is an n-dimensional vector, $f_0(\boldsymbol{x})$ is called the *objective* function, $f_i(\boldsymbol{x})$ are called the *inequality constraints* and $h_j(\boldsymbol{x})$ are called the *equality constraints*.

(b) The *Lagrangian* for a problem in standard form is given as

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x}) + \sum_{j=1}^{p} \nu_j h_j(\boldsymbol{x}).$$

Write out the Lagrangian for the rate allocation problem.

(c) When the problem is convex (as it is in our case), the key advantage of the Lagrange formulation is that we have the KKT (Karush–Kuhn–Tucker) conditions, saying that the optimal solution must satisfy[2]

$$\begin{aligned} \frac{\partial}{\partial x_\ell} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) &= 0, \ \text{for all } \ell = 1, 2, \ldots, n \\ f_i(\boldsymbol{x}) &\leq 0, \ \text{for all } i = 1, 2, \ldots, m \\ h_j(\boldsymbol{x}) &= 0, \ \text{for all } j = 1, 2, \ldots, p \\ \lambda_i &\geq 0, \ \text{for all } i = 1, 2, \ldots, m \\ \lambda_i f_i(\boldsymbol{x}) &= 0, \ \text{for all } i = 1, 2, \ldots, m \end{aligned}$$

Write out these conditions for the rate allocation problem.

[1]Here, we follow the terminology of the standard reference book: S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004, p.127.

[2]If you want to learn more, you may turn to Chapter 5 of Boyd's book.

(d) Show that your answer to (d) implies that under the optimal power allocation,

$$\begin{aligned} \text{Either } P_i \text{ satisfies:} \quad & \frac{\alpha_i \beta_i}{1+\beta_i P_i} = \text{constant} \\ \text{or} \quad & P_i = 0. \end{aligned}$$

(e) Now assume that $\alpha_i = 1$ and $\beta_i = \frac{|H_i|^2}{N_0}$, which is exactly the rate allocation problem discussed in the class. Show that the condition then becomes

$$\begin{aligned} \text{Either } P_i \text{ satisfies:} \quad & P_i + \frac{N_0}{|H_i|^2} = \text{constant} \\ \text{or} \quad & P_i = 0. \end{aligned}$$

(f) Let us denote $N_i = \frac{N_0}{|H_i|^2}$, which can be interpreted as the "equivalent" noise at channel i. Assume that $K = 5$ and $(N_1, N_2, N_3, N_4, N_5) = (8, 4, 6, 2, 10)$. Apply the "water-filling" solution of (e) to determine the optimal powers P_i for the following cases: the total power $P = 1, 4, 9, 16$. In addition, provide a water-filling figure for each case.

PROBLEM 6. (18 pts) Consider the scalar discrete-time inter symbol interference channel,

$$y[k] = \sum_{n=0}^{\nu} h_n x[k-n] + z[k], \quad k = 0, \ldots, N-1, \tag{1}$$

where $z[k]$ is circularly-symmetric complex white Gaussian noise independent of $x[k]$. Let us employ a cyclic prefix as done in OFDM, i.e.,

$$x[-l] = x[N-l], \quad l = 1, \ldots, \nu.$$

As done in class given the cyclic prefix,

$$\underbrace{\begin{bmatrix} y[N-1] \\ \vdots \\ y[0] \end{bmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{bmatrix} h_0 & \ldots & \ldots & h_\nu & 0 & \ldots & 0 & 0 \\ 0 & h_0 & \ldots & h_{\nu-1} & h_\nu & 0 & \ldots & 0 \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & \ldots & \ldots & 0 & h_0 & \ldots & \ldots & h_\nu \\ h_\nu & 0 & \ldots & 0 & 0 & h_0 & \ldots & h_{\nu-1} \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ h_1 & \ldots & h_\nu & 0 & \ldots & 0 & 0 & h_0 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} x[N-1] \\ \vdots \\ x[0] \end{bmatrix}}_{\boldsymbol{x}} + \underbrace{\begin{bmatrix} z[N-1] \\ \vdots \\ z[0] \end{bmatrix}}_{\boldsymbol{z}}. \tag{2}$$

In the derivation of OFDM we used the property that $F_N^{-1} D F_N = H$ where F_N is the DFT matrix and D is the diagonal matrix with

$$D_{l,l} = H_l = \frac{1}{\sqrt{N}} \sum_{n=0}^{\nu} h_n e^{-j\frac{2\pi}{N}nl}.$$

This yields the parallel channel model

$$Y[l] = H_l X[l] + Z[l]. \tag{3}$$

if $X[l]$ and $Y[l]$ are the Discrete Fourier Transforms of $\boldsymbol{x}$ and $\boldsymbol{y}$ respectively.

If the carrier synchronization is not accurate, then (1) gets modified as

$$y[k] = \sum_{n=0}^{\nu} e^{j2\pi f_0 k} h_n x[k-n] + z[k], \quad k = 0, \dots, N-1$$

where f_0 is the carrier frequency offset. If we still use the cyclic prefix for transmission, then (2) gets modified as

$$\boldsymbol{y} = \underbrace{\begin{bmatrix} h_0 e^{j2\pi f_0 (N-1)} & \dots & h_\nu e^{j2\pi f_0 (N-1)} & 0 & \dots & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & & \\ 0 & \ddots & \ddots & \ddots & e^{j2\pi f_0 \nu} h_0 & \dots & e^{j2\pi f_0 \nu} h_\nu \\ \ddots & \ddots & \ddots & \ddots & \ddots & & \\ e^{j2\pi f_0 0} h_1 & \dots & e^{j2\pi f_0 0} h_\nu & 0 & \dots & 0 & e^{j2\pi f_0 0} h_0 \end{bmatrix}}_{H'} \boldsymbol{x} + \boldsymbol{z}$$

Note that $H' = SH$, where S is a diagonal matrix with $S_{l,l} = e^{j2\pi f_0 (N-l)}$ and H is defined as in (2).

(a) Show that if $\boldsymbol{Y} = F_N \boldsymbol{y}$ and $\boldsymbol{X} = F_N \boldsymbol{x}$,

$$\boldsymbol{Y} = G\boldsymbol{X} + \boldsymbol{Z}$$

and prove that

$$G = F_N S F_N^{-1} D.$$

(b) If $f_0 \neq 0$, we see from part (a) that G is no longer a diagonal matrix and therefore we do not obtain the parallel channel result of (3). We get inter-carrier interference (ICI), *i.e.,* we have

$$Y[l] = G_{l,l} X[l] + \underbrace{\sum_{q \neq l} G_{l,q} X[q] + Z[l]}_{\text{ICI + noise}}, \quad l = 0, \dots, N-1,$$

which shows that the other carriers interfere with $X[l]$. Compute the SINR (signal-to-interference plus noise ratio). Assume $X[l]$ are i.i.d, with $\mathbb{E}[|X[l]|^2] = \mathcal{E}$. You can compute the SINR for the particular l and leave the expression in terms of $G_{l,q}$.

(c) Suppose we want to estimate $\boldsymbol{X}$ from $\boldsymbol{Y}$ using a linear filter $W \in \mathbb{C}^{N \times N}$, as $\hat{\boldsymbol{X}} = W^H \boldsymbol{Y}$. Find the filter W that minimizes the mean square error

$$\mathbb{E}[|W^H \boldsymbol{Y} - \boldsymbol{X}|^2].$$

You can again assume that $X[l]$ are i.i.d. with $\mathbb{E}[|X[l]|^2] = \mathcal{E}$ and that the receiver knows G.