ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Principles of Digital Communications: Assignment date: March 14, 2012
Summer Semester 2012 Due date: March 21, 2012

Solution of Homework 5

Problem 1. *(QAM with Erasure)*

$$
\begin{aligned}
P_{00} =& Pr(\{N_1 \geq -a\} \cap \{N_2 \geq -a\}) \\
=& Pr(\{N_1 \leq a\}) Pr(\{N_2 \leq a\}) \\
=& \left[1 - Q\left(\frac{a}{\sigma}\right)\right]^2 .
\end{aligned}
$$

By symmetry:

$$
\begin{aligned}
P_{01} = P_{03} =& Pr(\{N_1 \leq -(2b-a)\} \cap \{N_2 \geq -a\}) \\
=& Pr(\{N_1 \geq 2b-a\}) Pr(\{N_2 \leq a\}) \\
=& Q\left(\frac{2b-a}{\sigma}\right) \left[1 - Q\left(\frac{a}{\sigma}\right)\right] .
\end{aligned}
$$

$$
\begin{aligned}
P_{04} =& Pr(\{N_1 \leq -(2b-a)\} \cap \{N_2 \leq -(2b-a)\}) \\
=& Pr(\{N_1 \geq 2b-a\} \cap \{N_2 \geq 2b-a\}) \\
=& \left[Q\left(\frac{2b-a}{\sigma}\right)\right]^2 .
\end{aligned}
$$

$$
\begin{aligned}
P_{0\delta} =& 1 - Pr(\{Y \in \mathcal{R}_0\} \cup \{Y \in \mathcal{R}_1\} \cup \{Y \in \mathcal{R}_2\} \cup \{Y \in \mathcal{R}_3\} | \mathbf{s_0} \text{ was sent}) \\
=& 1 - P_{00} - P_{01} - P_{02} - P_{03} \\
=& 1 - \left[1 - Q\left(\frac{a}{\sigma}\right)\right]^2 - 2Q\left(\frac{2b-a}{\sigma}\right)\left[1 - Q\left(\frac{a}{\sigma}\right)\right] - \left[Q\left(\frac{2b-a}{\sigma}\right)\right]^2 \\
=& 1 - \left[1 - Q\left(\frac{a}{\sigma}\right) + Q\left(\frac{2b-a}{\sigma}\right)\right]^2
\end{aligned}
$$

Equivalently,

$$
\begin{aligned}
P_{0\delta} =& Pr(\{N_1 \in [a, 2b-a]\} \cup \{N_2 \in [a, 2b-a]\}) \\
=& Pr(N_1 \in [a, 2b-a]) + Pr(N_1 \in [a, 2b-a]) - Pr(\{N_1 \in [a, 2b-a]\} \cap \{N_2 \in [a, 2b-a]\}) \\
=& 2\left[Q\left(\frac{2b-a}{\sigma}\right) - Q\left(\frac{a}{\sigma}\right)\right] - \left[Q\left(\frac{2b-a}{\sigma}\right) - Q\left(\frac{a}{\sigma}\right)\right]^2,
\end{aligned}
$$

which gives the same result as before.

Problem 2. *(Gaussian Hypothesis testing)*

1. (a)

$$
\begin{aligned}
L(Y_1, ..., Y_n) &= \frac{P(Y_1, ...Y_N|H_1)}{P(Y_1, ..., Y_n|H_0)} \\
&= e^{\frac{|Y-\mu_0|^2 - |Y-\mu_1|^2}{2\sigma^2}} \\
&= e^{\frac{2(\mu_1-\mu_0)^T Y + ||\mu_0||^2 - ||\mu_1||^2}{2\sigma^2}}.
\end{aligned}
$$

Taking the logarithm on both sides we obtain the simplified decision rule

$$
(\mu_1 - \mu_0)^T Y \underset{H_1}{\overset{H_0}{\lesseqgtr}} \frac{||\mu_1||^2 - ||\mu_0||^2}{2}
$$

If we denote $a = \mu_1 - \mu_0$ and $b = ||\mu_1||^2 - ||\mu_0||^2$ we see that $a^T Y = b$ characterizes an n-dimensional hyperplane which passes through $\frac{\mu_1+\mu_0}{2}$ and separates the two decision regions.

(b) Under H_1,

$$
(\mu_1 - \mu_0)^T Y \sim N\left((\mu_1 - \mu_0)^T \mu_1, \sigma^2 ||\mu_1 - \mu_0||^2\right)
$$

Hence we obtain

$$
\begin{aligned}
P(E|H_1) &= Q\left(\frac{(\mu_1-\mu_0)^T \mu_1 - \left(\frac{||\mu_1||^2 - ||\mu_0||^2}{2}\right)}{\sqrt{\sigma^2 ||\mu_1 - \mu_0||^2}}\right) \\
&= Q\left(\frac{||\mu_1 - \mu_0||}{2\sigma}\right) \\
&= Q\left(\frac{d}{2\sigma}\right),
\end{aligned}
$$

where $d = ||\mu_1 - \mu_0||$ is the distance between the mean vectors. Similarly we can show that $P(E|H_0) = Q\left(\frac{d}{2\sigma}\right)$. Hence, $P(E) = Q\left(\frac{d}{2\sigma}\right)$.

2. (a)

$$L(Y_1,...,Y_n) = \frac{P(Y_1,...,Y_n|H_1)}{P(Y_1,...,Y_n|H_0)}$$
$$= \left(\frac{\sigma_0}{\sigma_1}\right)^n e^{\left(\frac{\sigma_1^2-\sigma_0^2}{2\sigma_1^2\sigma_0^2}\sum_{i=0}^n Y_i^2\right)} \underset{H_1}{\overset{H_0}{\lessgtr}} 1\ .$$

Taking logarithm and using $\sigma_1 > \sigma_0$ we obtain the simplified decision rule

$$\sum_{i=1}^n y_i^2 \underset{H_1}{\overset{H_0}{\lessgtr}} \frac{2\sigma_1^2\sigma_0^2}{\sigma_1^2-\sigma_0^2}\left(-n\log\frac{\sigma_0}{\sigma_1}\right) = \frac{2n\sigma_1^2\sigma_0^2}{\sigma_1^2-\sigma_0^2}\log\left(\frac{\sigma_1}{\sigma_0}\right).$$

The decision boundary is the n-dimensional hyper sphere where the points inside the sphere belong to H_0 whereas the points outside belong to H_1.

(b) i. Let define $Z = Y_1^2 + Y_2^2$. We first obtain the CDF of Z as follows

$$\begin{aligned} F_Z(z) &= P(Y_1^2+Y_2^2 \le z) \\ &= \int_{Y_1^2+Y_2^2\le z} \frac{1}{2\pi\sigma^2}e^{-\frac{y_1^2+y_2^2}{2\sigma^2}}dy_1dy_2 \end{aligned}$$

And we apply the following change of variables: $y_1 = rsin(\theta)$ and $y_2 = rcos(\theta)$ (thus, $r^2 = y_1^2 + y_2^2$) and obtain:

$$\begin{aligned} &= \int_{r=0}^{\sqrt{z}}\int_{\theta=0}^{2\pi}\frac{1}{2\pi\sigma^2}e^{-\frac{r^2}{2\sigma^2}}|detJ(\theta,r)|drd\theta \\ &= \int_{r=0}^{\sqrt{z}}\frac{1}{2\pi\sigma^2}e^{-\frac{r^2}{2\sigma^2}}2\pi r dr \\ &= \left[-e^{-\frac{r^2}{2\sigma^2}}\right]_0^{\sqrt{z}} \\ &= 1-e^{-\frac{z}{2\sigma^2}} \end{aligned}$$

Taking the derivative we obtain the probability density function $f_Z(z) = \frac{1}{2\sigma^2}e^{-\frac{z}{2\sigma^2}}$, which is an exponential with parameter $\frac{1}{2\sigma^2}$.

ii. For $n = 2$ and H_1, we have

$$\begin{aligned} P(E|H_1) &= P\left(Y_1^2+Y_2^2 \ge \frac{4\sigma_1^2\sigma_0^2}{\sigma_1^2-\sigma_0^2}\log(\frac{\sigma_1}{\sigma_0})|H_1\right) \\ &= \int_0^{\frac{4\sigma_1^2\sigma_0^2}{\sigma_1^2-\sigma_0^2}\log(\frac{\sigma_1}{\sigma_0})}\frac{1}{2\sigma_1^2}e^{-\frac{z}{2\sigma_1^2}}dz \\ &= 1-e^{-\frac{4\sigma_0^2}{\sigma_1^2-\sigma_0^2}\log(\frac{\sigma_1}{\sigma_0})} = 1-\left(\frac{\sigma_0}{\sigma_1}\right)^{\frac{4\sigma_0^2}{\sigma_1^2-\sigma_0^2}} \end{aligned}$$

Similarly, we obtain

$$
\begin{aligned}
P(E|H_0) &= \int_{\frac{4\sigma_1^2\sigma_0^2}{\sigma_1^2-\sigma_0^2}\log\frac{\sigma_1}{\sigma_0}}^{\infty} \frac{1}{2\sigma_0^2} e^{-\frac{z}{2\sigma_0^2}} dz \\
&= e^{-\frac{4\sigma_1^2}{\sigma_1^2-\sigma_0^2}\log\frac{\sigma_1}{\sigma_0}} = \left(\frac{\sigma_0}{\sigma_1}\right)^{\frac{4\sigma_1^2}{\sigma_1^2-\sigma_0^2}}
\end{aligned}
$$

Let us denote $\rho = \frac{\sigma_1}{\sigma_0}$. Then we have :

$$
\begin{aligned}
P(E|H_0) = 1 - \left(\frac{1}{\rho}\right)^{\frac{4}{\rho^2-1}}&, \; P(E|H_1) = \left(\frac{1}{\rho}\right)^{\frac{4\rho^2}{\rho^2-1}} \\
\text{and } P(E) = \frac{1}{2} + \frac{1}{2}&\left(\left(\frac{1}{\rho}\right)^{\frac{4\rho^2}{\rho^2-1}} - \left(\frac{1}{\rho}\right)^{\frac{4}{\rho^2-1}}\right)
\end{aligned}
$$

iii. Let us denote $A(\rho) = \left(\frac{1}{\rho}\right)^{\frac{4\rho^2}{\rho^2-1}}$ and $B(\rho) = \left(\frac{1}{\rho}\right)^{\frac{4}{\rho^2-1}}$. Using the Hopital rule we obtain:

$$\lim_{\rho\to\infty} \log B(\rho) = \lim_{\rho\to\infty} \frac{-4\log(\rho)}{\rho^2-1} = 0 \Rightarrow \lim_{\rho\to\infty} B(\rho) = 1, \text{ and } \lim_{\rho\to\infty} A(\rho) = 0^4 = 0$$

Hence

$$\lim_{\rho\to\infty} P(E) = 0.$$

Problem 3. *(Repeat Codes and Bhattacharyya Bound)*

1. First, we find the probability mass function of $(W_1, \dots, W_N)$ given each of the two hypotheses:

$$
\begin{aligned}
p_{W_1 \dots W_N|H}(w_1, \dots, w_N|0) &= Pr\{\text{sgn}(X_1+Z_1) = w_1, \dots, \text{sgn}(X_N+Z_N) = w_N | H = 0\} \\
&= Pr\{\text{sgn}(X_1+Z_1) = w_1, \dots, \text{sgn}(X_N+Z_N) = w_N \;| \\
&\qquad\qquad (X_1, \dots, X_N) = (1, \dots, 1)\} \qquad (1) \\
&= Pr\{\text{sgn}(1+Z_1) = w_1, \dots, \text{sgn}(1+Z_N) = w_N\} \qquad (2) \\
&= Pr\{\text{sgn}(1+Z_1) = w_1\} \cdot \ldots \cdot Pr\{\text{sgn}(1+Z_N) = w_N\}. \quad (3)
\end{aligned}
$$

But since the Z_i are independent of each other and have the same distribution (or, as we say more frequently, since the Z_i are iid random variables), we can write this also as

$$p_{W_1 \dots W_N|H}(w_1, \dots, w_N|0) = \prod_{i=1}^{N} Pr\{\text{sgn}(1+Z) = w_i\}. \qquad (4)$$

Notice that the event $\{W_i = \mathrm{sgn}(1+Z) = 0\}$ has probability zero. Therefore, it is of no interest to our consideration. This means that the random variables W_i can only assume values 1 or -1. Suppose that $(w_1, \ldots, w_N)$ contains k values of 1, and thus $(N-k)$ values of -1. With this definition, we can rewrite

$$p_{W_1 \ldots W_N|H}(w_1, \ldots, w_N|0) \quad = \quad (Pr\{1+Z \geq 0\})^k \, (Pr\{1+Z \leq 0\})^{N-k}. \qquad (5)$$

Let us introduce the following notation:

$$\epsilon \stackrel{def}{=} Pr\{(1+Z) \leq 0\} = 1 - Q\left(-\frac{1}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right). \qquad (6)$$

Then, we can write

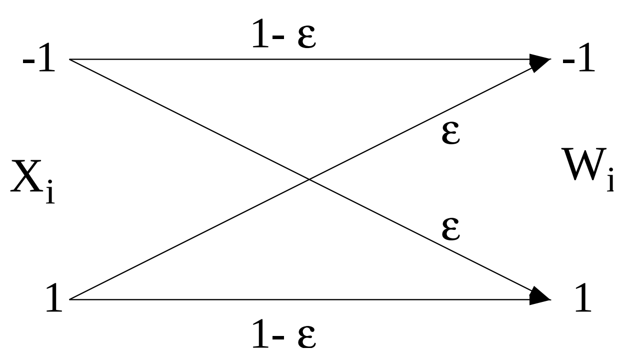
$$p_{W_1 \ldots W_N|H}(w_1, \ldots, w_N|0) \quad = \quad (1-\epsilon)^k \epsilon^{N-k}. \qquad (7)$$

Under hypothesis $H = 1$, we have essentially the same derivation. Let us give only a few steps, using the same definitions of k (number of ones) and ϵ as above:

$$\begin{aligned} p_{W_1 \ldots W_N|H}(w_1, \ldots, w_N|1) &= \prod_{i=1}^{N} Pr\{sign(-1+Z) = w_i\} && (8) \\ &= (Pr\{-1+Z \geq 0\})^k \, (Pr\{-1+Z \leq 0\})^{N-k} && (9) \\ &= \epsilon^k (1-\epsilon)^{N-k} && (10) \end{aligned}$$

Thus, we see that k is a sufficient statistic.

At this point, we give a simple picture that allows to derive all of this much more easily: The overall system between X and W may be viewed as a channel with input 1 or -1 and output also 1 or -1. There is a certain probability ϵ (called *transition* or *crossover* probability) that the channel converts 1 into -1 or vice versa:



This particular channel is called the *Binary Symmetric Channel*. From this figure, various results can be found easily. For instance, it is clear that if we put N consecutive values 1 onto the channel, the probability of getting, at the output, a particular sequence $(w_1, \ldots, w_N)$ which contains exactly k values of 1 is simply $(1-\epsilon)^k \epsilon^{N-k}$.

Similarly, the probability of getting, at the output, *any* sequence that contains exactly k values of 1 is $\binom{N}{k}(1-\epsilon)^k\epsilon^{N-k}$ because there are $\binom{N}{k}$ distinct particular sequences with exactly k ones each, and every one of them has probability $(1-\epsilon)^k\epsilon^{N-k}$.

Next, we have to develop the likelihood ratio:

$$L(w_1,\ldots,w_N) = \frac{p_{W_1\ldots W_N|H}(w_1,\ldots,w_N|1)}{p_{W_1\ldots W_N|H}(w_1,\ldots,w_N|0)} \tag{11}$$

$$= \frac{\epsilon^k(1-\epsilon)^{N-k}}{(1-\epsilon)^k\epsilon^{N-k}} = \left(\frac{\epsilon}{1-\epsilon}\right)^{2k-N}. \tag{12}$$

In terms of the loglikelihood ratio, we find

$$\log L(w_1,\ldots,w_N) = (2k-N)\log\left(\frac{\epsilon}{1-\epsilon}\right) \underset{0}{\overset{1}{\gtrless}} 0. \tag{13}$$

Since $\epsilon < 1/2$, we know that $\log\left(\frac{\epsilon}{1-\epsilon}\right) < 0$, and thus, when we divide by this term, the direction of the inequality is changed. Using this, the decision rule can be written as

$$k \underset{0}{\overset{1}{\lessgtr}} \frac{N}{2}. \tag{14}$$

That is, the best decision rule is simply *majority voting*: if the majority of the received values is 1, we decide for hypothesis $H = 0$ (i.e. transmitted value was 1); on the other hand, if the majority of the received values is -1, we decide for hypothesis $H = 1$ (i.e. transmitted value was -1).

2. Let us assume that N is odd. Then,

$$P_e(0) = Pr\{\text{ there are less than } N/2 \text{ ones in the received sequence } | \; N \text{ ones were transmitted } \} \tag{15}$$

$$= \sum_{m=0}^{(N-1)/2} \binom{N}{m}(1-\epsilon)^m\epsilon^{N-m} \tag{16}$$

By the symmetry of the problem, $P_e(1)$ turns out to be the exact same expression, thus

$$P_e = \sum_{m=0}^{(N-1)/2} \binom{N}{m}(1-\epsilon)^m\epsilon^{N-m} \tag{17}$$

In case N is even, we introduce a slight asymmetry because the term for $N/2$ has to be assigned to either $H = 0$ or $H = 1$ (cannot be assigned to both).

Clearly, this sum cannot be evaluated explicitly. There are various techniques to bound it. In this homework, we consider the *Bhattacharyya bound* as encountered in class.

3. In class, you have seen the following derivation of the Bhattacharyya bound. For the optimal decision rule, we can write the probability of error as follows:

$$P_e \;=\; \sum_w \min_w \left\{ p_{W|H}(w|0)p_H(0), p_{W|H}(w|1)p_H(1) \right\}, \tag{18}$$

where the sum is over all possible sequences w of length N. In our case, since $w_i \in \{-1, 1\}$, we have $w \in \{-1, 1\}^N$, and thus there are 2^N terms in the sum. But since for $a, b \geq 0$, we have that $\min a, b \leq \sqrt{ab}$, we get the following simple upper bound:

$$\begin{aligned} Pr\{e\} \;&\leq\; \sum_w \sqrt{p_{W|H}(w|0)p_H(0)p_{W|H}(w|1)p_H(1)} && (19) \\ &= \sqrt{p_H(0)p_H(1)} \sum_w \sqrt{p_{W|H}(w|0)p_{W|H}(w|1)} && (20) \\ &\leq \frac{1}{2} \sum_w \sqrt{p_{W|H}(w|0)p_{W|H}(w|1)}, && (21) \end{aligned}$$

where the last inequality follows because for $c, d \geq 0$, we have $\sqrt{cd} \leq (c+d)/2$. Now we have to plug in:

$$\begin{aligned} \tilde{P}_e \;&\leq\; \frac{1}{2} \sum_w \sqrt{p_{W|H}(w|0)p_{W|H}(w|1)} && (22) \\ &= \frac{1}{2} \sum_w \sqrt{(1-\epsilon)^{k(w)} \epsilon^{N-k(w)} \;\; \epsilon^{k(w)} (1-\epsilon)^{N-k(w)}}, && (23) \end{aligned}$$

where we used $k(w)$ to denote the number of values 1 in the sequence w. We find furthermore

$$\begin{aligned} \tilde{P}_e \;&\leq\; \frac{1}{2} \sum_w \sqrt{\epsilon^N (1-\epsilon)^N} = \frac{1}{2} \sqrt{\epsilon^N (1-\epsilon)^N} \sum_w 1 && (24) \\ &= \frac{1}{2} \sqrt{\epsilon^N (1-\epsilon)^N} \; 2^N = \frac{1}{2} \left(2\sqrt{\epsilon(1-\epsilon)} \right)^N. && (25) \end{aligned}$$

4. Again, we assume that N is odd; note however that the case when N is even would not add much insight. We used the following `matlab` program:

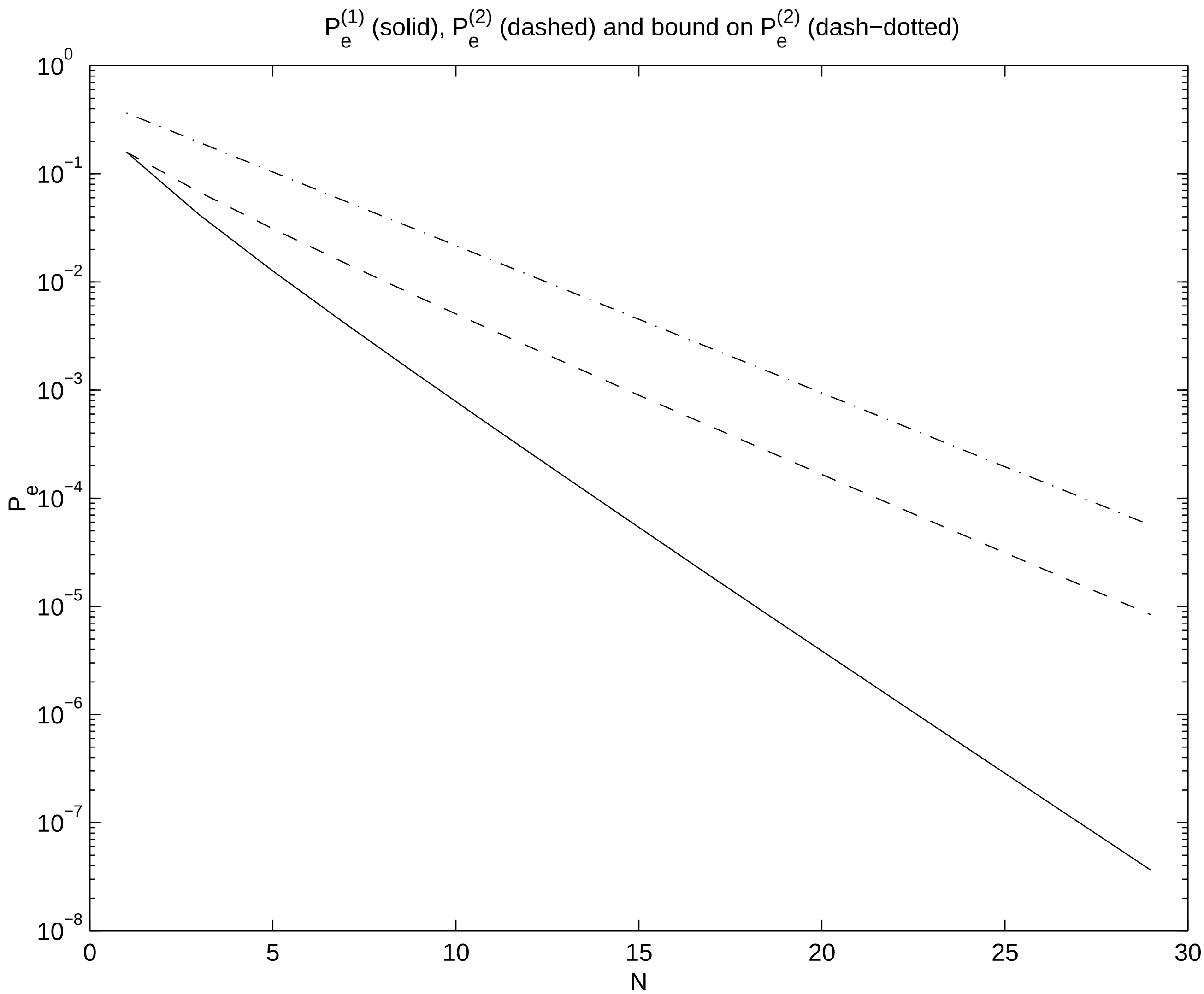
```
%
%  Principles of Digital Communications, Summer Semester 2001 (Prof. B. Rimoldi)
%
%  Solution to Homework 4, Problem 1.(iv)
%  Michael Gastpar
%
%  Notation: we are using Pe1 for $P_e$ and Pe2 for $\tilde P_e$
%
N = [ 1:2:30 ];
sigma = 1;
```

```
Pe1 =  1/2 * erfc(  sqrt(N)/sigma   /sqrt(2))

epsilon = 1/2 * erfc(  1/sigma  /sqrt(2));
Pe2 = zeros(1, length(N));
for ic = 1:length(N),
  for m = 0:(N(ic)-1)/2,
    Pe2(ic) = Pe2(ic) + prod(N(ic)-m+1:N(ic))/prod(1:m)  *  (1-epsilon)^m * epsilon^(N(ic)-m);
  end;
end;
Pe2Bhatt =  1/2 * (2*sqrt(epsilon*(1-epsilon))).^N

semilogy( N, Pe1, N, Pe2Bhatt, '--', N, Pe2, '-.');
title('P_e^{(1)} (solid), P_e^{(2)} (dashed) and bound on P_e^{(2)} (dash-dotted)');
xlabel('N');
ylabel('P_e');
```

which gives the following plot for the error probabilities:



Problem 4. *(Tighter Union Bhattacharyya Bound: Binary Case)*

1. From the definition of the decision region $\mathcal{R}_i$

$$\mathcal{R}_i = \left\{y : P_H(i) f_{Y|H}(y|i) \geq P_H(j) f_{Y|H}(y|j)\right\} \qquad i \neq j,$$

it is easy to see that in region $\mathcal{R}_0$

$$P_H(0) f_{Y|H}(y|0) \geq P_H(1) f_{Y|H}(y|1)$$

and vice-versa. Thus we can write

$$\begin{aligned}
Pr\{e\} &= P_H(0)\int_{\mathcal{R}_1} f_{Y|H}(y|0)dy + P_H(1)\int_{\mathcal{R}_0} f_{Y|H}(y|1)dy \\
&= \int_{\mathcal{R}_1} \min\left\{P_H(0)f_{Y|H}(y|0), P_H(1)f_{Y|H}(y|1)\right\} dy \\
&\quad + \int_{\mathcal{R}_0} \min\left\{P_H(0)f_{Y|H}(y|0), P_H(1)f_{Y|H}(y|1)\right\} dy \\
&= \int_{\mathcal{R}_0+\mathcal{R}_1} \min\left\{P_H(0)f_{Y|H}(y|0), P_H(1)f_{Y|H}(y|1)\right\} dy \\
&= \int_{y} \min\left\{P_H(0)f_{Y|H}(y|0), P_H(1)f_{Y|H}(y|1)\right\} dy.
\end{aligned}$$

2. To show that for $a, b \geq 0, \sqrt{ab} \leq \frac{a+b}{2}$, we proceed as follows. Let $m = (a+b)/2$ be the midpoint of an imaginary segment of the real line that goes from a to b. Let $d = (b-a)/2$ be half the distance between a and b. Writing a and b in terms of m and d we obtain: $ab = (m-d)(m+d) = m^2 - d^2 \leq m^2$ which is the desired result.

   Using this and the hint, namely, for $a, b \geq 0,\ \min(a, b) \leq \sqrt{ab}$, we can write

$$\begin{aligned}
Pr\{e\} &= \int_{y} \min\left\{P_H(0)f_{Y|H}(y|0), P_H(1)f_{Y|H}(y|1)\right\} dy \\
&\leq \int_{y} \sqrt{P_H(0)f_{Y|H}(y|0)P_H(1)f_{Y|H}(y|1)}dy \\
&= \sqrt{P_H(0)P_H(1)}\int_{y} \sqrt{f_{Y|H}(y|0)f_{Y|H}(y|1)}dy \\
&\leq \frac{P_H(0)+P_H(1)}{2}\int_{y} \sqrt{f_{Y|H}(y|0)f_{Y|H}(y|1)}dy \\
&= \frac{1}{2}\int_{y} \sqrt{f_{Y|H}(y|0)f_{Y|H}(y|1)}dy.
\end{aligned}$$

3. In class we upper bounded $Pr\{e|H=0\}$ and $Pr\{e|H=1\}$ individually instead of upperbounding the almost final result $Pr\{e\} = P_H(0)Pr\{e|H=0\} + P_H(1)Pr\{e|H=1\}$, as we did here. More precisely, what we did in class, written differently, is

$$\begin{aligned}
Pr\{e|H=0\} &= \int_{\mathcal{R}_1} f_{Y|H}(y|0)dy \\
&= \int_{\mathcal{R}_1} \min\left\{f_{Y|H}(y|0), f_{Y|H}(y|1)\right\} dy \\
&\leq \int_{\mathcal{R}_1} \sqrt{f_{Y|H}(y|0)f_{Y|H}(y|1)}dy \\
&\leq \int_{\mathbb{R}^n} \sqrt{f_{Y|H}(y|0)f_{Y|H}(y|1)}dy
\end{aligned}$$

The last step, which further loosens the bound, is necessary to find a bound of $Pr\{e|H=0\}$ that does not depend on $\mathcal{R}_1$. This "overbounding" is avoided in *(ii)* by finding the bound over the whole $Pr\{e\}$.

Problem 5. *(Application of Tight Bhattacharyya Bound)*

(i) Using the *Tight Bhattacharyya Bound*, we get

$$\begin{aligned}
Pr\{e\} &\leq \frac{1}{2}\int_y \sqrt{f_{Y|H}(y|0)f_{Y|H}(y|1)}dy \\
&= \frac{1}{2}\int_y \sqrt{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(y+a)^2}{2\sigma^2}\right\}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(y-a)^2}{2\sigma^2}\right\}}dy \\
&= \frac{1}{2}\int_y \frac{1}{\sqrt{2\pi\sigma^2}}\sqrt{\exp\left\{-\frac{y^2+a^2}{\sigma^2}\right\}}dy \\
&= \frac{1}{2}\exp\left\{-\frac{a^2}{2\sigma^2}\right\}\int_y \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{y^2}{2\sigma^2}\right\}dy \\
&= \frac{1}{2}\exp\left\{-\frac{a^2}{2\sigma^2}\right\}.
\end{aligned}$$

(ii) The above bound is the same as the one derived in class, which was obtained working specifically with the expression for the Q-function. It is surprising that the *Bhattacharyya Bound*, which applies to arbitrary channels, yields the same result.

Problem 6. *(Bhattacharyya Bound for DMCs)*

1. Inequality (a) follows from the tight *Bhattacharyya Bound.*

   Using the definition of DMC, it is straightforward to see that

$$\begin{aligned}
P(\mathbf{y}|\mathbf{s_0}) &= \prod_{i=1}^{n} P(y_i|s_{0i}) \qquad \text{and} \\
P(\mathbf{y}|\mathbf{s_1}) &= \prod_{i=1}^{n} P(y_i|s_{1i}).
\end{aligned}$$

   Inequality (b) follows by substituting the above values in (a).

   Equality (c) is obtained by observing that $\sum_{\mathbf{y}}$ is the same as $\sum_{y_1,\ldots,y_n}$ (the first one being a vector notation for the sum over all possible $y_1,\ldots,y_n$).

   In (c) we see that we want the sum of all possible products. It can also be obtained by summing over each y_i and taking the product of the resulting sum for all y_i. This results in inequality (d). We get equality (e) by writing (d) in a more concise form.

When $s_{0i} = s_{1i}$, $\sqrt{P(y|s_{0i})P(y|s_{1i})} = P(y|s_{0i})$ and thus $\sum_y \sqrt{P(y|s_{0i})P(y|s_{1i})} = \sum_y P(y|s_{0i}) = 1$. This would contribute unity to the product (not really useful!). Thus we are only interested in terms where $s_{0i} \neq s_{1i}$. We form the product of all such sums where $s_{0i} \neq s_{1i}$. We then look out for terms where $s_{0i} = a$ and $s_{1i} = b, a \neq b$ and raise the sum to the appropriate power. (Eg. If we have the product $prpqrpqrr$, we would write it as $p^3q^2r^4$). Hence equality (f).

2. For a binary input channel, we have only two source alphabets $\mathcal{X} = \{a, b\}$. Thus

$$\begin{aligned} Pr\{e\} &\leq z^{n(a,b)} z^{n(b,a)} \\ &= z^{n(a,b)+n(b,a)} \\ &= z^{d_H(\mathbf{s_0}, \mathbf{s_1})} \end{aligned}$$

3. The value of

   (a) For a binary input Gaussian channel,

$$\begin{aligned} z &= \int_y \sqrt{f_{Y|X}(y|0) f_{Y|X}(y|1)} dy \\ &= \exp\left(-\frac{E}{2\sigma^2}\right) \end{aligned}$$

   (b) For the Binary Symmetric Channel (BSC),

$$\begin{aligned} z &= \sqrt{P(y=0|x=0)P(y=0|x=1)} + \sqrt{P(y=1|x=0)P(y=1|x=1)} \\ &= 2\sqrt{\delta(1-\delta)}. \end{aligned}$$

   (c) For the Binary Erasure Channel (BEC),

$$\begin{aligned} z &= \sqrt{P(y=0|x=0)P(y=0|x=1)} + \sqrt{P(y=E|x=0)P(y=E|x=1)} \\ &\quad + \sqrt{P(y=1|x=0)P(y=1|x=1)} \\ &= 0 + \delta + 0 \\ &= \delta. \end{aligned}$$