## Homework 3

Solutions

## Problem Assignments

1. (10 points) Note that for $\alpha \geq x$, and since $x \geq 0$,

$$
\begin{aligned}
\alpha^{2} & =(\alpha-x)^{2}+2 \alpha x-x^{2} \\
& \geq(\alpha-x)^{2}+2 x^{2}-x^{2} \\
& =(\alpha-x)^{2}+x^{2}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
Q(x) & =\int_{x}^{\infty} \frac{1}{\sqrt{2 \pi}} e^{\frac{-\alpha^{2}}{2}} d \alpha \\
& \leq \int_{x}^{\infty} \frac{1}{\sqrt{2 \pi}} e^{-\frac{(\alpha-x)^{2}+x^{2}}{2}} d \alpha \\
& =e^{\frac{-x^{2}}{2}} \int_{x}^{\infty} \frac{1}{\sqrt{2 \pi}} e^{-\frac{(\alpha-x)^{2}}{2}} d \alpha \\
& =\frac{1}{2} e^{\frac{-x^{2}}{2}} .
\end{aligned}
$$

2. $(10+10=20$ points $)$
(a) The ML rule states the following decision

$$
f_{\mathbf{y}}(\mathrm{y} \mid \alpha=1) \stackrel{\hat{\alpha}=1}{\gtrless} f_{\mathbf{y}}(\mathrm{y} \mid \alpha=-1) .
$$

After certain manipulations, the maximum likelihood decision rule is

$$
\mathbf{y}^{T} \mathbf{h} \underset{\alpha=-1}{\stackrel{\alpha=1}{\gtrless}} 0 .
$$

And, the probability of error is

$$
\begin{equation*}
P_{e}=Q\left(\frac{\|\mathbf{h}\|}{\sigma}\right) . \tag{1}
\end{equation*}
$$

Note that the error probability depends on the ratio of the signal energy $E_{b}=\|\mathbf{h}\|^{2}$ and the noise power $\frac{N_{0}}{2}=\sigma^{2}$, commonly referred to as the signal to noise ratio (SNR). So, in the literature of communications theory, we often see

$$
P_{e}=Q\left(\sqrt{2 E_{b} / N_{0}}\right)
$$

(b) Adding one more hypothesis complicates the problem a little bit. The likelihood function for each value of $\alpha$ is as follows:

$$
\begin{aligned}
L(\alpha=1 \mid y) & \propto \exp \left(-\frac{1}{2 \sigma^{2}}\|\mathrm{y}-\mathbf{h}\|^{2}\right) \\
L(\alpha=0 \mid \mathrm{y}) & \propto \exp \left(-\frac{1}{2 \sigma^{2}}\|\mathrm{y}\|^{2}\right) \\
L(\alpha=-1 \mid \mathrm{y}) & \propto \exp \left(-\frac{1}{2 \sigma^{2}}\|\mathrm{y}+\mathbf{h}\|^{2}\right)
\end{aligned}
$$

where the multiplicative constant is omitted. The maximum likelihood principle requires us finding the $\alpha$ that is maximal among the three. So, the detector chooses $\hat{\alpha}=1$ if $L(\alpha=1 \mid \mathrm{y})>L(\alpha=0 \mid \mathrm{y})$ and $L(\alpha=1 \mid \mathrm{y})>L(\alpha=$ $-1 \mid y)$. Likewise for choosing $\hat{\alpha}=0$ and $\hat{\alpha}=-1$. With some algebraic efforts, we have the following decision rule:

$$
\begin{array}{cl}
\hat{\alpha}=1 & \text { if } \quad \mathrm{y}^{T} \mathbf{h} \geq \frac{1}{2}\|\mathbf{h}\|^{2} \\
\hat{\alpha}=0 & \text { if } \quad-\frac{1}{2}\|\mathbf{h}\|^{2} \leq \mathrm{y}^{T} \mathbf{h} \leq \frac{1}{2}\|\mathbf{h}\|^{2} \\
\hat{\alpha}=-1 & \text { if } \quad \mathrm{y}^{T} \mathbf{h} \leq-\frac{1}{2}\|\mathbf{h}\|^{2} .
\end{array}
$$

To compute the error probability, we need

$$
\begin{aligned}
P[\operatorname{error} \mid \alpha=1] & =P\left[\left.\mathrm{y}^{T} \mathbf{h}<\frac{1}{2}\|\mathbf{h}\|^{2} \right\rvert\, \alpha=1\right] \\
& =P\left[(\mathbf{h}+\mathbf{z})^{T} \mathbf{h}<\frac{1}{2}\|\mathbf{h}\|^{2}\right] \\
& =Q\left(\frac{\|\mathbf{h}\|}{2 \sigma}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
P[\operatorname{error} \mid \alpha=0] & =1-P\left[\left.-\frac{1}{2}\|\mathbf{h}\|^{2} \leq \mathrm{y}^{T} \mathbf{h} \leq \frac{1}{2}\|\mathbf{h}\|^{2} \right\rvert\, \alpha=0\right] \\
& =1-P\left[-\frac{1}{2}\|\mathbf{h}\|^{2} \leq \mathbf{z}^{T} \mathbf{h} \leq \frac{1}{2}\|\mathbf{h}\|^{2}\right] \\
& =1-\left(Q\left(-\frac{\|\mathbf{h}\|}{2 \sigma}\right)-Q\left(\frac{\|\mathbf{h}\|}{2 \sigma}\right)\right) \\
& =2 Q\left(\frac{\|\mathbf{h}\|}{2 \sigma}\right)
\end{aligned}
$$

We also need $P[\operatorname{error} \mid \alpha=-1]$, which is equal to $P[\operatorname{error} \mid \alpha=1]$ by symmetry. So, the probability of error is

$$
\begin{aligned}
P_{e} & =P[\operatorname{error} \mid \alpha=1] P[\alpha=1]+P[\operatorname{error} \mid \alpha=0] P[\alpha=0]+P[\operatorname{error} \mid \alpha=-1] P[\alpha=-1] \\
& =p_{1} \cdot Q\left(\frac{\|\mathbf{h}\|}{2 \sigma}\right)+2 p_{0} \cdot Q\left(\frac{\|\mathbf{h}\|}{2 \sigma}\right)+\left(1-p_{0}-p_{1}\right) \cdot Q\left(\frac{\|\mathbf{h}\|}{2 \sigma}\right) \\
& =\left(1+p_{0}\right) \cdot Q\left(\frac{\|\mathbf{h}\|}{2 \sigma}\right)
\end{aligned}
$$

where $p_{0}, p_{1}$, and $1-p_{0}-p_{1}$ are the prior probabilities of $\alpha=1, \alpha=0$, and $\alpha=-1$, respectively. When the prior probabilities are equal, the error probability is

$$
P_{e}=\frac{4}{3} Q\left(\frac{\|\mathbf{h}\|}{2 \sigma}\right)
$$

3. $(10 \times 4=40$ points $)$ This problem intends to let you have a taste of what "spacetime coding" is about in a multiple-transmit and receive antennas system. Basically, when equipping with more antennas, the system's error performance can be improved as compared to the single antenna counterpart. More specifically, when the number of antennas grows, this benefit also gets prominent as the error probability curve (versus signal-to-noise ratio) decreases in the large SNR regime at a rate at most proportional to the product of the number of transmit antennas and the number of receive antennas, as we shall exploit in this problem. This is the maximal achievable diversity gain of the system.
(a) Let $\mathbf{y}=\left[\mathbf{y}_{1}^{t}, \mathbf{y}_{2}^{t}, \ldots, \mathbf{y}_{T}^{t}\right]^{t}$, where the superscript means transpose. The ML detection rule for $\mathbf{S}=\left[\mathbf{s}_{1}, \mathbf{s}_{2}, \ldots, \mathbf{s}_{T}\right]$ is

$$
\hat{\mathbf{S}}_{M L}=\arg \max _{\mathbf{s}_{1}, \mathbf{s}_{2}, \ldots, \mathbf{s}_{T}} \log f_{\mathbf{y}}(\boldsymbol{y} \mid \mathbf{S})
$$

where $f_{\mathbf{y}}(\boldsymbol{y} \mid \mathbf{S})=f_{\mathbf{y}}\left(\boldsymbol{y} \mid \mathbf{s}_{1}, \mathbf{s}_{2}, \ldots, \mathbf{s}_{T}\right)=\prod_{i=1}^{T} f_{\mathbf{y}_{i}}\left(\boldsymbol{y}_{i} \mid \mathbf{s}_{i}\right)$ is the likelihood function.
Since the receiver is assumed to know the channel $\mathbf{H}$, we have

$$
f_{\mathbf{y}_{i}}\left(\boldsymbol{y}_{i} \mid \mathbf{s}_{i}\right)=\frac{1}{\pi^{M} N_{0}{ }^{M}} \exp \left\{-\frac{1}{N_{0}}\left\|\mathbf{y}_{i}-\sqrt{\frac{E_{s}}{N_{0}}} \mathbf{H s}_{i}\right\|^{2}\right\}
$$

It follows that

$$
\begin{aligned}
\hat{\mathbf{S}}_{M L} & =\arg \max _{\mathbf{s}_{1}, \mathbf{s}_{2}, \ldots, \mathbf{s}_{T}} \sum_{i=1}^{T} \log f_{\mathbf{y}_{i}}\left(\boldsymbol{y}_{i} \mid \mathbf{s}_{i}\right) \\
& =\arg \min _{\mathbf{s}_{1}, \mathbf{s}_{2}, \ldots, \mathbf{s}_{T}} \sum_{i=1}^{T}\left\|\mathbf{y}_{i}-\sqrt{\frac{E_{s}}{N_{0}}} \mathbf{H s}_{i}\right\|^{2} \\
& =\arg \min _{\mathbf{S}}\left\|\mathbf{Y}-\sqrt{\frac{E_{s}}{N}} \mathbf{H S}\right\|^{2} .
\end{aligned}
$$

(b) Let's see a FACT before we go into the derivation.

FACT
Let $\mathbf{W}$ be an $M \times T$ matrix where all entries are independent with $\mathbf{w}_{i, j} \sim$ $\mathcal{C N}\left(0, N_{0}\right)$. Then, $\operatorname{Tr}(\mathbf{G W})$ is a complex gaussian random variable with

$$
\operatorname{Tr}(\mathbf{G} \mathbf{W}) \sim \mathcal{C N}\left(0, N_{0} \cdot\|\mathbf{G}\|_{F}^{2}\right),
$$

where $\mathbf{G}$ is a non-random, possibly complex, $T \times M$ matrix.
(proof)

$$
\begin{aligned}
\operatorname{Tr}(\mathbf{G W}) & =\sum_{p=1}^{T}(\mathbf{G} \mathbf{W})_{p p} \\
& =\sum_{p=1}^{T} \sum_{q=1}^{T} \mathbf{G}_{p q} \mathbf{W}_{q p}
\end{aligned}
$$

from which we know $\operatorname{Tr}(\mathbf{G W})$ is a linear combination of independent complex gaussian random variable, and is therefore a complex gaussian R.V. It's straightforward to see the mean of $\operatorname{Tr}(\mathbf{G W})$ is zero.
The variance is

$$
\begin{aligned}
\operatorname{Var}(\operatorname{Tr}(\mathbf{G W})) & =\operatorname{Var}\left(\sum_{p} \sum_{q} \mathbf{G}_{p q} \mathbf{W}_{q p}\right)=\sum_{p} \sum_{q}\left|\mathbf{G}_{p q}\right|^{2} \operatorname{Var}\left(\mathbf{W}_{q p}\right) \\
& =N_{0} \sum_{p} \sum_{q}\left|\mathbf{G}_{p q}\right|^{2} \\
& =N_{0}\|\mathbf{G}\|_{F}^{2} .
\end{aligned}
$$

With the fact, we can obtain the conditional pairwise error probability (PEP) as

$$
\begin{aligned}
& P\left(\mathbf{S}^{(i)} \rightarrow \mathbf{S}^{(j)} \mid \mathbf{H}\right)=P\left(\left.\left\|\mathbf{Y}-\sqrt{\frac{E_{s}}{N}} \mathbf{H} \mathbf{S}^{(j)}\right\|_{F}^{2}<\left\|\mathbf{Y}-\sqrt{\frac{E_{s}}{N}} \mathbf{H} \mathbf{S}^{(i)}\right\|_{F}^{2} \right\rvert\, \mathbf{H}, \mathbf{S}=\mathbf{S}^{(i)}\right) \\
&=P(\operatorname{Tr}[(\mathbf{Y}\left.\left.-\sqrt{\frac{E_{s}}{N}} \mathbf{H} \mathbf{S}^{(j)}\right)^{H}\left(\mathbf{Y}-\sqrt{\frac{E_{s}}{N}} \mathbf{H} \mathbf{S}^{(j)}\right)\right] \\
&\left.\left.<\operatorname{Tr}\left[\left(\mathbf{Y}-\sqrt{\frac{E_{s}}{N}} \mathbf{H} \mathbf{S}^{(i)}\right)^{H}\left(\mathbf{Y}-\sqrt{\frac{E_{s}}{N}} \mathbf{H} \mathbf{S}^{(i)}\right)\right] \right\rvert\, \mathbf{H}, \mathbf{S}=\mathbf{S}^{(i)}\right) \\
&=P\left(\operatorname{Tr}\left[\left(\mathbf{S}^{(i)}-\mathbf{S}^{(j)}\right)^{H} \mathbf{H}^{H} \mathbf{Y}+\mathbf{Y}^{H} \mathbf{H}\left(\mathbf{S}^{(i)}-\mathbf{S}^{(j)}\right)\right]\right. \\
&\left.\left.>\sqrt{\frac{E_{s}}{N}} \operatorname{Tr}\left[\mathbf{S}^{(i)^{H}} \mathbf{H}^{H} \mathbf{H} \mathbf{S}^{(i)}-\mathbf{S}^{(j)^{H}} \mathbf{H}^{H} \mathbf{H} \mathbf{S}^{(j)}\right] \right\rvert\, \mathbf{H}, \mathbf{S}=\mathbf{S}^{(i)}\right)
\end{aligned}
$$

By inserting $\mathbf{Y}=\sqrt{\frac{E_{s}}{N}} \mathbf{H} \mathbf{S}^{(i)}+\mathbf{W}$ into the above and having rearrangement of several terms, it follows

$$
\begin{aligned}
P\left(\mathbf{S}^{(i)} \rightarrow \mathbf{S}^{(j)} \mid \mathbf{H}\right)= & P\left(\operatorname{Tr}\left[\mathbf{E}_{i, j}^{H} \mathbf{H}^{H} \mathbf{W}+\mathbf{W}^{H} \mathbf{H} \mathbf{E}_{i, j}\right]\right. \\
& \left.\quad>\sqrt{\frac{E_{s}}{N}} \operatorname{Tr}\left[\left(\mathbf{S}^{(i)}-\mathbf{S}^{(j)}\right)^{H} \mathbf{H}^{H} \mathbf{H}\left(\mathbf{S}^{(i)}-\mathbf{S}^{(j)}\right)\right]\right) \\
= & P\left(\left.2 \operatorname{Re}\left(\operatorname{Tr}\left[\mathbf{E}_{i, j}^{H} \mathbf{H}^{H} \mathbf{W}\right]\right)>\sqrt{\frac{E_{s}}{N}}\left\|\mathbf{H E} \mathbf{E}_{i, j}\right\|_{F}^{2} \right\rvert\, \mathbf{H}\right),
\end{aligned}
$$

where $\operatorname{Re}(\cdot)$ means the real part operator and, by the fact, $\operatorname{Tr}\left[\mathbf{E}_{i, j}^{H} \mathbf{H}^{H} \mathbf{W}\right]$ is a complex gaussian R.V. with mean 0 and variance $N_{0}\left\|\mathbf{E}_{i, j}^{H} \mathbf{H}^{H}\right\|_{F}^{2}=$ $N_{0}\left\|\mathbf{H E}_{i, j}\right\|_{F}^{2}$.
We can let $X=2 \operatorname{Re}\left(\operatorname{Tr}\left(\mathbf{E}_{i, j}^{H} \mathbf{H}^{H} \mathbf{W}\right)\right)$. We know $X \sim \mathcal{N}\left(0,2 N_{0}\left\|\mathbf{H E}_{i, j}\right\|_{F}^{2}\right)$.
Thus, the conditional PEP is given by

$$
\begin{aligned}
P\left(\mathbf{S}^{(\mathbf{i})} \rightarrow \mathbf{S}^{(\mathbf{j})} \mid \mathbf{H}\right) & =P\left(X>\sqrt{\frac{E_{s}}{N}}\left\|\mathbf{H E}_{i, j}\right\|_{F}^{2}\right) \\
& =P\left(\frac{X}{\sigma_{X}}>\sqrt{\frac{E_{s}}{2 N_{0} N}}\left\|\mathbf{H E}_{i, j}\right\|_{F}^{2}\right) \\
& =Q\left(\sqrt{\frac{\rho\left\|\mathbf{H} \mathbf{E}_{i, j}\right\|_{F}^{2}}{2 N}}\right),
\end{aligned}
$$

where $\frac{X}{\sigma_{X}}$ is a standard gaussian R.V.
(c) The average PEP can be upper bounded as

$$
\begin{aligned}
P\left(\mathbf{S}^{(i)} \rightarrow \mathbf{S}^{(j)}\right) & =E\left[P\left(\mathbf{S}^{(i)} \rightarrow \mathbf{S}^{(j)} \mid \mathbf{H}\right)\right] \\
& \leq E\left[e^{-\frac{\rho\left\|\mathbf{H E} \mathbf{H}_{i, j}\right\|_{F}^{2}}{4 N}}\right]
\end{aligned}
$$

in which the random term is $\left\|\mathbf{H E}_{i, j}\right\|_{F}^{2}$.

By the definition of Frobenious norm, we know

$$
\begin{aligned}
\left\|\mathbf{H E}_{i, j}\right\|_{F}^{2} & =\left\|\left(\mathbf{H E}_{i, j}\right)^{H}\right\|_{F}^{2} \\
& =\left\|\mathbf{E}_{i, j}^{H} \mathbf{H}^{H}\right\|_{F}^{2} \\
& =\sum_{p=1}^{M}\left\|\mathbf{E}_{i, j}^{H} \mathbf{h}_{p}\right\|^{2}
\end{aligned}
$$

where $\mathbf{h}_{p}$ is the $p$ th column of $\mathbf{H}^{H}$ with $\mathbf{h}_{p} \sim \mathcal{C N}\left(\mathbf{0}, \mathbf{I}_{N}\right)$. Therefore we have

$$
P\left(\mathbf{S}^{(i)} \rightarrow \mathbf{S}^{(j)}\right) \leq E\left[\exp \left(-\frac{\rho}{4 N} \sum_{p=1}^{M}\left\|\mathbf{E}_{i, j}^{H} \mathbf{h}_{p}\right\|^{2}\right)\right]
$$

where the average is taken over $\mathbf{h}=\left[\mathbf{h}_{1}^{t}, \mathbf{h}_{2}^{t}, \ldots, \mathbf{h}_{M}^{t}\right]^{t}$ with a joint density

$$
f_{\mathbf{h}}(\boldsymbol{h})=\prod_{p=1}^{M} \frac{1}{\pi^{N}} \exp \left(-\boldsymbol{h}_{p}^{H} \boldsymbol{h}_{p}\right)
$$

Thus, it yields

$$
\begin{aligned}
P\left(\mathbf{S}^{(i)} \rightarrow \mathbf{S}^{(j)}\right) & \leq E\left[\exp \left(-\frac{\rho}{4 N} \sum_{p=1}^{M}\left\|\mathbf{E}_{i, j}^{H} \mathbf{h}_{p}\right\|^{2}\right)\right] \\
& =\int_{-\infty}^{\infty} \exp \left\{-\frac{\rho}{4 N} \sum_{p=1}^{M} \mathbf{h}_{p}^{H} \mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H} \mathbf{h}_{p}\right\} \cdot f_{\mathbf{h}}(\boldsymbol{h}) d \boldsymbol{h} \\
& =\int_{-\infty}^{\infty} \frac{1}{\pi^{M N}} \exp \left\{-\sum_{p=1}^{M} \boldsymbol{h}_{p}^{H}\left(\mathbf{I}_{N}+\frac{\rho}{4 N} \mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}\right) \boldsymbol{h}_{p}\right\} d \boldsymbol{h}_{p}
\end{aligned}
$$

where the integration in the 2 nd equality is taken with respect to all components in $\boldsymbol{h}$ and, likewise, the integration is with respect to all entries in $\boldsymbol{h}_{p}$ in the 3rd equality. Setting $\mathbf{K}^{-1}=\mathbf{I}_{N}+\frac{\rho}{4 N} \mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}$ gives

$$
\begin{aligned}
P\left(\mathbf{S}^{(\mathbf{i})} \rightarrow \mathbf{S}^{(\mathbf{j})}\right) & \leq \int_{-\infty}^{\infty} \frac{1}{\pi^{M N}} \exp \left(-\sum_{p=1}^{M} \boldsymbol{h}_{p}^{H} \mathbf{K}^{-1} \boldsymbol{h}_{p}\right) d \boldsymbol{h}_{p} \\
& =\int_{-\infty}^{\infty} \frac{1}{\pi^{M N} \operatorname{det}(\mathbf{K})^{M}} \exp \left(-\sum_{p=1}^{M} \boldsymbol{h}_{p}^{H} \mathbf{K}^{-1} \boldsymbol{h}_{p}\right) d \boldsymbol{h}_{p} \cdot \operatorname{det}(\mathbf{K})^{M} \\
& =\underbrace{\int_{-\infty}^{\infty} \prod_{p=1}^{M} \frac{1}{\pi^{N} \operatorname{det}(\mathbf{K})} \exp \left(-\boldsymbol{h}_{p}^{H} \mathbf{K}^{-1} \boldsymbol{h}_{p}\right) d \boldsymbol{h}_{p} \cdot \operatorname{det}(\mathbf{K})^{M}}_{=1 \text { (intergral of a density from }-\infty \text { to } \infty)} \\
& =\operatorname{det}(\mathbf{K})^{M} \\
& =\left(\frac{1}{\operatorname{det}\left(\mathbf{I}_{N}+\frac{\rho}{4 N} \mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}\right)}\right)^{M}
\end{aligned}
$$

(d) Since $\mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}$ is a Hermitian matrix, we can decompose it to

$$
\mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}=\mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^{H}
$$

where $\mathbf{U}$ is a unity matrix consisting of eigenvectors of $\mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}$ and $\boldsymbol{\Lambda}=$ $\operatorname{diag}\left(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{N}\right)$ is a diagonal matrix in which we arrange the eigenvalues
in a decreasing order $\lambda_{1} \geq \lambda_{2} \geq \ldots \geq \lambda_{N} \geq 0$. If $\operatorname{rank}\left(\mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}\right)=r \leq N$, then there are $r$ nonzero eigenvalues $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{r}$.
We then can represent the PEP as

$$
\begin{aligned}
P\left(\mathbf{S}^{(i)} \rightarrow \mathbf{S}^{(j)}\right) & \leq\left(\operatorname{det}\left(\mathbf{I}_{N}+\frac{\rho}{4 N} \mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}\right)\right)^{-M} \\
& =\left(\prod_{k=1}^{r}\left(1+\frac{\rho}{4 N} \lambda_{k}\right)\right)^{-M}
\end{aligned}
$$

For SNR $\rho \gg 1$, the PEP approaches

$$
P\left(\mathbf{S}^{(i)} \rightarrow \mathbf{S}^{(j)}\right) \leq\left(\frac{\rho}{4 N}\right)^{-r M}\left(\prod_{k=1}^{r} \lambda_{k}\right)^{-M}
$$

Note that when the matrix $\mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}$ has full rank $N$, the error probability scales with $\rho^{-N M}$, in which the product $N M$ is the maximum achievable diversity gain, which can be realized by proper designs of $\mathbf{E}_{i, j}$ that makes the matrix $\mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}$ of full-rank, or equivalently the matrix $\mathbf{E}_{i, j}=\mathbf{S}^{(i)}-\mathbf{S}^{(j)}$ of full rank $N$ for $T \geq N$. (since $\operatorname{rank}\left(\mathbf{E}_{i, j}\right)=\operatorname{rank}\left(\mathbf{E}_{i, j} \mathbf{E}_{i, j}^{H}\right)$. We see the design of $\mathbf{E}_{i, j}$ is across spatial domain over $N$ multiple transmit antennas and across temporal domain over $T$ transmit time slots. Thus, we call the design of the codeword matrix $\mathbf{E}_{i, j}$ as "space-time coding." Current WLAN standard IEEE 802.11n adopts the space-time technology in wireless routers, as you can see the most updated wireless routers in the market are all equipped with multiple antennas.
4. $(10+10=20$ points $)$
(a) The number $X$ of Poisson events with parameter $\lambda$ occurred in the time span $[0, t]$ has the probability mass function

$$
P[X(t)=n]=e^{-\lambda t} \frac{(\lambda t)^{n}}{n!}
$$

Let $A$ be the event that the number of Poisson events occurred in the interval $[0, y]$ is greater than $k$. The probability distribution function is

$$
\begin{aligned}
F_{Y_{k}}(y) & =P\left[Y_{k} \leq y\right] \\
& =P[A] \\
& =P[X(y) \geq k] \\
& =1-P[X(y)<k] \\
& =1-\sum_{n=0}^{k-1} e^{-\lambda y} \frac{(\lambda y)^{n}}{n!}
\end{aligned}
$$

(b) The probability density function of $Y_{k}$ is

$$
\begin{aligned}
f_{Y_{k}}(y) & =d F_{Y_{k}}(y) / d y \\
& =\frac{(\lambda y)^{k-1} \lambda e^{-\lambda y}}{(k-1)!}
\end{aligned}
$$

5. (10 points) By checking the following expansion

$$
\begin{aligned}
\sum_{i=1}^{n}\left(X_{i}-\bar{X}_{n}\right)^{2} & =\sum_{i=1}^{n}\left(X_{i}-\mu-\left(\bar{X}_{n}-\mu\right)\right)^{2} \\
& =\sum_{i=1}^{n}\left(X_{i}-\mu\right)^{2}+\sum_{i=1}^{n}\left(\bar{X}_{n}-\mu\right)^{2}-2 \sum_{i=1}^{n}\left(X_{i}-\mu\right)\left(\bar{X}_{n}-\mu\right) \\
& =\sum_{i=1}^{n}\left(X_{i}-\mu\right)^{2}+n\left(\bar{X}_{n}-\mu\right)^{2}-2\left(n \bar{X}_{n}-n \mu\right)\left(\bar{X}_{n}-\mu\right) \\
& =\sum_{i=1}^{n}\left(X_{i}-\mu\right)^{2}-n\left(\bar{X}_{n}-\mu\right)^{2}
\end{aligned}
$$

and taking the expectation

$$
\begin{aligned}
E\left[\sum_{i=1}^{n}\left(X_{i}-\bar{X}_{n}\right)^{2}\right] & =\sum_{i=1}^{n} E\left[\left(X_{i}-\mu\right)^{2}\right]-n E\left[\left(\bar{X}_{n}-\mu\right)^{2}\right] \\
& =n \sigma^{2}-n \cdot \operatorname{Var}\left(\bar{X}_{n}\right)=(n-1) \sigma^{2},
\end{aligned}
$$

we can see the unbiasedness of the sample variance.

## Extra Problems

You do NOT need to turn in your solutions for the following problems. However, I strongly encourage you to work through them.

1. (a) We see $Z+W \sim \mathcal{N}\left(0, A^{2}+\sigma^{2}\right)$. So, the ML rule is simply choosing $\hat{X}=+1$ if $y>0$ and $\hat{X}=-1$ otherwise. And the error probability

$$
P_{e}=Q\left(\frac{1}{\sqrt{A^{2}+\sigma^{2}}}\right) \rightarrow 0.5 \text { as } A \rightarrow \infty
$$

(b) The likelihood ratio is

$$
\begin{aligned}
\Lambda(y) & =\frac{\exp \left(-\frac{\left(y-(1+A)^{2}\right)}{2 \sigma^{2}}\right)+\exp \left(-\frac{\left(y-(1-A)^{2}\right)}{2 \sigma^{2}}\right)}{\exp \left(-\frac{\left(y-(-1+A)^{2}\right)}{2 \sigma^{2}}\right)+\exp \left(-\frac{\left(y-(-1-A)^{2}\right)}{2 \sigma^{2}}\right)} \\
& =\exp \left(\frac{2 y}{\sigma^{2}}\right) \frac{\cosh \left(\frac{A(y-1)}{\sigma^{2}}\right)}{\cosh \left(\frac{A(y+1)}{\sigma^{2}}\right)}
\end{aligned}
$$

So, the ML rule is choosing $\hat{X}=+1$ if $\Lambda(y)>1$ and $\hat{X}=-1$ otherwise.
(c) When $A \rightarrow \infty$, we have

$$
\Lambda(y) \rightarrow \begin{cases}\exp \left(\frac{2 y}{\sigma^{2}}\right) \exp \left(\frac{-2 A}{\sigma^{2}}\right)=\exp \left(\frac{2(y-A)}{\sigma^{2}}\right) & \text { if } y>1 \\ \exp \left(\frac{2 y}{\sigma^{2}}\right) \exp \left(\frac{-2 A y}{\sigma^{2}}\right)=\exp \left(\frac{2(1-A) y}{\sigma^{2}}\right) & \text { if }-1<y<1 \\ \exp \left(\frac{2 y}{\sigma^{2}}\right) \exp \left(\frac{2 A}{\sigma^{2}}\right)=\exp \left(\frac{2(y+A)}{\sigma^{2}}\right) & \text { if } y<-1\end{cases}
$$

And, the ML rule becomes

$$
\begin{aligned}
& \hat{X}=+1, \text { if } y>A \text { or }-A<y<0 \\
& \hat{X}=-1, \text { if } y<-A \text { or } 0<y<+A
\end{aligned}
$$

It follows the error probability is

$$
\begin{aligned}
P_{e}= & \frac{1}{4} P(\hat{X}=+1 \mid X=-1, Z=A)+\frac{1}{4} P(\hat{X}=+1 \mid X=-1, Z=-A) \\
& +\frac{1}{4} P(\hat{X}=-1 \mid X=+1, Z=A)+\frac{1}{4} P(\hat{X}=-1 \mid X=+1, Z=-A) \\
= & Q\left(\frac{1}{\sigma}\right)+\frac{1}{2}\left(Q\left(\frac{A-1}{\sigma}\right)+Q\left(\frac{A+1}{\sigma}\right)\right)-\frac{1}{2}\left(Q\left(\frac{2 A-1}{\sigma}\right)+Q\left(\frac{2 A+1}{\sigma}\right)\right) \\
\rightarrow & Q\left(\frac{1}{\sigma}\right) \text { as } A \rightarrow \infty .
\end{aligned}
$$

(d) In (a), interference is Gaussian so it acts like noise. In (c), interference has structure so when it is very strong, one can make use of its structure to distinguish it from the desired signal.

