Chapter 1

Solutions to Exercises

Chapter 2

Solutions to Exercises

EXERCISE 2.1. 1. Let $r(t) = \sqrt{r_0^2 + (vt)^2 + 2r_0 vt\cos(\phi)}$. Then,

$$E_r(f, t, r(t), \theta, \psi)) = \frac{\Re[\alpha(\theta, \psi, f)\exp\{j2\pi f(1 - r(t)/c)\}]}{r(t)}.$$

Moreover, if we assume that $r_0 \gg vt$, then we get that $r(t) \approx r_0 + vt\cos(\phi)$. Thus, the doppler shift is $fv\cos(\phi)/c$.

2. Let (x, y, z) be the position of the mobile in Cartesian coordinates, and (r, ψ, θ) the position in polar coordinates. Then

$$\begin{aligned}
(x, y, z) &= (r\sin\theta\cos\psi, r\sin\theta\sin\psi, r\cos\theta) \\
(r, \psi, \theta) &= \left(\sqrt{x^2 + y^2 + z^2}, \arctan(y/x), \arccos(z/\sqrt{x^2 + y^2 + z^2})\right) \\
\dot{\psi} &= \frac{x\dot{y} - \dot{x}y}{x^2 + y^2} \\
\dot{\theta} &= -\frac{\dot{z}r - z\dot{r}}{r^2\sqrt{1 - (z/r)^2}}
\end{aligned}$$

We see that $\dot{\psi}$ is small for large $x^2 + y^2$. Also $\dot{\theta}$ is small for $|z/r| < 1$ and r large. If $|r/z| = 1$ then $\theta = 0$ or $\theta = \pi$ and $v <= r|\dot{\theta}|$ so v/r large assures that $\dot{\theta}$ is small. If r is not very large then the variation of θ and ψ may not be negligible within the time scale of interest even for moderate speeds v. Here large depends on the time scale of interest.

EXERCISE 2.2.

$$\begin{aligned}
E_r(f, t) &= \frac{\alpha\cos[2\pi f(t - r(t)/c)]}{2d - r(t)} + \frac{2\alpha[d - r(t)]\cos[2\pi f(t - r(t)/c)]}{r(t)[2d - r(t)]} \\
&\quad -\frac{\alpha\cos[2\pi f(t + (r(t) - 2d)/c)]}{2d - r(t)}
\end{aligned}$$

$$= \frac{2\alpha \sin\left[2\pi f\left(t - d/c\right)\right] \sin\left[2\pi f\left(r(t) - d\right)/c\right]}{2d - r(t)} + \frac{2\alpha\left[d - r(t)\right]\cos\left[2\pi f\left(t - r(t)/c\right)\right]}{r(t)[2d - r(t)]} \tag{2.1}$$

where we applied the identity

$$\cos x - \cos y = 2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{y-x}{2}\right)$$

We observe that the first term of (2.1) is similar in form to equation (2.13) in the notes. The second term of (2.1) goes to 0 as $r(t) \to d$ and is due to the difference in propagation losses in the 2 paths.

Exercise 2.3. If the wall is on the other side, both components arrive at the mobile from the left and experience the same Doppler shift.

$$E_r(f,t) = \frac{\Re[\alpha \exp\{j2\pi[f(1-v/c)t - fr_0/c]\}]}{r_0 + vt} - \frac{\Re[\alpha \exp\{j2\pi[f(1-v/c)t - f(r_0+2d)/c]\}]}{r_0 + 2d + vt}$$

We have the interaction of 2 sinusoidal waves of the same frequency and different amplitude.

Over time, we observe the composition of these 2 waves into a single sinusoidal signal of frequency $f(1-v/c)$ and constant amplitude that depends on the attenuations $(r_0 + vt)$ and $(r_0 + 2d + vt)$ and also on the phase difference $f2d/c$.

Over frequency, we observe that when $f2d/c$ is an integer both waves interfere destructively resulting in a small received signal. When $f2d/c = (2k+1)/2, k \in Z$ these waves interfere constructively resulting in a larger received signal. So when f is varied by $c/4d$ the amplitude of the received signal varies from a minimum to a maximum.

The variation over frequency is similar in nature to that of section 2.1.3, but since the delay spread is different the coherence bandwidth is also different.

However there is no variation over time because the Doppler spread is zero.

Exercise 2.4. 1. i) With the given information we can compute the Doppler spread:

$$D_s = |f_1 - f_2| = \frac{fv}{c}|\cos\theta_1 - \cos\theta_2|$$

from which we can compute the coherence time

$$T_c = \frac{1}{4D_s} = \frac{c}{4fv|\cos\theta_1 - \cos\theta_2|}$$

ii) There is not enough information to compute the coherence bandwidth, as it depends on the delay spread which is not given. We would need to know the difference in path length to compute the delay spread T_d and use it to compute W_c.

2. From part 1 we see that a larger angular range results in larger delay spread and smaller coherence time. Then, in the richly scattered environment the channel would show a smaller coherence time than in the environment where the reflectors are clustered in a small angular range.

EXERCISE 2.5. 1.

$$\begin{aligned} r_1 &= \sqrt{r^2 + (h_s - h_r)^2} = r\sqrt{1 + (h_s - h_r)^2/r^2} \approx r(1 + \frac{(h_s - h_r)^2}{2r^2}) \\ r_2 &= \sqrt{r^2 + (h_s + h_r)^2} = r\sqrt{1 + (h_s + h_r)^2/r^2} \approx r(1 + \frac{(h_s + h_r)^2}{2r^2}) \\ r_2 - r_1 &\approx \frac{(h_s + h_r)^2 - (h_s - h_r)^2}{2r} = \frac{h_s^2 + h_r^2 + 2h_sh_r - h_s^2 - h_r^2 + 2h_sh_r}{2r} \\ &= \frac{2h_sh_r}{r} \end{aligned}$$

Therefore $b = 2h_sh_r$.

2.

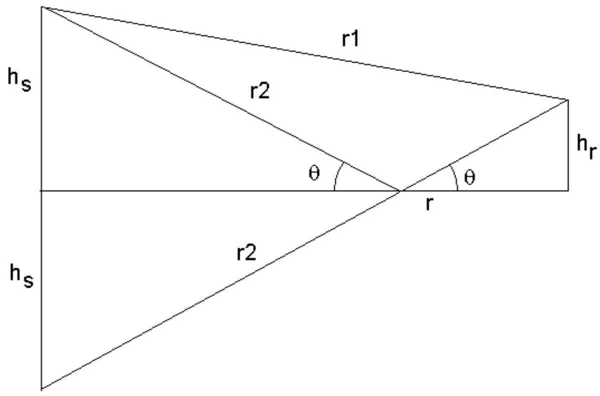
$$\begin{aligned} E_r(f,t) &\approx \frac{Re[\alpha[\exp\{j2\pi(ft - fr_1/c)\] - \exp\{j2\pi(ft - fr_2/c)\]]}{r_1} \\ &= \frac{Re[\alpha[\exp\{j2\pi(ft - fr_1/c)\][1 - \exp(j2\pi f(r_1 - r_2)/c)]}{r_1} \\ &\approx \frac{Re[\alpha[\exp\{j2\pi(ft - fr_1/c)\][1 - \exp(j2\pi f/c * b/r)]}{r_1} \\ &\approx \frac{Re[\alpha[\exp\{j2\pi(ft - fr_1/c)\][1 - (1 - j2\pi f/c * b/r)]}{r_1} \\ &= \frac{2\pi f|\alpha|b}{cr^2}\Re[j\exp(j\angle\alpha)\exp[j2\pi(ft - fr_1/c)]] \\ &= -\frac{2\pi f|\alpha|b}{cr^2}\sin[2\pi(ft - fr_1/c) + \angle\alpha]] \end{aligned}$$

Therefore $\beta = 2\pi f|\alpha|b/c$.

3.

$$\frac{1}{r_2} = \frac{1}{r_1 + (r_2 - r_1)} = \frac{1}{r_1[1 + (r_2 - r_1)/r_1]} \approx \frac{1}{r_1}\left(1 - \frac{r_2 - r_1}{r_1}\right) \approx \frac{1}{r_1}\left(1 - \frac{b}{r_1^2}\right)$$

Therefore if we don't make the approximation of b) we get another term in the expansion that decays as r^{-3}. This term is negligible for large enough r as compared to β/r^2.



EXERCISE 2.6. 1. Let f_2 be the probability density of the distance from the origin at which the photon is absorbed by exactly the 2nd obstacle that it hits. Let x be the location of the first obstacle, then

$$\begin{aligned} f_2(r) &= \mathbb{P}\{\text{photon absorbed by 2nd obstacle at } r\} \\ &= \int_x \mathbb{P}\{\text{absorbed by 2nd obstacle at } r \mid \text{not absorbed by 1st obstacle at } x\} \\ &\qquad \times \ \mathbb{P}\{\text{not absorbed by 1st obstacle at } x\} \quad dx \end{aligned}$$

Since the obstacle are distributed according to poisson process which has memoryless distances between consecutive points, the first term inside the integral is $f_1(r-x)$. The second term is the probability that the first obstacle is at x and the photon is not absorbed by it. Thus, it is given by $(1-\gamma)q(x)$. Thus,

$$f_2(r) = \int_{x=-\infty}^{\infty} (1-\gamma)q(x) f_1(r-x) dx$$

2. Similarly, we observe that $f_{k+1}(r)$ is given by

$$\begin{aligned} f_{k+1}(r) &= \int_x \mathbb{P}\{\text{absorbed by } (k+1)\text{th obst at } r \mid \text{not absorbed by 1st obst at } x\} \\ &\qquad \times \ \mathbb{P}\{\text{not absorbed by 1st obstacle at } x\} \quad dx \\ &= \int_{x=-\infty}^{\infty} (1-\gamma)q(x) f_k(r-x) dx \end{aligned} \tag{2.2}$$

3. Summing up (2.2) for $k=1$ to ∞, we get:

$$\sum_{k=2}^{\infty} f_k(r) = \int_{x=-\infty}^{\infty} (1-\gamma)q(x) \left(\sum_{k=1}^{\infty} f_k(r-x) \right) dx$$

Thus,

$$f(r) - f_1(r) = \int_{x=-\infty}^{\infty} (1-\gamma)q(x)f(r-x)dx,$$

or equivalently,

$$f(r) = \gamma q(r) + \int_{x=-\infty}^{\infty} (1-\gamma)q(x)f(r-x)dx \tag{2.3}$$

4. Using (2.3), we get that

$$F(\omega) = (1-\gamma)Q(\omega) + F(\omega)Q(\omega), \tag{2.4}$$

where F and Q denote the Fourier transform of f and q respectively. Since the $q(x)$ is known explicitly, its Fourier transform can be directly calculated and it turns out to be:

$$Q(\omega) = \frac{\eta^2}{\eta^2 + \omega^2}.$$

Substituting thin in (2.4), we get

$$F(\omega) = \frac{\gamma\eta^2}{\gamma\eta^2 + \omega^2}.$$

Thus, F is of the same form as Q, except for a different parameter η. Thus,

$$f(r) = \frac{\sqrt{\gamma}\eta}{2} e^{-\sqrt{\gamma}\eta|r|}$$

5. Without any loss of generality we can assume that r is positive, then power density at r is given by

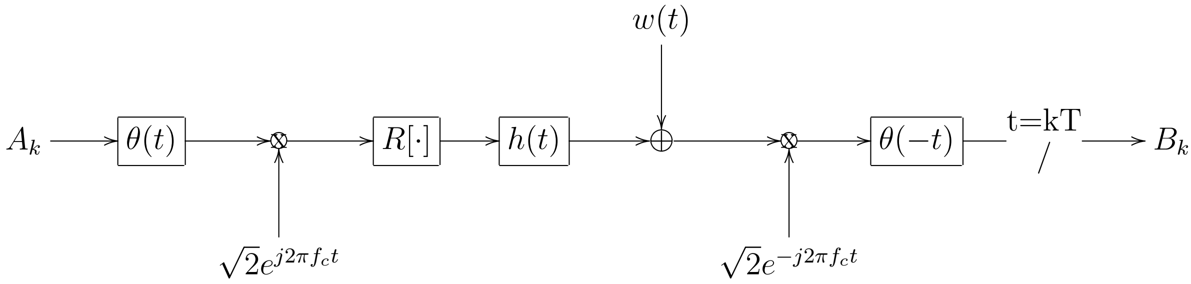
$$\begin{aligned}\int_{x=r}^{\infty} f(x)dx &= \int_{x=r}^{\infty} \frac{\sqrt{\gamma}\eta}{2} e^{-\sqrt{\gamma}\eta x} dx \\ &= \frac{1}{2} e^{-\sqrt{\gamma}\eta r}.\end{aligned}$$

A similar calculation for a negative r gives power density at distance r to be

$$\frac{e^{-\sqrt{\gamma}\eta|r|}}{2}.$$

Exercise 2.7.

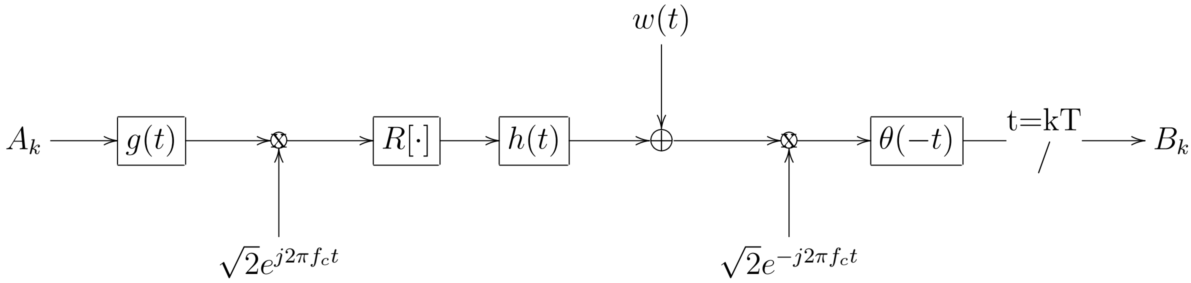
EXERCISE 2.8. The block diagram for the (unmodified) system is:



1. Which filter should one redesign?

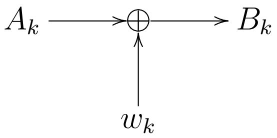
One should redesign the filter at the transmitter. Modifying the filter at the receiver may cause $\{\theta(t-kT)\}_k$ no longer to be an orthonormal set, resulting in noise on the samples not to be i.i.d. By leaving $\{\theta(t-kT)\}_k$ at the receiver as an orthonormal set, we are assured the the noise on the samples is i.i.d.

Let the modified filter be g(t). The block diagram for the modified system is:

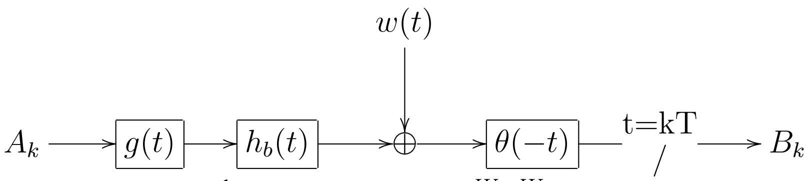


(Solution to Part 3: Figure of the various filters at passband).

We want to find g(t) such that there is no ISI between samples. Before we continue to find g(t), we depict the desired simplified block diagram for the system with no ISI:



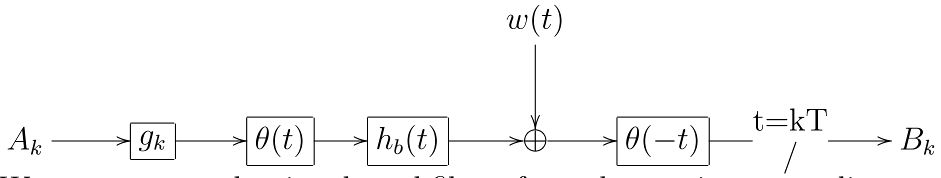
For ease of manipulation, we transform the passband representation of the system to a baseband representation



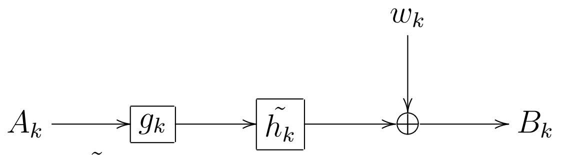
where $H_b(f) = \begin{cases} H(f+f_c) & \in [-\frac{W}{2}, \frac{W}{2}] \\ 0 & \text{otherwise} \end{cases}$

$H(f)$ is assumed bandlimited between $[f_c - \frac{W}{2}, f_c + \frac{W}{2}]$

We let $g(t) = \sum_k g_k \theta(t-kT)$, and redraw the block diagram:



We now convert the signals and filters from the continuous to discrete time domain:



where $\tilde{h}_k = \theta * h_b * \theta_-|_{t=kT}$.

We justify interchanging the order of w(t) and $\theta(-t)$, since we know the noise on the samples is i.i.d.

$G(z) = \tilde{H}^{-1}(z)$ gives the desired result.

In summary, $g(t) = \sum_k g_k \theta(t - kT)$ where g_k is given by $G(z) = \tilde{H}^{-1}(z)$, and $\tilde{H}(z)$ is given by the Z-Transform of $\tilde{h}_k = \theta * h_b * \theta_-|_{t=kT}$

Exercise 2.9. Part 1)

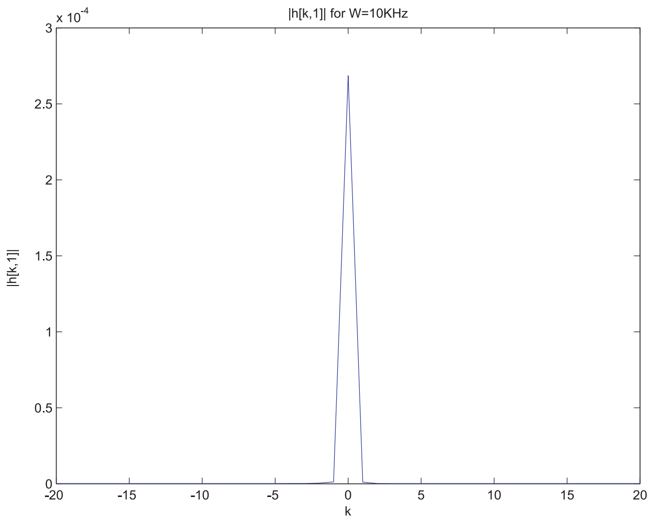


Figure 2.1: Magnitude of taps, W = 10kHz, time = 1 sec. Two paths are completely lumped together

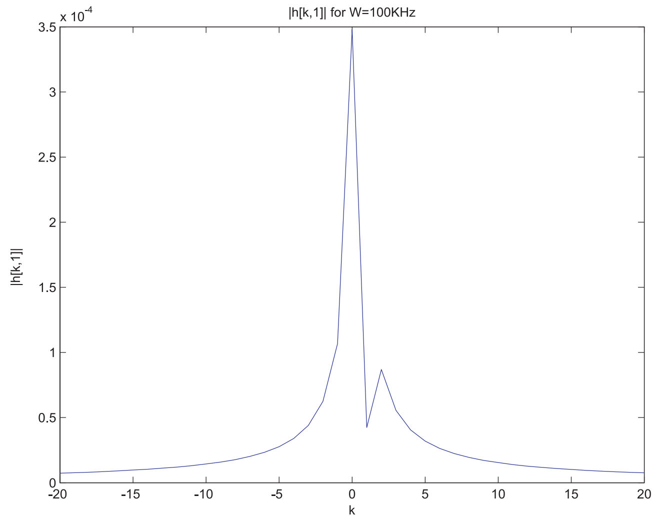


Figure 2.2: Magnitude of taps, W = 100kHz, time = 1 sec. Two paths are starting to become resolved.

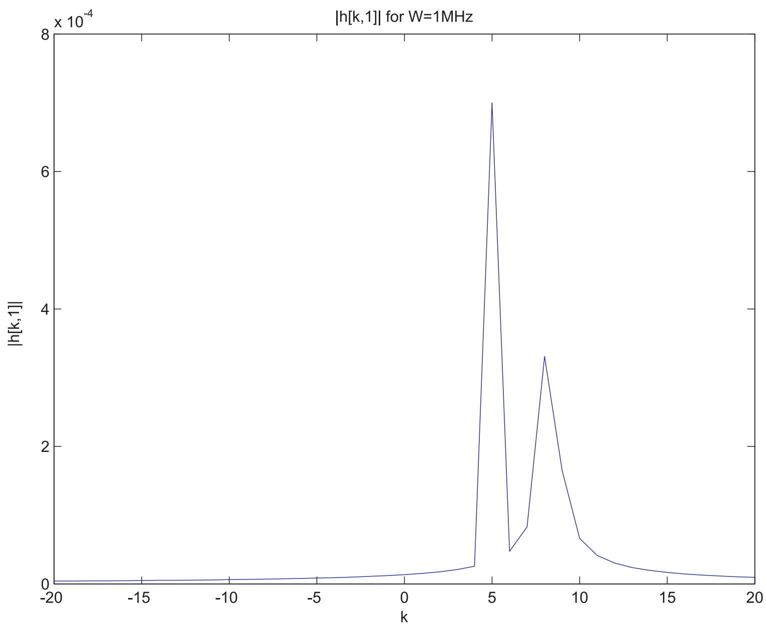


Figure 2.3: Magnitude of taps, W = 1MHz, time = 1 sec. Two paths are more resolved.

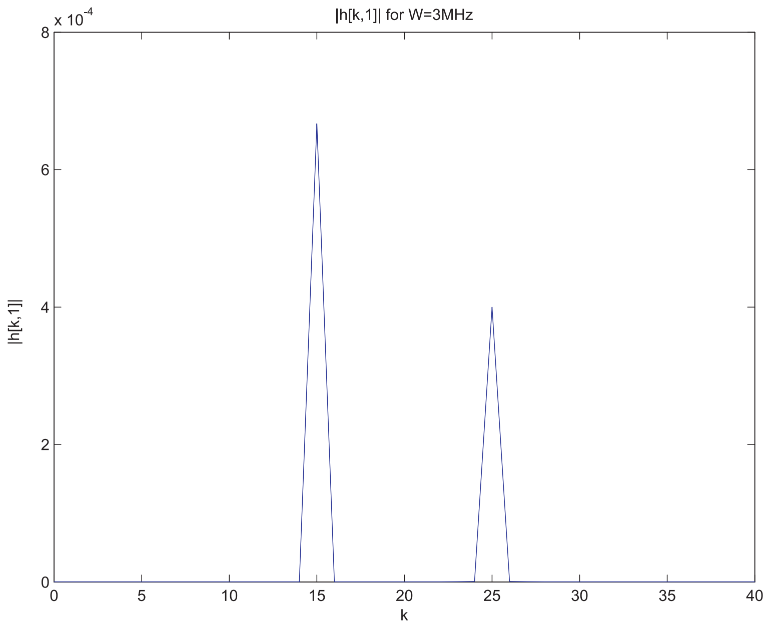


Figure 2.4: Magnitude of taps, W = 3MHz, time = 1 sec. Two paths are clearly resolved.

Part 2) We see that the time variations have the same frequency in both cases (flat fading in Figure 2.5 and frequency selective fading in Figure 2.7), but are much more pronounced in the case of flat fading. This is because in frequency selective fading (large W) each of the signal paths corresponds to a different tap, so they don't interfere significantly and the taps have small fluctuations. On the other hand in the case of flat fading, we sample the channel impulse response with low resolution and all the signal paths are lumped into the same tap. They interfere constructively and destructively generating large fluctuations in the tap values. If the model included more signal paths, then the number of paths contributing significantly to each tap would vary as a function of the bandwidth W, so the frequency of the tap variations would depend on the bandwidth, smaller bandwidth corresponding to larger Doppler spread and faster fluctuations (smaller T_c). Finally we could analyze this effect in the frequency domain. In frequency selective fading, the channel frequency response varies within the bandwidth of interest. There is an averaging effect and the resulting signal is never faded too much. This is an example of diversity over frequency.

Exercise 2.10. Consider the environment in Figure 2.9.

The shorter paths (dotted lines) contribute to the first tap and the longer paths (dashed) contribute to the second tap. Then the delay spread for the first tap is given by:

$$\frac{fv}{c}|\cos\phi_1 - \cos\phi_2|,$$

and the delay spread for the second tap is given by:

$$\frac{fv}{c}|\cos\theta_1 - \cos\theta_2|.$$

By appropriately choosing $\theta_1, \theta_2, \phi_1$ and ϕ_2, we can construct examples where the doppler spreads for both the taps are same or different.

Exercise 2.11. Let $H(f) = 1$ for $|f| < W/2$ and 0 otherwise. Then if $h(t) \leftrightarrow H(f)$ it follows that $h(t) = W\text{sinc}(Wt)$. Then we can write:

$$\begin{aligned}
\Re\{w[m]\} &= \left\{[w(t)\sqrt{2}\cos(2\pi f_c t)] * h(t)\ |_{t=m/W}\right\} \\
&= \left[\int_{-\infty}^{\infty} w(\tau)\sqrt{2}W\cos(2\pi f_c \tau)\text{sinc}(W(t-\tau))d\tau\right]_{t=m/W} \\
&= \int_{-\infty}^{\infty} w(\tau)\sqrt{2}W\cos(2\pi f_c \tau)\text{sinc}(m-W\tau)d\tau \\
&= \int_{-\infty}^{\infty} w(\tau)\psi_{m,1}(\tau)d\tau
\end{aligned}$$

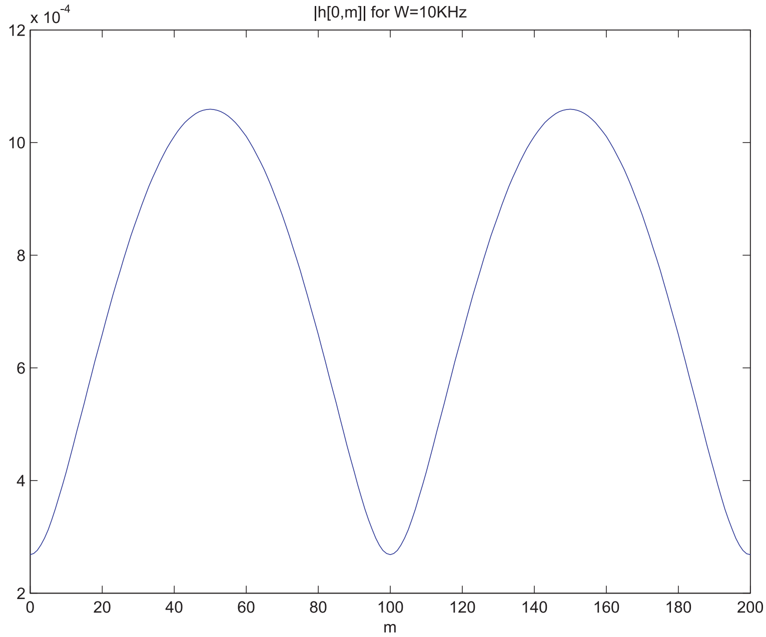


Figure 2.5: Flat fading: time variation of magnitude of 1 tap. (x-axis is the time index m).

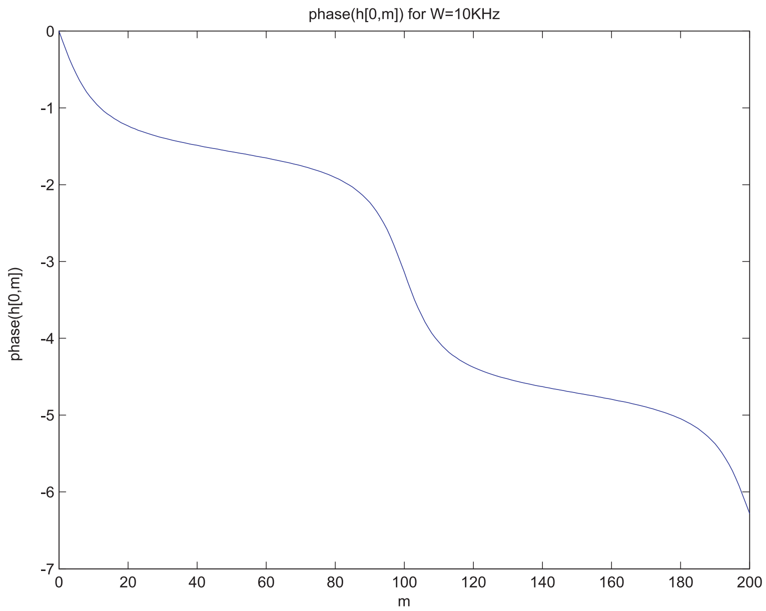


Figure 2.6: Flat fading: time variation of phase of 1 tap. (x-axis is the time index m).

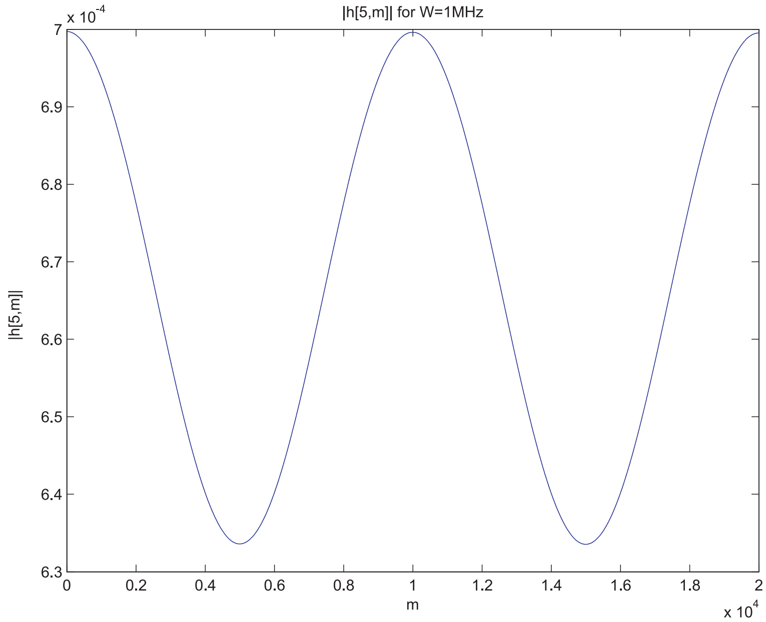


Figure 2.7: Frequency selective fading: time variation of magnitude of 1 tap. Note: scale of y-axis is much finer here than in the flat fading case. (x-axis displays time with units of seconds. x-axis label of time index 'm' is a typo. Should be 'time.')

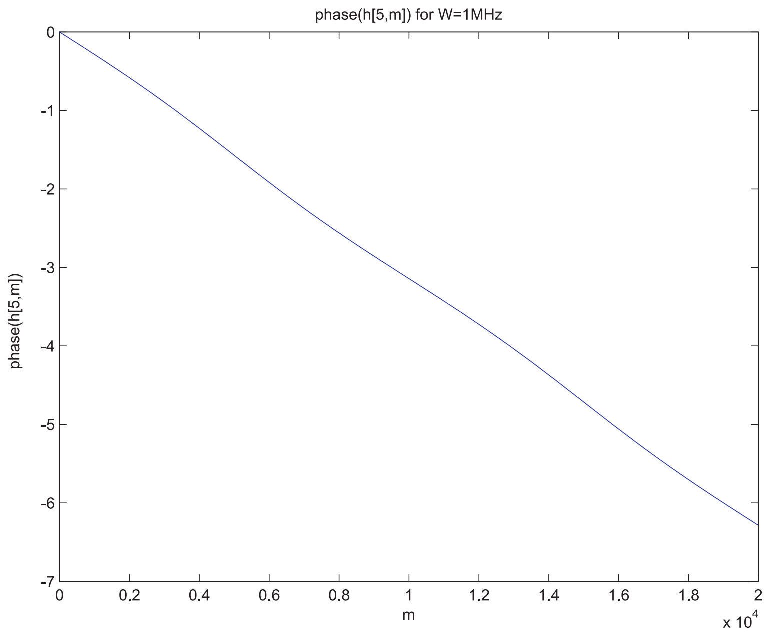


Figure 2.8: Frequency selective fading: time variation of phase of 1 tap. (x-axis displays time with units of seconds. x-axis label of time index 'm' is a typo. Should be 'time.')

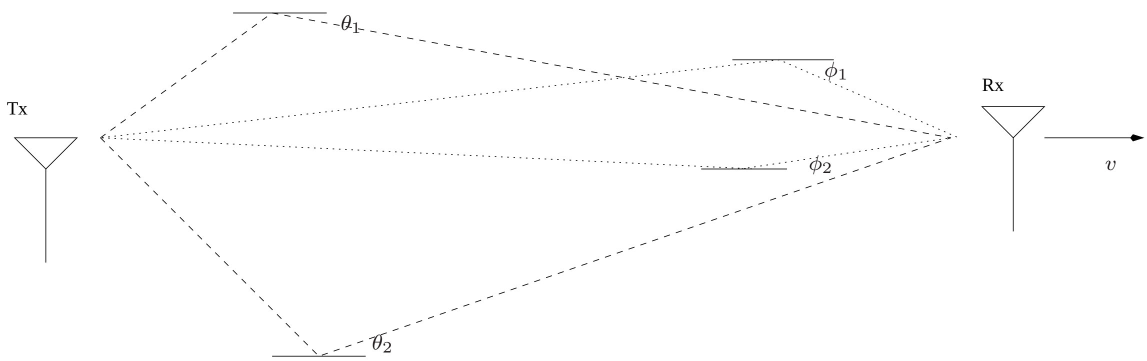


Figure 2.9: Location of reflectors, transmitter and receiver

where $\psi_{m,1}(\tau) = \sqrt{2}W cos(2\pi f_c\tau)\text{sinc}(m - W\tau)$.

Similarly,

$$\begin{aligned}
\Im\{w[m]\} &= -\left\{[w(t)\sqrt{2}\sin(2\pi f_c t)] * h(t) \mid_{t=m/W}\right\} \\
&= -\left[\int_{-\infty}^{\infty} w(\tau)\sqrt{2}W\sin(2\pi f_c\tau)\text{sinc}(W(t-\tau))d\tau\right]_{t=m/W} \\
&= -\int_{-\infty}^{\infty} w(\tau)\sqrt{2}W\sin(2\pi f_c\tau)\text{sinc}(m - W\tau)d\tau \\
&= \int_{-\infty}^{\infty} w(\tau)\psi_{m,2}(\tau)d\tau
\end{aligned}$$

where $\psi_{m,2}(\tau) = -\sqrt{2}W sin(2\pi f_c\tau)\text{sinc}(m - W\tau)$.

EXERCISE 2.12. 1) Let $\theta_n(t)$ denote $\theta(t - nT)$.

Show that if the waveforms $\{\theta_n(t)\}_n$ form an orthogonal set, then the waveforms $\{\psi_{n,1}, \psi_{n,2}\}_n$ also form an orthogonal set, provided $\theta(t)$ is band-limited to $[-f_c, f_c]$. $\psi_{n,1}, \psi_{n,2}$ are defined as

$$\begin{aligned}
\psi_{n,1}(t) &= \theta_n(t)\cos 2\pi f_c t \\
\psi_{n,2}(t) &= \theta_n(t)\sin 2\pi f_c t
\end{aligned} \tag{2.5}$$

By definition $\{\theta_n(t)\}_n$ forms an orthogonal set

$$\begin{aligned}
&\Longleftrightarrow \quad \int_{-\infty}^{\infty}\theta_n^*(t)\theta_m(t)dt = a\,\delta[m-n] \quad \text{for some } a \in \mathbb{R} \\
&\Longleftrightarrow \quad \int_{-\infty}^{\infty}\Theta_n^*(f)\Theta_m(f)df = a\,\delta[m-n] \quad \text{for some } a \in \mathbb{R}, \text{ by Parseval's Theorem} \qquad (2.6)
\end{aligned}$$

where $\Theta_n(f)$ is the Fourier Transform of $\theta_n(t)$.

We would like to show

1) $< \psi_{n,1}(t), \psi_{m,1}(t) > \quad \propto \quad \delta[m-n] \quad \forall\, m,n \in \mathbb{Z}$
waveforms modulated by $\cos 2\pi f_c t$ remain orthogonal to each other

2) $< \psi_{n,2}(t), \psi_{m,2}(t) > \quad \propto \quad \delta[m-n] \quad \forall\, m,n \in \mathbb{Z}$
waveforms modulated by $\sin 2\pi f_c t$ remain orthogonal to each other

3) $< \psi_{n,1}(t), \psi_{m,2}(t) > \quad = \quad 0 \quad \forall\, m,n \in \mathbb{Z}$
waveforms modulated by $\cos 2\pi f_c t$ are orthog. to waveforms modulated by $\sin 2\pi f_c t$.

We will show these three cases individually:

Case 1)

$$\begin{aligned} < \psi_{n,1}(t), \psi_{m,1}(t) > &= \int_{-\infty}^{\infty} \psi_{n,1}^*(t)\psi_{m,1}(t)dt \\ &= \int_{-\infty}^{\infty} \Psi_{n,1}^*(f)\Psi_{m,1}(f)df \text{ by Parseval's} \end{aligned} \tag{2.7}$$

where

$$\begin{aligned} \Psi_{n,1}(f) &= \int_{-\infty}^{\infty} \psi_{n,1}(t)e^{-j2\pi ft}dt \\ &= \int_{-\infty}^{\infty} \theta_n(t)\cos(2\pi f_c t)e^{-j2\pi ft}dt, \qquad \text{from (2.5)} \\ &= \Theta_n(f) * (\frac{1}{2}\delta(f-f_c) + \frac{1}{2}\delta(f+f_c)) \\ &= \frac{1}{2}(\Theta_n(f-f_c) + \Theta_n(f+f_c)) \end{aligned}$$

Substituting into (3)

$$\begin{aligned} &= \frac{1}{4}\int_{-\infty}^{\infty} [\Theta_n^*(f-f_c) + \Theta_n^*(f+f_c)][\Theta_m(f-f_c) + \Theta_m(f+f_c)]df \\ &= \frac{1}{4}\int_{-\infty}^{\infty} \Theta_n^*(f-f_c)\Theta_m(f-f_c) + \underbrace{\Theta_n^*(f-f_c)\Theta_m(f+f_c)}_{=0} + \\ &\quad + \underbrace{\Theta_n^*(f+f_c)\Theta_m(f-f_c)}_{=0} + \Theta_n^*(f+f_c)\Theta_m(f+f_c)df \end{aligned}$$

The second and third terms equal zero since $\theta(t)$ is bandlimited to $[-f_c, f_c]$ resulting in no overlap in the region of support of $\Theta(f+f_c)$ and $\Theta(f-f_c)$, as seen in Figure 2.10(b).

$$= \frac{1}{4}\int_{-\infty}^{\infty} \Theta_n^*(f-f_c)\Theta_m(f-f_c) + \Theta_n^*(f+f_c)\Theta_m(f+f_c)df$$

$$= \frac{1}{4}\int_{-\infty}^{\infty} \Theta_n^*(f)\Theta_m(f) + \Theta_n^*(f)\Theta_m(f)df$$

since integrals from $-\infty$ to ∞ are invariant to shifts of the integrand along the x-axis.

$$\begin{aligned} &= \frac{1}{4}2\int_{-\infty}^{\infty} \Theta_n^*(f)\Theta_m(f)df \\ &= \frac{a}{2}\delta[m-n], \text{ by equation (2.6)} \qquad (2.8) \\ &\propto \delta[m-n] \end{aligned}$$

Case 2)
$< \psi_{n,2}(t), \psi_{m,2}(t) >$

$$\begin{aligned} &= \int_{-\infty}^{\infty} \psi_{n,2}^*(t)\psi_{m,2}(t)dt \\ &= \int_{-\infty}^{\infty} \Psi_{n,2}^*(f)\Psi_{m,2}(f)df \text{ by Parseval's} \\ &= \int_{-\infty}^{\infty} (\frac{1}{2j})^*[\Theta_n^*(f-f_c) - \Theta_n^*(f+f_c)](\frac{1}{2j})[\Theta_m(f-f_c) - \Theta_m(f+f_c)]df \\ &= \frac{1}{4}\int_{-\infty}^{\infty} \Theta_n^*(f-f_c)\Theta_m(f-f_c) - \underbrace{\Theta_n^*(f-f_c)\Theta_m(f+f_c)}_{=0} + \\ &\quad - \underbrace{\Theta_n^*(f+f_c)\Theta_m(f-f_c)}_{=0} + \Theta_n^*(f+f_c)\Theta_m(f+f_c)df \\ &= \frac{1}{4}\int_{-\infty}^{\infty} \Theta_n^*(f-f_c)\Theta_m(f-f_c) + \Theta_n^*(f+f_c)\Theta_m(f+f_c)df \\ &= \frac{1}{4}\int_{-\infty}^{\infty} \Theta_n^*(f)\Theta_m(f) + \Theta_n^*(f)\Theta_m(f)df \\ &= \frac{1}{4}2\int_{-\infty}^{\infty} \Theta_n^*(f)\Theta_m(f)df \\ &= \frac{a}{2}\delta[m-n], \text{ by equation (2.6)} \qquad (2.9) \\ &\propto \delta[m-n] \end{aligned}$$

Case 3)
$< \psi_{n,1}(t), \psi_{m,2}(t) >$

$$\begin{aligned} &= \int_{-\infty}^{\infty} \psi_{n,1}^*(t)\psi_{m,2}(t)dt \\ &= \int_{-\infty}^{\infty} \Psi_{n,1}^*(f)\Psi_{m,2}(f)df \text{ by Parseval's} \end{aligned}$$

$$
\begin{aligned}
&= (\frac{1}{4j})\int_{-\infty}^{\infty}[\Theta_n^*(f-f_c)+\Theta_n^*(f+f_c)][\Theta_m(f-f_c)-\Theta_m(f+f_c)]df \\
&= (\frac{1}{4j})\int_{-\infty}^{\infty}\Theta_n^*(f-f_c)\Theta_m(f-f_c)-\underbrace{\Theta_n^*(f-f_c)\Theta_m(f+f_c)}_{=0}+ \\
&\quad +\underbrace{\Theta_n^*(f+f_c)\Theta_m(f-f_c)}_{=0}-\Theta_n^*(f+f_c)\Theta_m(f+f_c)df \\
&= (\frac{1}{4j})\int_{-\infty}^{\infty}\Theta_n^*(f-f_c)\Theta_m(f-f_c)-\Theta_n^*(f+f_c)\Theta_m(f+f_c)df \\
&= (\frac{1}{4j})\int_{-\infty}^{\infty}\Theta_n^*(f)\Theta_m(f)-\Theta_n^*(f)\Theta_m(f)df \\
&= 0 \quad \forall\, m,n\in\mathbb{Z}
\end{aligned}
$$

For $\psi(t)$ to be *orthonormal*, set $\frac{a}{2}=1$ in (2.8) and (2.9), which implies $a=2$. We should scale $\theta_n(t)$ by $\sqrt{2}$.

Part 2) $\tilde{\theta}(t)=4f_c\text{sinc}(4f_ct)$ is an example $\theta(t)$ that is *not* band-limited to $[-f_c,f_c]$. See Figure 2.10(c). For this example, there will be an overlap in the region of support of $\tilde{\Theta}(f+f_c)$ and $\tilde{\Theta}(f-f_c)$. See Figure 2.10(d). The cross terms $\tilde{\Theta}_n^*(f-f_c)\tilde{\Theta}_m(f+f_c)$ and $\tilde{\Theta}_n^*(f+f_c)\tilde{\Theta}_m(f-f_c)$ will no longer $=0$ and $\{\psi_{n,1},\psi_{n,2}\}_n$ will no longer by orthogonal.

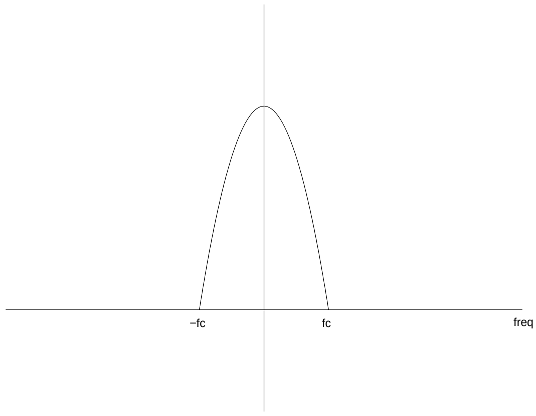
2 take away messages:
1) The orthogonality property of a set of waveforms is unchanged if the waveforms experience a frequency shift, or in other words are multiplied by $e^{j2\pi f_ct}$.
2) WGN projected onto $\{\psi_{n,1},\psi_{n,2}\}_n$ will yield i.i.d. gaussian noise samples.

EXERCISE 2.13. Let $\mathbf{F}[\cdot]$ denote the Fourier transform operator, $*$ denote convolution, $u(\cdot)$ the unit step function and

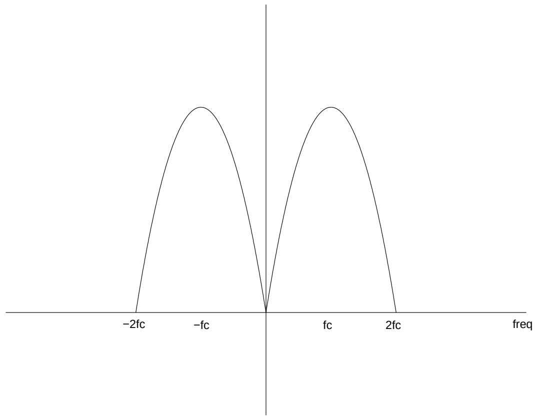
$$
H(f)=\begin{cases} 1/j & if \quad f>0 \\ 0 & if \quad f=0 \\ -1/j & if \quad f<0 \end{cases}
$$

with $h(t)\leftrightarrow H(f)$. Then we can write:

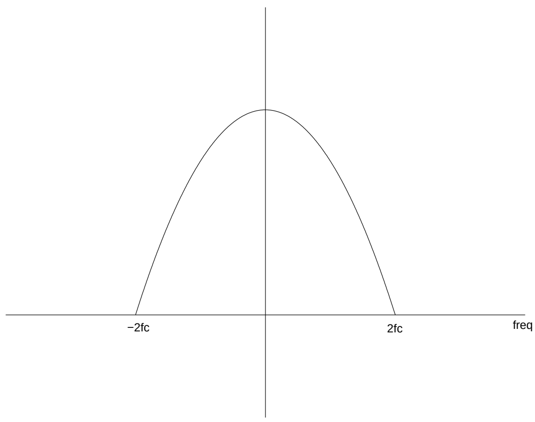
$$
\begin{aligned}
\Im[y_b(t)e^{j2\pi f_ct}] &= \frac{1}{2j}[y_b(t)e^{j2\pi f_ct}-(y_b(t)e^{j2\pi f_ct})^*]=\frac{1}{2j}\mathbf{F}^{-1}[Y_b(f-f_c)-Y_b^*(-f-f_c)] \\
&= \frac{\sqrt{2}}{2j}\mathbf{F}^{-1}[Y(f)u(f)-Y(f)u(-f)]=\frac{\sqrt{2}}{2}\mathbf{F}^{-1}[Y(f)H(f)]=\frac{\sqrt{2}}{2}y(t)*h(t) \\
&= \frac{\sqrt{2}}{2}\sum_i[a_i(t)x(t-\tau_i(t))]*h(t)
\end{aligned}
$$



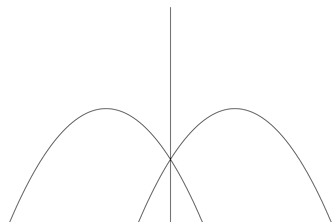
(a) Frequency range of $\Theta(f)$ band-limited from $-f_c, f_c$



(b) Frequency range of $\Theta(f+fc)$ and $\Theta(f-fc)$. Notice no overlap in region of support.



(c) Frequency range of $\tilde{\Theta}(f)$ *not* band-limited from $-f_c, f_c$



$$
\begin{aligned}
&= \frac{\sqrt{2}}{2}\sum_i \{a_i(t)\sqrt{2}\Re[x_b(t-\tau_i(t))e^{j2\pi f_c(t-\tau_i(t))}]\} * h(t) \\
&= \frac{1}{2}\sum_i \{a_i(t)[x_b(t-\tau_i(t))e^{j2\pi f_c(t-\tau_i(t))} + x_b^*(t-\tau_i(t))e^{-j2\pi f_c(t-\tau_i(t))}]\} * h(t) \\
&=_a \frac{1}{2j}\sum_i \{a_i(t)[x_b(t-\tau_i(t))e^{j2\pi f_c(t-\tau_i(t))} - x_b^*(t-\tau_i(t))e^{-j2\pi f_c(t-\tau_i(t))}]\} \\
&= \sum_i \{a_i(t)\Im[x_b(t-\tau_i(t))e^{j2\pi f_c(t-\tau_i(t))}]\} \\
&= \Im\left\{\left[\sum_i a_i(t)x_b(t-\tau_i(t))e^{-j2\pi f_c\tau_i(t)}\right]e^{j2\pi f_c t}\right\}
\end{aligned}
$$

The equality (a) follows because the first term between the braces is zero for negative frequencies and the second term is zero for positive frequencies.

Yes. Both equations together allow to equate the complex arguments of the $\Re$ and $\Im$ operators, thus allowing to obtain the baseband equivalent of the impulse response of the channel.

Exercise 2.14.

Exercise 2.15. Effects that make the tap gains vary with time:

- Doppler shifts and Doppler spread: $D = f_c\tau_i'(t)$, $T_c \sim 1/D = 1/(f_c\tau_i'(t))$ The coherence time is determined by the Doppler spread of the paths that contribute to a given tap. As W increases the paths are sampled at higher resolution and fewer paths contribute to each tap. Therefore the Doppler spread decreases for increasing W and its influence on the variation of the tap gains decreases.

- Variation of $\{a_i(t)\}_i$ with time. $a_i(t)$ changes slowly, with a time scale of variation much larger than the other effects discussed. However as W increases and it becomes comparable to f_c assuming that a single gain affects the corresponding path equally across all frequencies may not be a good approximation. The reflection coefficient of the scatterers may be frequency dependent and for very large bandwidths we need to change the model.

- Movement of paths from tap to tap. $\tau_i(t)$ changes with t and the corresponding path moves from one tap to another. As W increases fewer paths contribute to each tap and the tap gains change significantly when a path moves from tap to tap. A path moves from tap to tap when $\Delta\tau_i(t)W = 1$ or $\Delta\tau_i(t)/\Delta t \cdot W = 1/\Delta t$. So this effect takes place in a time scale of $\Delta t \sim 1/(W\tau_i'(t))$. As W increases this effect starts taking place in a small time scale and it becomes the dominant cause of time variation in the channel tap gains.

The third effect dominates when $\Delta t < T_c$ or equivalently when $W > f_c$.

EXERCISE 2.16.

$$h_\ell[m] = \sum_{i=1}^{N} a_i(m/W)e^{-j2\pi f_c \tau_i(m/W)}\text{sinc}(\ell - \tau_i(m/W)W)$$

Let $\bar{\tau} = \frac{1}{N}\sum_{i=1}^{N}\tau_i(0)$ and $\Delta\tau_i(m/W) = \tau_i(m/W) - \bar{\tau}$. Then,

$$h_\ell[m] = e^{-j2\pi f_c\bar{\tau}}\sum_{i=1}^{N} a_i(m/W)e^{-j2\pi f_c\Delta\tau_i(m/W)}\text{sinc}(\ell - \bar{\tau}W - \Delta\tau_i(m/W)W)$$

Often in practice $f_c\bar{\tau} \sim f_c r/c >> 1$ [1] so it is a reasonable assumption to model $e^{-j2\pi f_c\bar{\tau}} = e^{-j\theta}$ where $\theta \sim Uniform[0, 2\pi]$ and θ is independent of everything else. Note that $\bar{\tau}$ does not depend on m so a particular realization of θ is the same for all components of $\mathbf{h}$. Since $e^{-j\theta}$ has uniformly distributed phase, its distribution does not change if we introduce an arbitrary phase shift ϕ. So $e^{j\phi}e^{-j\theta} \sim e^{-j\theta}$.

It follows that

$$\begin{aligned}
e^{j\phi}\mathbf{h} &= e^{j\phi}e^{-j\theta}\begin{bmatrix} \sum_{i=1}^{N} a_i(m/W)e^{-j2\pi f_c\Delta\tau_i(m/W)}\text{sinc}(\ell - \bar{\tau}W - \Delta\tau_i(m/W)W) \\ \sum_{i=1}^{N} a_i((m+1)/W)e^{-j2\pi f_c\Delta\tau_i((m+1)/W)}\text{sinc}(\ell - \bar{\tau}W - \Delta\tau_i((m+1)/W)W) \\ \vdots \\ \sum_{i=1}^{N} a_i((m+n)/W)e^{-j2\pi f_c\Delta\tau_i((m+n)/W)}\text{sinc}(\ell - \bar{\tau}W - \Delta\tau_i((m+n)/W)W) \end{bmatrix} \\
&=_d e^{-j\theta}\begin{bmatrix} \sum_{i=1}^{N} a_i(m/W)e^{-j2\pi f_c\Delta\tau_i(m/W)}\text{sinc}(\ell - \bar{\tau}W - \Delta\tau_i(m/W)W) \\ \sum_{i=1}^{N} a_i((m+1)/W)e^{-j2\pi f_c\Delta\tau_i((m+1)/W)}\text{sinc}(\ell - \bar{\tau}W - \Delta\tau_i((m+1)/W)W) \\ \vdots \\ \sum_{i=1}^{N} a_i((m+n)/W)e^{-j2\pi f_c\Delta\tau_i((m+n)/W)}\text{sinc}(\ell - \bar{\tau}W - \Delta\tau_i((m+n)/W)W) \end{bmatrix} \\
&=_d \mathbf{h}
\end{aligned}$$

Since this is true for all ϕ, under the previous assumptions $\mathbf{h}$ is circularly symmetric.

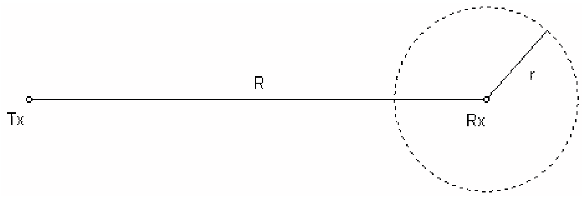
EXERCISE 2.17. 1. $h(\tau, t)$ is the response of the channel to an impulse that occurs at time $t - \tau$, i.e., $\delta(t - (t - \tau))$. Replacing $x(t)$ by $\delta(t - (t - \tau))$ in the given expression we obtain:

$$h(\tau, t) = \frac{a}{\sqrt{K}}\sum_{i=0}^{K-1}\delta(\tau - \tau_{\theta_i}(t)).$$

The projection of the velocity vector $\mathbf{v}$ onto the direction of the path at angle θ has a magnitude:

$$v_\theta = |\mathbf{v}|\cos\theta.$$

[1] r is the distance between transmit and receive antennas



The distance travelled by the mobile in the direction θ in time t is $v_\theta t$, which is the reduction in the distance between transmitter and receiver for the path of angle θ. Then,

$$\tau_\theta(t) = \tau_\theta(0) - \frac{|\mathbf{v}|\cos\theta \cdot t}{c}.$$

2. $T_d \ll 1/W$ means that most of the paths arrive in an interval much smaller than the sample time of the signal. Since the signal remains approximately constant over the interval T_d, it can be pulled out from the summation in part (a). In this way we can lump together the influence of all the paths into a single tap $h_0[m]$.

   We assume that $\delta_\theta(t) \ll 1/W$, For this we assume that $\delta_0(0) = 0$ and $\delta_0(t) \ll 1/W$ for the time scale of interest. Thus,

   $$h_0[m] = \int_0^{2\pi} a_\theta e^{-j2\pi f_c \tau_\theta(m/W)} \text{sinc}[-\tau_\theta(m/W) \cdot W] d\theta,$$

   where we use the fact that $a_\theta(t) = a_\theta, \forall t$. Finally we note that

   $$\lim_{t=0} \text{sinc}(t) = 1$$

   and since $\tau_\theta(m/W) \cdot W \ll 1$ we obtain:

   $$h_0[m] = \int_0^{2\pi} a_\theta e^{-j2\pi f_c \tau_\theta(m/W)} d\theta.$$

3. The independence assumption requires that different paths come from difference scatters. For this to be true even for small variations in angle of arrival θ, it is necessary that the scatters be located far away from the receiver. How far depends on the size of the scatters, and the angle difference $\Delta\theta$ over which we require the paths to be independent. The identically distributed assumption requires that the lengths of the paths from transmitter to receiver be comparable for all angle θ. This occurs when $r \ll R$ in the following figure.

4. The stationarity of $h_0[m]$ can be seen from previous formula and the uniformity of the phase. To calculate $R_0[n]$,

   $$R_0[n] = E[\int_0^{2\pi}\int_0^{2\pi} a^*_{\theta_1} a_{\theta_2} e^{j2\pi f_c(\tau_{\theta_1}(0) - \frac{vm+vn}{cW}\cos\theta_1 - \tau_{\theta_2}(0) + \frac{vm}{cW}\cos\theta_2)} d\theta_1 d\theta_2]$$

$$= \int_0^{2\pi}\int_0^{2\pi} E[a_{\theta_1}^* a_{\theta_2} e^{j2\pi f_c(\tau_{\theta_1}(0)-\tau_{\theta_2}(0))}] e^{j2\pi f_c(-\frac{vm+vn}{cW}\cos\theta_1 + \frac{vm}{cW}\cos\theta_2)} d\theta_1 d\theta_2$$
$$= \int_0^{2\pi} E[|a_\theta|^2] e^{-j2\pi f_c(\frac{vn}{cW}\cos\theta)} d\theta.$$

5.

$$\begin{aligned} R_0[n] &= \int_{-\infty}^{\infty} S(f) e^{j2\pi f n} df \\ &= \int_{-D_s/2W}^{D_s/2W} 4a^2 W/(D_s\sqrt{1-(2fW/D_s)^2}) e^{j2\pi f n} df \\ &= \int_0^{\pi} 2a^2 e^{j\pi n D_s \cos(\theta)/W} d\theta, \end{aligned}$$

where we use the substitution $\cos(\theta) = 2fW/D_s$.

6. From the definition of PSD.

EXERCISE 2.18. 1. The key difference is that in Clarke's model, the attenuations of the signals are random and have the same distribution in all directions from the receiver, whereas in the present model, they are deterministic and direction dependent. The key similarity is that in both cases, the phases are i.i.d. in all directions.

2. The delay spread T_d is $(4-2)$ km $/c = 6.7\mu$s. Therefore the channel is flat if $W \ll 1/T_d = 150$ kHz.

3. We assume that only paths that arrive with delay in $[\ell/W - 1/(2W), \ell/W + 1/(2W)]$ contribute to tap ℓ. Let θ be the angle the path makes with the line between the transmitter and the receiver at the receiver. The paths that arrive with the desired delay lie in between angles θ_1 and θ_2 and between $-\theta_1$ and $-\theta_2$, where θ_1, θ_2 are such that the delay of the path is $\ell/W \pm 1(2W)$, i.e.,

$$\begin{aligned} 1 + r(\theta_1) &= c(\ell/W - 1/(2W)) \\ 1 + r(\theta_2) &= c(\ell/W + 1/(2W)) \end{aligned}$$

where $r(\theta) = \sqrt{5 - 4\cos\theta}$ is the distance between the Tx and the scatterer at the angle θ.

The total power received is

$$\mathbb{E}[|h_\ell|^2] = 2\int_{\theta_1}^{\theta_2} G \frac{1}{r(\theta)^2} d\theta \qquad (2.10)$$

4. The received power in the delay range $[\tau, \tau + d\tau]$ is
$$2G/(5 - 4\cos\theta)|d\theta|$$
where $\cos\theta = [5 - (c\tau - 1)^2]/4$ and $\sin\theta d\theta = -c(c\tau - 1)/2d\tau$. Hence, the received power is
$$\frac{cG}{(c\tau - 1)\sqrt{1 - (\frac{5-(c\tau-1)^2}{4})^2}}$$
for $c\tau$ between 2 and 4. This gives the power-delay profile.

5. The Doppler shift at angle θ is $v\cos\theta/\lambda$, where $\lambda = c/f_c$. Thus the Doppler spread for the tap ℓ is
$$+f_c v/c|\cos\theta_1 - \cos\theta_2|$$
.

   The power of the received signal in the range of Doppler shifts $[f, f + df]$ is
$$2G/(5 - 4\cos\theta)|d\theta|$$
where $\cos\theta = \lambda/v \cdot f$ and $\sin\theta d\theta = \lambda/vdf$. Hence the Doppler spectrum is
$$S(f) = \frac{2G\lambda/v}{(5 - 4\lambda f/v)\sqrt{1 - (\lambda f/v)^2}}.$$
The Doppler spectrum of tap ℓ picks off a section of this corresponding to the range of Doppler shifts for the paths that contribute to this tap.

6. No, since the location of the scatterers should determine exactly the phase of the arriving path, so there cannot be any randomness of the phase once the location of the scatterers is fixed.

   On the other hand, the phase varies rapidly as a function of the scatterer positions at a spatial scale of the order of λ (cm's) while the large scale path loss and delay varies at the scale of kilometers. So what our assumptions are saying is that we are assuming that the scatters are "approximately" 1 km from the receiver, where the approximation is accurate up to the order of λ's . The randomness at the small scale validates the random phase assumption.

Exercise 2.19. 1.

$$\begin{aligned} f_m &= \frac{vf_c}{c} \\ D &= 2f_m = \frac{2vf_c}{c} \end{aligned}$$

$$T_c = \frac{1}{4D} = \frac{c}{8vf_c}$$

b) The antennas should be spaced at least by $d = vT_c = c/(8f_c)$ to get independenlty faded signals.

2. The signal arriving at the base station antenna from an angle α relative to the direction of v' experiences a Doppler shift

$$f_\alpha = \frac{v' f_c \cos(\alpha)}{c}$$

α ranges from $\pi - \arctan(R/d)$ to $\pi + \arctan(R/d)$. The Doppler shift is maximum for $\alpha = \pi$ and minimum for $\alpha = \pi - \arctan(R/d)$ or $\alpha = \pi + \arctan(R/d)$. Therefore the Doppler spread is

$$D = \frac{v' f_c[1 - \cos(\arctan(R/d))]}{c} = \frac{v' f_c 2[\sin(\arctan(R/d)/2)]^2}{c}$$

and the corresponding T_c:

$$T_c = \frac{1}{4D} = \frac{c}{v' f_c 8[\sin(\arctan(R/d)/2)]^2}$$

3. The minimum base station antenna spacing for uncorrelated fading is $d = v'T_c = c/(f_c 8[\sin(\arctan(R/d)/2)]^2)$. In practice the base station antenna is located in a high tower with no obstructions in its vicinity, so most of the scattering takes place around the mobile. In this case we can assume $R << d$ and approximate $\sin(\arctan(R/d)/2) \approx R/2d$ to get $d = (cd^2)/(2f_cR^2)$. In this particular setting this means that the minimum antenna spacing at the base station must be in the order of d^2/R^2 larger than that at the mobile to get independently faded signals.

Chapter 3

Solutions to Exercises

EXERCISE 3.1. We have

$$\begin{aligned}
P_e &= E_h[Q(\sqrt{2|h|^2\mathsf{SNR}})], && (3.1)\\
&= \int_0^\infty e^{-x}\int_{\sqrt{2x\mathsf{SNR}}}^\infty \frac{1}{\sqrt{2\pi}}e^{-t^2/2}dtdx, && (3.2)\\
&= \frac{1}{\sqrt{2\pi}}\int_0^\infty \int_0^{t^2/(2\mathsf{SNR})} e^{-t^2/2}e^{-x}dxdt, && (3.3)\\
&= \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-t^2/2}(1-e^{-t^2/(2\mathsf{SNR})})dt, && (3.4)\\
&= \frac{1}{2}-\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-t^2(1+1/\mathsf{SNR})/2}dt, && (3.5)\\
&= \frac{1}{2}\left(1-\sqrt{\frac{\mathsf{SNR}}{1+\mathsf{SNR}}}\right), && (3.6)
\end{aligned}$$

where the third step follows from changing the order of integration. Now, for large SNR, we also have

$$\sqrt{\frac{\mathsf{SNR}}{1+\mathsf{SNR}}} \approx 1-\frac{1}{2\mathsf{SNR}},$$

which implies

$$P_e \approx \frac{1}{4\mathsf{SNR}}$$

EXERCISE 3.2. 1. Let $\rho = \mathsf{SNR}$. For Rayleigh fading $|h[0]|^2 \sim Exp(1)$ so we have:

$$P_e = E\left[Q\left(\sqrt{2|h[0]|^2\rho}\right)\right] = \int_0^\infty \int_{\sqrt{2x\rho}}^\infty \frac{1}{\sqrt{2\pi}}e^{-t^2/2}e^{-x}dtdx$$

$$
\begin{aligned}
&= \int_0^\infty \int_0^{t^2/(2\rho)} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} e^{-x} dx dt = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \left[1 - e^{-t^2/(2\rho)}\right] dt \\
&= \frac{1}{2}\left[1 - \sqrt{\frac{\rho}{1+\rho}} \int_{-\infty}^\infty \sqrt{\frac{1+\rho}{2\pi\rho}} e^{-\frac{t^2}{2}(1+1/\rho)} dt\right] = \frac{1}{2}\left[1 - \sqrt{\frac{1}{1+1/\rho}}\right]
\end{aligned}
$$

We can approximate $\sqrt{1/(1+x)} = 1 - x/2 + o(x)$ for $x \to 0$ [1]. Then,

$$P_e = \frac{1}{2}\left[1 - 1 + \frac{1}{2}\frac{1}{\rho} + o(1/\rho)\right] = \frac{1}{4\rho} + o(1/\rho)$$

and

$$\lim_{\rho\to\infty} P_e \rho = \frac{1}{4}$$

2. We will need the following result:

$$\int_0^\infty Q(\sqrt{y})dy = \int_0^\infty \int_{\sqrt{y}}^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt dy = \int_0^\infty \int_0^{t^2} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dy dt = \int_0^\infty \frac{t^2}{\sqrt{2\pi}} e^{-t^2/2} dt = \frac{1}{2}$$

Let $f(\cdot)$ be the pdf of $|h[0]|^2$. Then,

$$P_e = E\left[Q\left(\sqrt{2|h[0]|^2\rho}\right)\right] = \int_0^\infty Q(\sqrt{2x\rho}) f(x) dx$$

Assuming that $f(\cdot)$ is right continuous at 0, that $f(0) > 0$ and that $f(\cdot)$ is bounded (this last condtion enables us to use the bounded convergence theorem to exchange limit and integral):

$$
\begin{aligned}
\lim_{\rho\to\infty} P_e\rho &= \lim_{\rho\to\infty} \int_0^\infty Q(\sqrt{2x\rho}) f(x) \rho dx = \lim_{\rho\to\infty} \int_0^\infty Q(\sqrt{y}) f\left(\frac{y}{2\rho}\right) \frac{1}{2} dy \\
&= \frac{f(0)}{2} \int_0^\infty Q(\sqrt{y}) dy = \frac{f(0)}{4}
\end{aligned}
$$

3. Let $g_\ell(\cdot)$ be the pdf of $|h[\ell]|^2$, and assume that it is right continuous and strictly positive at 0, for $\ell = 1, \ldots, L$. Let $f_\ell(\cdot)$ be the pdf of $\sum_{i=1}^{\ell} |h[i]|^2$. Then using the fact that the pdf of the sum of independent random variables equals the convolution of the corresponding pdfs we can write for $x \to 0$:

$$f_2(x) = \int_0^x g_1(t) g_2(x-t) dt = g_1(0) g_2(0) x + o(x)$$

[1] $\lim_{x\to 0} o(x)/x = 0$

$$
\begin{aligned}
f_3(x) &= \int_0^x f_2(t)g_3(x-t)dt = \int_0^x [g_1(0)g_2(0)t + o(t)]g_3(x-t)dt = g_1(0)g_2(0)g_3(0)\frac{x^2}{2} + o(x^2 \\
&\vdots \\
f_L(x) &= \int_0^x f_{L-1}(t)g_L(x-t)dt = \int_0^x \left[\left(\prod_{\ell=1}^{L-1} g_\ell(0)\right)\frac{t^{L-2}}{(L-2)!} + o(t^{L-2})\right] g_L(x-t)dt \\
&= \left(\prod_{\ell=1}^{L} g_\ell(0)\right)\frac{x^{L-1}}{(L-1)!} + o(x^{L-1}) = \beta x^{L-1} + o(x^{L-1})
\end{aligned}
$$

where we defined $\beta = [\prod_{\ell=1}^{L} g_\ell(0)]/(L-1)!$. The probability of error is given by:

$$
P_e = E\left[Q\left(\sqrt{2\|\mathbf{h}\|^2\rho}\right)\right] = \int_0^\infty Q(\sqrt{2x\rho})f_L(x)dx
$$

Multiplying by ρ^L, taking limit for $\rho \to \infty$ and assuming that we can exchange the order of limits and integrals we have:

$$
\begin{aligned}
\lim_{\rho\to\infty} P_e\rho^L &= \lim_{\rho\to\infty}\int_0^\infty Q(\sqrt{2x\rho})\rho^L f_L(x)dx = \lim_{\rho\to\infty}\int_0^\infty Q(\sqrt{y})\rho^{L-1} f_L\left(\frac{y}{2\rho}\right)\frac{1}{2}dy \\
&= \lim_{\rho\to\infty}\int_0^\infty Q(\sqrt{y})\frac{f_L\left(\frac{y}{2\rho}\right)}{\left(\frac{y}{2\rho}\right)^{L-1}}\left(\frac{y}{2}\right)^{L-1}\frac{1}{2}dy = \int_0^\infty Q(\sqrt{y})\beta\left(\frac{y}{2}\right)^{L-1}\frac{1}{2}dy \\
&= \frac{\beta}{2^L}\int_0^\infty\int_{\sqrt{y}}^\infty \frac{1}{\sqrt{2\pi}}e^{-t^2/2}y^{L-1}dtdy = \frac{\beta}{2^L}\int_0^\infty\int_0^{t^2}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}y^{L-1}dydt \\
&= \frac{\beta}{2^L}\int_0^\infty \frac{t^{2L}}{L}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt = \frac{\beta}{2^L}\frac{1}{2L}\int_{-\infty}^\infty t^{2L}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt \\
&= \frac{\beta}{2^L}\frac{1}{2L}\frac{(2L)!}{L!2^L} = \binom{2L-1}{L}\frac{1}{4^L}\prod_{\ell=1}^{L} g_\ell(0)
\end{aligned}
$$

4. The K parameter of a Ricean distribution is defined as the ratio of the powers in the specular (or constant) component and the fading component. If the specular component has amplitude μ and the fading component is $CN(0,\sigma^2)$ and we normalize the total power to be 1 we obtain:

$$
1 = \mu^2 + \sigma^2 = \sigma^2(K+1) \Rightarrow \ \mu = \sqrt{\tfrac{K}{K+1}} \quad \sigma^2 = \tfrac{1}{K+1}
$$

   For Ricean fading with parameter K the pdf of $|h[\ell]|^2$ is given by (see Proakis (2.1-140)):

$$
f(y) = (K+1)e^{-\left(\frac{K}{K+1}+y\right)(K+1)} I_0\left[\sqrt{yK(K+1)}\right], y \geq 0
$$

which evaluated at $y = 0$ yields $f(0) = (K+1)e^{-K}$.

Using the result from part (3) we get:

$$\lim_{\rho\to\infty} P_e \rho^L = \binom{2L-1}{L} \frac{(K+1)^L e^{-LK}}{4^L}$$

As $K \to \infty$ the expression above decays exponentially in K and the Ricean channel converges to the AWGN channel.

Exercise 3.3. Since $|h|^2$ is exponential, for large SNR, we have

$$\mathbb{P}(E_\epsilon) \;\approx\; \mathsf{SNR}^{-(1-\epsilon)}.$$

Therefore

$$\lim_{\mathsf{SNR}\to\infty} \frac{\log \mathbb{P}(E_\epsilon)}{\log \mathsf{SNR}} \;=\; -(1-\epsilon).$$

Similarly,

$$\lim_{\mathsf{SNR}\to\infty} \frac{\log \mathbb{P}(E_{-\epsilon})}{\log \mathsf{SNR}} \;=\; -(1+\epsilon).$$

1. By conditioning on E_ϵ, probability of error can be written as:

$$P_e \;=\; \mathbb{P}(\text{error}|E_\epsilon)\mathbb{P}(E_\epsilon) + \mathbb{P}(\text{error}|E_\epsilon^c)\mathbb{P}(E_\epsilon^c).$$

Now, the for the second term, we see that $|h|^2\mathsf{SNR} > \mathsf{SNR}^\epsilon$ whenever E_ϵ^c happens. Thus, because of the exponential tail of the Q function, the second term goes to zero exponentially fast and does not contribute to the limit. Also, upper bounding $\mathbb{P}(\text{error}|E_\epsilon)$ by 1, we get

$$\begin{aligned}\lim_{\mathsf{SNR}\to\infty} \frac{\log P_e}{\log \mathsf{SNR}} &\le \lim_{\mathsf{SNR}\to\infty} \frac{\log \mathbb{P}(E_\epsilon)}{\log \mathsf{SNR}}, \\ &= -(1-\epsilon).\end{aligned}$$

2. Now, similarly conditioning on $E_{-\epsilon}$, we get

$$\begin{aligned}P_e &= \mathbb{P}(\text{error}|E_{-\epsilon})\mathbb{P}(E_{-\epsilon}) + \mathbb{P}(\text{error}|E_{-\epsilon}^c)\mathbb{P}(E_{-\epsilon}^c), \\ &\ge \mathbb{P}(\text{error}|E_{-\epsilon})\mathbb{P}(E_{-\epsilon}).\end{aligned}$$

Now, we see that $|h|^2\mathsf{SNR} < \mathsf{SNR}^{-\epsilon}$ whenever $E_{-\epsilon}$ happens. Thus, the probability of error then can be lower bounded by a nonzero constant (e.g. $Q(1)$). Thus, we get

$$\begin{aligned}\lim_{\mathsf{SNR}\to\infty} \frac{\log P_e}{\log \mathsf{SNR}} &\ge \lim_{\mathsf{SNR}\to\infty} \frac{\log(Q(1)) + \log(\mathbb{P}(E_\epsilon))}{\log \mathsf{SNR}}, \\ &= -(1+\epsilon).\end{aligned}$$

3. Combining the last two part, we get

$$\lim_{\mathsf{SNR}\to\infty} \frac{\log P_e}{\log \mathsf{SNR}} = 1.$$

EXERCISE 3.4. To keep the same probability of error, the separation between consecutive points should be the same for both PAM and QAM. Let this separation be $2a$. Then, the average energy for a PAM with 2^k points is given by:

$$\begin{aligned} E_{av}(2^k - \text{PAM}) &= \frac{1}{2^{k-1}} \sum_{i=1}^{2^{k-1}} (2i-1)^2 a^2, \\ &= \frac{a^2}{3}(2^{2k} - 1). \end{aligned}$$

Since a 2^k QAM can be thought as two independent $2^{k/2}$ PAMs, we get that the average energy for a QAM with 2^k points is:

$$\begin{aligned} E_{av}(2^k - \text{QAM}) &= 2E_{av}(2^{k/2} - \text{PAM}), \\ &= \frac{2a^2}{3}(2^k - 1). \end{aligned}$$

Thus, the loss in energy is given by:

$$10 \log \left(\frac{2^k + 1}{2} \right),$$

which grows linearly in k.

EXERCISE 3.5. Consider the following scheme which works for both BPSK and QPSK. We have

$$\begin{aligned} y[m] &= hx[m-1]u[m] + w[m], \\ y[m-1] &= hx[m-1] + w[m-1]. \end{aligned}$$

Substituting the value of $hx[m-1]$ from the second equation into the first, we get

$$\begin{aligned} y[m] &= y[m-1]u[m] + w[m] - u[m]w[m-1], \\ y[m] &= \tilde{h}u[m] + \tilde{w}[m]. \end{aligned}$$

Where $\tilde{w}[m]$ is $\mathcal{CN}(0, 2N_0)$ and is independent of the input (because of symmetry of $w[m-1]$). The effective channel $\tilde{h}$ is $\mathcal{CN}(0, a^2 + N_0)$ and is *known* at the receiver. Thus, the performance is same as the performance on a coherent channel with signal to noise ratio given by $(\mathsf{SNR}+1)/2$. Thus, the probability of error for BPSK is given by

$$\frac{1}{2(1+\mathsf{SNR})},$$

and the probability of error for QPSK is given by:

$$\frac{1}{(1+\mathsf{SNR})}.$$

EXERCISE 3.6. 1. Let $\rho = \mathsf{SNR}$.

$$\begin{aligned}
P_e &= E\left[Q\left(\sqrt{2\|h[0]\|^2\rho}\right)\right] = \int_0^\infty \int_{\sqrt{2x\rho}}^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \frac{x^{L-1}}{(L-1)!} e^{-x} dt dx \\
&= \int_0^\infty \int_0^{t^2/(2\rho)} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \frac{x^{L-1}}{(L-1)!} e^{-x} dx dt \qquad (3.7) \\
&\stackrel{(a)}{=} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \left[1 - \sum_{k=0}^{L-1} \left(\frac{t^2}{2\rho}\right)^k \frac{e^{-t^2/(2\rho)}}{k!}\right] dt \\
&= \frac{1}{2}\left[1 - \sqrt{\frac{\rho}{1+\rho}} \sum_{k=0}^{L-1} \frac{1}{k!(2\rho)^k} \int_{-\infty}^\infty \sqrt{\frac{1+\rho}{2\pi\rho}} e^{-\frac{t^2}{2}(1+1/\rho)} dt\right] \\
&\stackrel{(b)}{=} \frac{1}{2}\left[1 - \sqrt{\frac{\rho}{1+\rho}} \sum_{k=0}^{L-1} \frac{1}{k!(2\rho)^k} \left(\frac{\rho}{1+\rho}\right)^k \frac{(2k)!}{k!2^k}\right] \qquad (3.8) \\
&= \frac{1}{2}\left[1 - \sum_{k=0}^{L-1} \binom{2k}{k} \mu \left(\frac{1-\mu}{2}\right)^k \left(\frac{1+\mu}{2}\right)^k\right]
\end{aligned}$$

where $\mu = \sqrt{\rho/(1+\rho)}$. Also (a) follows from the equivalence between the distribution functions of the Gamma and Poisson random variables, and (b) results from the formula for the even moments of the standard Gaussian distribution with proper scaling.

2. We start with the sufficient statistics:

$$\begin{aligned}
r_A^{(\ell)} &= h[\ell]x_1 + w_A^{(\ell)}, \quad \ell = 1, 2, \ldots, L \\
r_B^{(\ell)} &= h[\ell]x_2 + w_B^{(\ell)}, \quad \ell = 1, 2, \ldots, L \qquad (3.9)
\end{aligned}$$

where $x_1 = \|x_A\|$ and $x_2 = 0$ if $\mathbf{x}_A$ is transmitted, and $x_1 = 0$ and $x_2 = \|x_B\|$ if $\mathbf{x}_B$ is transmitted. As in the notes assume $\|x_A\|^2 = \|x_B\|^2 = \mathcal{E}_b$.

Since we are analyzing coherent reception, the receiver knows $\mathbf{h}$ and can further project onto $\mathbf{h}/\|\mathbf{h}\|$ obtaining the new sufficient statistics:

$$\begin{aligned} \tilde{r}_A &= \frac{\mathbf{h}}{\|\mathbf{h}\|}\mathbf{r}_A = \|h\|x_1 + \tilde{w}_A \\ \tilde{r}_B &= \frac{\mathbf{h}}{\|\mathbf{h}\|}\mathbf{r}_b = \|h\|x_2 + \tilde{w}_B \end{aligned} \tag{3.10}$$

where $\tilde{w}_A$ and $\tilde{w}_B$ are independent $CN(0, N_0)$ random variables. We can finally compare $\Re(\tilde{r}_A)$ and $\Re(\tilde{r}_B)$ to decide whether $\mathbf{x}_A$ or $\mathbf{x}_B$ was transmitted.

The error probability is given by:

$$\begin{aligned} P_e &= E\left[Pr\left(\Re\{\tilde{r}_A\} > \Re\{\tilde{r}_B\} \mid \mathbf{x}_B\right)\right] \\ &= E\left[Pr\left(\Re\{\tilde{w}_A - \tilde{w}_B\} > \|\mathbf{h}\|\sqrt{\mathcal{E}_b}\right)\right] = E\left[Q\left(\sqrt{2L\|\mathbf{h}\|^2\frac{\mathcal{E}_b}{2LN_0}}\right)\right] \end{aligned} \tag{3.11}$$

Since $\sqrt{L}\mathbf{h} \sim CN(0, \mathbf{I}_L)$ if follows that the formula of part (1) still holds by replacing $\rho \to \mathcal{E}_b/(2LN_0)$, which results in the desired answer.

Exercise 3.7. 1. We have (for simplicity we denote the squared-distances as d_1 and d_2)

$$\begin{aligned} \mathbb{P}[x_A \to x_B] &= E_{h_1,h_2}\left[Q\left(\sqrt{\frac{\mathsf{SNR}(|h_1|^2 d_1 + |h_2|^2 d_2)}{2}}\right)\right] \\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty \int_0^\infty \int_{\sqrt{\frac{\mathsf{SNR}(xd_1+yd_2)}{2}}}^\infty e^{-t^2/2}e^{-x}e^{-y}dtdxdy \\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty \int_0^{2t^2/(d_2\mathsf{SNR})} \int_0^{(2t^2/\mathsf{SNR}-d_2y)/d_1} e^{-x}dxe^{-y}dye^{-t^2/2}dt \\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty \int_0^{2t^2/(d_2\mathsf{SNR})} (1 - e^{-(2t^2/\mathsf{SNR}-d_2y)/d_1})e^{-y}dye^{-t^2/2}dt \\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty \left(1 - e^{-2t^2/(d_2\mathsf{SNR})} - \frac{e^{-2t^2/d_1\mathsf{SNR}}}{1-d_2/d_1}(1 - e^{-(1/d_2-/d_1)2t^2/\mathsf{SNR}})\right) e^{-t^2/2}dt, \\ &= 0.5 + \frac{1}{\sqrt{2\pi}(d_2-d_1)}\int_0^\infty e^{-t^2/2}(d_1e^{-2t^2/(d_1\mathsf{SNR})} - d_2e^{-2t^2/(d_2\mathsf{SNR})})dt, \\ &= 0.5 + \frac{0.5}{d_2-d_1}\left(\frac{d_1}{\sqrt{1+4/(d_1\mathsf{SNR})}} - \frac{d_2}{\sqrt{1+4/(d_2\mathsf{SNR})}}\right) \end{aligned}$$

2. For the high SNR scenario, we get (using Taylor series)

$$\frac{1}{\sqrt{1+4/(d_2\mathsf{SNR})}} = 1 - 2/(d_2\mathsf{SNR}) + 6/(\mathsf{SNR}^2 d_2^2).$$

Which implies

$$\mathbb{P}[x_A \to x_B] \quad = \quad 3/(d_1 d_2 \mathsf{SNR}^2).$$

EXERCISE 3.8. 1. For QPSK over an AWGN channel $P_{correct} = \left[1 - Q\left(\sqrt{\frac{2a^2}{N_0}}\right)\right]^2$ so

$$P_e = 1 - P_{correct} = 2Q\left(\sqrt{\frac{2a^2}{N_0}}\right) - \left[Q\left(\sqrt{\frac{2a^2}{N_0}}\right)\right]^2$$

To compare the performance to that of a scheme that uses only real symbols we fix the bit rate (2 bits per channel use) and total transmitted power. For 4-PAM with signal points located in $\{\pm b, \pm 3b\}$ the symbol error probability is (see Proakis (5.2-42)):

$$P_e = \frac{3}{2}Q\left(\sqrt{\frac{2b^2}{N_0}}\right)$$

The average transmitted power is $5b^2$ so normalizing the total transmitted power to 1 we have $a = 1/\sqrt{2}$, $b = 1/\sqrt{5}$. The power loss of 4-PAM over QPSK is $5/2 = 4dB$.

2. The conditional probability of error of QPSK conditioned on the channel realization $\mathbf{h}$ is:

$$P_{e|\mathbf{h}} = 2Q\left(\sqrt{\frac{2a^2\|\mathbf{h}\|^2}{N_0}}\right) - \left[Q\left(\sqrt{\frac{2a^2\|\mathbf{h}\|^2}{N_0}}\right)\right]^2$$

Averaging over the distribution of $\|\mathbf{h}\|^2$, noting that the second term of the above expression does not affect the high SNR error performance, upper bounding the Q function as done in the lectures, and using the characteristic function of the exponential distribution to evaluate the expectations we obtain:

$$\begin{aligned} P_e &= E\left[P_{e|\mathbf{h}}\right] \leq E\left[2Q\left(\sqrt{\frac{2a^2\|\mathbf{h}\|^2}{N_0}}\right)\right] \leq E\left[2e^{-\frac{\|\mathbf{h}\|^2 a^2}{N_0}}\right] = 2\prod_{\ell=1}^{L} E\left[e^{-\frac{|h[\ell]|^2 a^2}{N_0}}\right] \\ &= 2\left(\frac{1}{1+a^2/N_0}\right)^L = 2\left(1+\frac{\mathsf{SNR}}{2}\right)^{-L} \end{aligned}$$

where $\mathsf{SNR} = 2a^2/N_0$ is the signal to noise ratio.

For 4-PAM we can find the conditional error probability conditioned on $\mathbf{h}$ using the result of question 4.b) of homework 2 (which was derived for binary antipodal signaling but can easily be extended for 4-PAM) and scaling appropriately by b:

$$P_{e|\mathbf{h}} = \frac{3}{2}Q\left(\sqrt{\frac{2b^2\|\mathbf{h}\|^2}{N_0}}\right)$$

Taking expectation over $\mathbf{h}$ and bounding as before we get:

$$P_e \leq \frac{3}{2}\left(\frac{1}{1+b^2/N_0}\right)^L = \frac{3}{2}\left(1+\frac{\mathsf{SNR}}{5}\right)^{-L}$$

Therefore we get in both cases the same diversity gain, but the SNR performance is degraded by $4dB$ in the 4-PAM case.

3. Let $\mathcal{A} = \{a(1+j), a(1-j), a(-1+j), a(-1-j)\}$, $\mathcal{B} = \{x \in \mathcal{C}^2 : x_1 \in \mathcal{A}, x_2 \in \mathcal{A}\}$. Then the transmitted symbols are in $\mathcal{D} = \{\mathbf{U}x : x \in \mathcal{B}\}$, for some fixed unitary matrix $\mathbf{U} \in \mathcal{C}^{2x2}$. Let $\mathbf{x} = [x[1], x[2]]^T$ be one of such symbols. Then the received signal is:

$$y[\ell] = h[\ell]x[\ell] + z[\ell], \ell = 1, 2$$

from which we can extract the sufficient statistics:

$$r[\ell] = \frac{h[\ell]^*}{|h[\ell]|}y[\ell] = |h[\ell]|x[\ell] + \tilde{z}[\ell], \ell = 1, 2$$

where $\mathbf{z} =_d \tilde{\mathbf{z}}$. Let $v[\ell] = |h[\ell]|x[\ell]$ for $\ell = 1, 2$ be the modified symbol as seen at the receiver.

We are interested in the pairwise error probability, that is the probability of detecting a symbol $\mathbf{x}_2$ when the transmitted symbol is $\mathbf{x}_1$. The pairwise error probability conditioned on the channel realization $\mathbf{h}$ depends only on $\|\mathbf{v}_1 - \mathbf{v}_2\|^2$ through the expression:

$$P_{(\mathbf{x}_1 \to \mathbf{x}_2)|\mathbf{h}} = Q\left(\sqrt{\frac{\|\mathbf{v}_2 - \mathbf{v}_1\|^2}{2N_0}}\right)$$

which can be rewritten and upper bounded as:

$$P_{(\mathbf{x}_1 \to \mathbf{x}_2)|\mathbf{h}} = Q\left[\sqrt{\frac{|h[1]|^2|x_2[1] - x_1[1]|^2 + |h[2]|^2|x_2[2] - x_1[2]|^2}{2N_0}}\right]$$

$$\leq \exp\left[-\frac{|h[1]|^2|x_2[1]-x_1[1]|^2+|h[2]|^2|x_2[2]-x_1[2]|^2}{4N_0}\right]$$

Taking expectation over $\mathbf{h}$ and using the characteristic function of the exponential distribution to evaluate it we obtain:

$$\begin{aligned}
P_{(\mathbf{x}_1\to\mathbf{x}_2)} &\leq E\left\{\exp\left[-\frac{|h[1]|^2|x_2[1]-x_1[1]|^2+|h[2]|^2|x_2[2]-x_1[2]|^2}{4N_0}\right]\right\}\\
&= E\left\{\exp\left[-\frac{|h[1]|^2|x_2[1]-x_1[1]|^2}{4N_0}\right]\right\}E\left\{\exp\left[-\frac{|h[2]|^2|x_2[2]-x_1[2]|^2}{4N_0}\right]\right\}\\
&= \left[\frac{4N_0}{1+|x_2[1]-x_1[1]|^2}\right]\left[\frac{4N_0}{1+|x_2[2]-x_1[2]|^2}\right]\\
&\leq \frac{16N_0}{|x_2[1]-x_1[1]|^2|x_2[2]-x_1[2]|^2}
\end{aligned}$$

We see that the pairwise error probability depends only on the product distance $|x_2[1]-x_1[1]|^2|x_2[2]-x_1[2]|^2$. The rotation matrix $\mathbf{U}$ should be chosen in such a way as to maximize the minimum product distance between any pair of codewords in the constellation $\mathcal{D}$.

EXERCISE 3.9.

EXERCISE 3.10. 1. The code can be represented by a permutation of the 16-point QAM. What is transmitted on the second subchannel can be obtained as a simple permutation of what is transmitted on the first sub-channel.

2. Data rate = 2 bits/channel use (since all the information is contained in one of the QAMs itself).

3. Since the product distance is non-zero, the diversity gain is 2. The minimum product distance is given by by $64a^4$ where $2a$ is the minimum distance between the QAM symbols. Then, by normalizing the average receiver SNR to be 1 per time symbol, we get:

$$\begin{aligned}
2\times 2\times(4\times 20a^2)/16 &= 1,\\
a^2 &= 0.05.
\end{aligned}$$

Therefore, the product distance is given by 0.32.

4. The power used for the rotation code is $4a^2$ per time and that of the permutation code is $20a^2$, but for a fair comparison of power used, we need to normalize by the optimal product distance of the rotation code. Numerical results indicate that the rotation code outperforms a permutation code by a factor of 1.05.

EXERCISE 3.11. 1.

$$\mathbb{P}\{\mathbf{x}_A \to \mathbf{x}_B\} \leq [\frac{1}{1+\mathsf{SNR}}]^d$$

where d is the Hamming distance between the binary codewords $\mathbf{x}_A$ and $\mathbf{x}_B$. The diversity gain of the code is the minimum Hamming distance d_{min} between the codewords .

2. It's 2. Same diversity gain as the repetition code but higher rate 3/2 rather than 1/2.

3. The probability a symbol gets decoded incorrectly is of the order of SNR^{-1}. If $\lceil d_{min}/2 \rceil$ errors are made, then there is a significant probability (i.e., the probability does not decay with SNR) that an overall error is made, as an incorrect codeword may be closer to $\hat{\mathbf{c}}$ than the transmitted codeword. We are ok if fewer than that is made. Hence, the diversity gain is $\lceil d_{min}/2 \rceil$.

   For the example, the diversity is only 1.

4. The typical error event for each symbol is when the channel is in a deep fade. If we declare an erasure whenever the channel is in a deep fade, then the typical error is that there are only erasures and no hard decision errors. We can decode whenever the number of erasures is less than d_{min}, since there is at most one codeword that is consistent with the erasure pattern. Hence the diversity gain is back to d_{min}, same as soft decision decoding.

   How do we know the channel is in deep fade? Heuristically, when $|h_\ell|^2 < 1/\mathsf{SNR}$. More rigorously, we can fix a $\epsilon > 0$ and use the threshold $|h_\ell|^2 < 1/\mathsf{SNR}^{1-\epsilon}$ to decide on an erasure. This will give us a diversity gain of $d_{min}(1-\epsilon)$. But we can choose ϵ arbitrarily close to zero so we can get close to the desired diversity gain.

EXERCISE 3.12. 1. For repetition coding, probability of error is given by

$$\mathbb{P}\left\{|\mathbf{h}|^2 < 1/\mathsf{SNR}\right\}.$$

Now, let the singular value of decomposition of $\mathbf{K_h}$ be

$$\mathbf{K_h} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*.$$

Now, define: $\tilde{\mathbf{h}} = \mathbf{U}^*\mathbf{h}$. Then $|\tilde{\mathbf{h}}| = |\mathbf{h}|$ and $\tilde{\mathbf{h}}$ has i.i.d. complex Gaussian entries with variance given by the singular values of $\mathbf{K_h}$. Thus, the diversity order is given by the number non-zero singular values of $\mathbf{K_h}$.

2. We can write the time diversity channel as:

$$\begin{aligned}\mathbf{y} &= [h_1, h_2, \cdots, h_L] \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & x_L \end{bmatrix} + \mathbf{w}, \\ &= [h_1, h_2, \cdots, h_L]\mathbf{K_h}^{-1/2}\mathbf{K_h}^{1/2} \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & x_L \end{bmatrix} + \mathbf{w},\end{aligned}$$

   Now, $[h_1, h_2, \cdots, h_L]\mathbf{K_h}^{-1/2}$ is i.i.d. Gaussian and is similar to a standard MISO channel. Thus the code design criterion is given by the determinant of $\mathbf{K_h}^{0.5}(\mathbf{X}_A - \mathbf{X}_B)(\mathbf{X}_A - \mathbf{X}_B)^*\mathbf{K_h}^{0.5^*}$. As $\mathbf{K_h}$ is a fixed matrix and only contributes a constant we can eliminate it from the code design criterion. Now, in our case $\mathbf{X}_A$ and $\mathbf{X}_B$ are restricted to be diagonal which implies that the difference determinant is the the product distance. Thus, the criterion remains unchanged. Note, that here we have assumed $\mathbf{K_h}$ is full rank. Otherwise, one will have to work with the corresponding reduced problem by ignoring some of the h_is.

3. As seen in the previous part, the criterion remains unchanged.

Exercise 3.13. The channel equation is $\mathbf{y} = \mathbf{h}x + \mathbf{w}$. Let $l = \arg\max_i |h[i]|$. The selection combiner bases its decision on the l^{th} branch only discarding the rest so the decision is based on $y[\ell] = h[\ell]x + w[\ell]$.

Let $\{x_i\}_{i=1}^L$ be i.i.d. $Exp(1)$ random variables, and $x = \max_i x_i$. Then the pdf of x, $f(\cdot)$, for $x \to 0$ is given by:

$$\begin{aligned}f(x) &= L(1-e^{-x})^{L-1}e^{-x} = L[1-(1-x+o(x))]^{L-1}[1-x+o(x)] \\ &= Lx^{L-1} + o(x^{L-1})\end{aligned}$$

Noting that $|h[\ell]|^2$ has the above pdf, we can use the derivation of Ex. 3.2, part 3) replacing β with L:

$$\lim_{\rho\to\infty} P_e\rho^L = \frac{L}{4^L}\frac{(2L-1)!}{L!} = \frac{(2L-1)!}{4^L(L-1)!}$$

We observe that this scheme still achieves a diversity gain of L but the error performance degrades by the factor $L/\beta = L!/(\prod_{i=1}^L g_i[0]) = L!$ with respect to that of optimal combining.

Exercise 3.14.

EXERCISE 3.15. 1. We obtain the same diversity gain over the MISO channel (assuming the same statistical characterization of the fading gains) as we have operationally converted the MISO channel into a parallel channel. Also, since the determinant of a diagonal matrix is the product of its diagonal elements, the determinant metric for the MISO channel is same as the product distance metric of the time diversity code.

2. For the rotation code the worst case pairwise error probability is given by

$$\frac{16}{\min \delta_{ij}} \mathsf{SNR}^{-2},$$

where the optimal δ_{ij} is given by 16/5, which gives $5/\mathsf{SNR}^2$ as the upper bound for the rotation code.

On the other hand, for Alamouti scheme, the worst case probability of error is given by:

$$\frac{16}{\mathsf{SNR}^2 \det(\mathbf{X}_A - \mathbf{X}_B)^2}.$$

If u_1 and u_2 are the BPSK $(+/-a)$ symbols used for the Alamouti scheme, then the average power per time symbol is given by $2a^2$. The determinant is given by $u_1^2 + u_2^2$, thus the worst case determinant is given by $4a^2$. Thus, after normalizing we get worst case probability of error is given by $4/\mathsf{SNR}^2$. Thus, the Alamouti scheme is better.

For QPSK symbols, we are just using an additional degree of freedom which will change the power used for both the schemes, but the relative difference in the worst case error performance (a factor of 1.25) remains the same.

3. See Figure 3.1.

EXERCISE 3.16. 1. We have

$$\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{d} + \mathbf{w}, \\ &= \sum_i \mathbf{a}_i d_i + \mathbf{w}. \end{aligned}$$

Since $\mathbf{A}$ is orthogonal, all the $\mathbf{a}_i$s are orthogonal, thus for detecting d_i, we can project along $\mathbf{a}_i$:

$$\mathbf{a}_i^* \mathbf{y} = ||\mathbf{a}_i||^2 d_i + \mathbf{a}_i^* \mathbf{w}. \tag{3.12}$$

Since $\mathbf{A}$ is orthogonal, the noise $\mathbf{a}_i^* \mathbf{w}$ is independent of other noise terms (and hence the other projections).Thus, each of the d_is can be decoded separately.

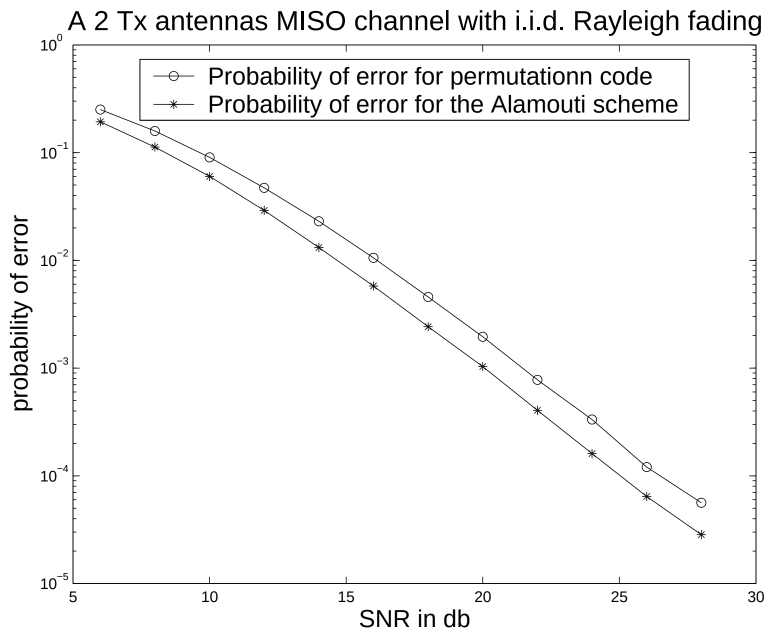


Figure 3.1: The error probability of uncoded QAM with the Alamouti scheme and that of a permutation code over one antenna at a time for the Rayleigh fading MISO channel with two transmit antennas: the permutation code is only about 1.5 dB worse than the Alamouti scheme over the plotted error probability range.

2. If $||\mathbf{a}_m|| = ||\mathbf{h}||$, then we can normalize the equation (3.12) and get a fading coefficient of $||\mathbf{h}||$ which implies a full diversity gain for each symbol.

3. We have

$$\mathbf{h}^*\mathbf{X} = \mathbf{d}^t\mathbf{A}^t,$$

which along with orthogonality and the full diversity property of $\mathbf{A}$ implies that

$$\begin{aligned} \mathbf{h}^*\mathbf{X}\mathbf{X}^*\mathbf{h} &= ||\mathbf{d}||^2||\mathbf{h}||^2\mathbf{I}_L, \\ \mathbf{h}^*(\mathbf{X}\mathbf{X}^* - ||\mathbf{d}||^2\mathbf{I}_L)\mathbf{h} &= 0, \end{aligned}$$

for every $\mathbf{h}$. Thus, $\mathbf{X}\mathbf{X}^*$ must be $||\mathbf{d}||^2\mathbf{I}_L$.

Exercise 3.17.

Exercise 3.18. First, lets calculate the deep fade probability for Ricean fading which we will use for calculating both diversity order and product distance criterion. Let

$$h = Ae^{i\theta} + n,$$

where A is a fixed number and θ is uniform in $[0, 2\pi]$ and n is $\mathcal{CN}(0,1)$. Since we are interested only in the $|h|^2$ whose distribution does not depend on θ (because n is circularly symmetric), without any loss of generality, we can take $\theta = 0$. Then we can write h as

$$h = A + n_R + jn_I,$$

where $n = n_R + jn_I$ and n_R and n_I are i.i.d. Gaussian. Then

$$\begin{aligned}\mathbb{P}\left\{|h|^2 < 1/\mathsf{SNR}\right\} &= \mathbb{P}\left\{(A+n_R)^2 + (n_I)^2 < 1/\mathsf{SNR}\right\}, \\ &\approx \mathbb{P}\left\{(A+n_R)^2 < 1/\mathsf{SNR} \text{ and } (n_I)^2 < 1/\mathsf{SNR}\right\}, \\ &\approx \frac{e^{-A^2/4}}{\sqrt{\mathsf{SNR}}}\frac{1}{\sqrt{\mathsf{SNR}}}, \\ &\approx \frac{e^{-A^2/4}}{\mathsf{SNR}}.\end{aligned}$$

1. Now, the deep fade event for a time diversity channel is when all the sub-channels are in deep fade. The probability of which is given by

$$\frac{e^{-LA^2/4}}{\mathsf{SNR}^L}.$$

   Thus, the diversity order does not change for Ricean fading.

2. For a pairwise code design criterion, we want to calculate

$$\begin{aligned}\mathbb{P}\left\{\sum_i |h_i|^2 d_i^2 < 1/\mathsf{SNR}\right\} &\approx \prod_i \mathbb{P}\left\{|h_i|^2 d_i^2 < 1/\mathsf{SNR}\right\}, \\ &\approx \frac{e^{-LA^2/4}}{\mathsf{SNR}^L d_1^2 d_2^2 \cdots d_L^2}.\end{aligned}$$

   Thus, we get the same product distance criterion for Ricean fading as well.

Exercise 3.19. 1. Let $\mathbf{H}$ be the fading matrix for the MIMO channel. Then the channel model can be written as:

$$\mathbf{Y} = \mathbf{HX} + \mathbf{W}.$$

Now, this channel model can be rewritten as a MISO channel with block-length $n_r N$. Let $\tilde{\mathbf{X}}$ and $\mathbf{h}$ be

$$\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{X} \end{bmatrix},$$

$$\mathbf{h} = \left[\begin{array}{cccccc} H(1,1) & \cdots & H(1,L) & H(2,1) & \cdots & \cdots & H(n_r,L) \end{array} \right].$$

Then the received signal can be rewritten as

$$\mathbf{y} = \mathbf{h}\tilde{\mathbf{X}} + \mathbf{w},$$

with $\mathbf{y}$ and $\mathbf{w}$ appropriately defined in term of $\mathbf{Y}$ and $\mathbf{W}$. Then the probability of pairwise error can be written as:

$$E\left[Q\left(\sqrt{\frac{\mathsf{SNR}\mathbf{h}(\tilde{\mathbf{X}}_A - \tilde{\mathbf{X}}_B)(\tilde{\mathbf{X}}_A - \tilde{\mathbf{X}}_B)^*\mathbf{h}^*}{2}}\right)\right]$$

2. Since we have reduced the MIMO problem with i.i.d. Rayleigh fading to a MISO problem with i.i.d. Rayleigh fading, probability of pairwise error can be upper bounded as:

$$\begin{aligned} \mathbb{P}(\mathbf{X}_A \to \mathbf{X}_B) &\leq \frac{4^{Ln_r}}{\mathsf{SNR}^{Ln_r} \det\left((\tilde{\mathbf{X}}_A - \tilde{\mathbf{X}}_B)(\tilde{\mathbf{X}}_A - \tilde{\mathbf{X}}_B)^*\right)}, \\ &= \left(\frac{4^L}{\mathsf{SNR}^L \det\left((\mathbf{X}_A - \mathbf{X}_B)(\mathbf{X}_A - \mathbf{X}_B)^*\right)}\right)^{n_r}, \end{aligned}$$

where the last step follows from the diagonal structure of $\tilde{\mathbf{X}}_A$ and $\tilde{\mathbf{X}}_B$.

3. Thus, the code design criterion of maximizing the minimum determinant remains unchanged.

Exercise 3.20. Using the same notation as in the text, and using a subindex to denote the receive antenna (either 1 or 2) we have:

$$\begin{bmatrix} y_1[1] \\ y_1[2]^* \\ y_2[1] \\ y_2[2]^* \end{bmatrix} = \begin{bmatrix} h_{11} & h_{21} \\ h_{21}^* & -h_{11}^* \\ h_{12} & h_{22} \\ h_{22}^* & -h_{12}^* \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} w_1[1] \\ w_1[2]^* \\ w_2[1] \\ w_2[2]^* \end{bmatrix} \tag{3.13}$$

where h_{ij} is the complex channel gain from transmit antenna i to receive antenna j. In vector notation we can rewrite (3.13) as $\mathbf{y} = \mathbf{H}\mathbf{u} + \mathbf{w}$, where $\mathbf{w} \sim CN(0, N_0 I_4)$. Noting that the columns of $\mathbf{H}$ (which we call $\mathbf{h}_i$) are orthogonal, we can project the (slightly modified) received vector $\mathbf{y}$ onto the normalized columns of $\mathbf{H}$ to obtain the 2 sufficient statistics:

$$r_i = \frac{\mathbf{h}_i^*}{\|\mathbf{h}_i\|}\mathbf{y} = \|h\| u_i + \tilde{w}_i \tag{3.14}$$

for $i = 1, 2$, where $\tilde{w}_i \sim CN(0, N_0)$ independent across i, and $\|\mathbf{h}\| = \|\mathbf{h}_1\| = \|\mathbf{h}_2\|$ is the effective gain.

EXERCISE 3.21. 1. For spatial multiplexing, each stream comes from a BPSK constellation $(+/-a)$. Then, to normalize the average transmit power to 1, we get $a = \sqrt{0.5}$. Then, the probability of pairwise error is upper bounded as:

$$\begin{aligned}\mathbb{P}(\mathbf{x}_1 \to \mathbf{x}_2) &\leq \frac{16}{\mathsf{SNR}^2 ||\mathbf{x}_1 - \mathbf{x}_2||^4}, \\ &\leq \frac{16}{\mathsf{SNR}^2 (2a)^4}, \\ &\leq \frac{4}{\mathsf{SNR}^2}.\end{aligned}$$

For the Alamouti scheme, we should be using the same rate. Thus, each symbol should come from a four point PAM $(+/-a, +/-3a)$. Then to normalize the average transmit power we have:

$$\begin{aligned}\frac{1}{4} 2 \left(2a^2 + 18a^2\right) &= 1, \\ a^2 &= 1/10.\end{aligned}$$

Now, the probability of pair-wise error is upper bounded by:

$$\begin{aligned}\mathbb{P}(\mathbf{X}_A \to \mathbf{X}_B) &\leq \left(\frac{4^2}{\mathsf{SNR}^2 \det\left((\mathbf{X}_A - \mathbf{X}_B)(\mathbf{X}_A - \mathbf{X}_B)^*\right)}\right)^2, \\ &\leq \frac{4^4}{\mathsf{SNR}^4 ((u_{1_a} - u_{1_b})^2 + (u_{2_a} - u_{2_b})^2)^2}, \\ &\leq \frac{256}{\mathsf{SNR}^4 4^4 a^8}, \\ &\leq \frac{10000}{\mathsf{SNR}^4}.\end{aligned}$$

For $\mathsf{SNR} > 50$, Alamouti scheme outperforms the spatial multiplexing scheme.

2. For general R, the only thing that changes is the value of the minimum distance between the points. For a $2^{R/2}$ point PAM in case of spatial multiplexing, the minimum distance is given by:

$$a^2 = \frac{3}{2(2^R - 1)}.$$

Then, the probability of error is upper bounded by:

$$\mathbb{P}(\mathbf{x}_1 \to \mathbf{x}_2) \leq \frac{16}{\mathsf{SNR}^2 (2a)^4},$$

$$\leq \frac{4(2^R-1)^2}{9\mathsf{SNR}^2}.$$

For the Alamouti scheme, the minimum distance is given by:

$$a^2 = \frac{3}{2(2^{2R}-1)}.$$

Thus, the probability of error is given by

$$\begin{aligned}\mathbb{P}(\mathbf{X}_A \to \mathbf{X}_B) &\leq \frac{1}{\mathsf{SNR}^4 a^8}, \\ &= \frac{16(2^{2R}-1)^4}{81\mathsf{SNR}^4}.\end{aligned}$$

The SNR threshold for general R turns out to be approximately 0.662^{3R} which increases exponentially in R.

Exercise 3.22. When using QAMs, the only thing that changes is the minimum distance. For the spatial multiplexing scheme, the QAM constellation size is $2^{R/2}$. Thus, energy for the the $2^{R/2}$ QAM is given by

$$2a^2\frac{2^{R/2}-1}{3}.$$

Thus, to normalize the total transmit power per unit time, we get:

$$a^2 = \frac{3}{4(2^{R/2}-1)}.$$

Then, the probability of error is upper bounded by:

$$\begin{aligned}\mathbb{P}(\mathbf{x}_1 \to \mathbf{x}_2) &\leq \frac{1}{\mathsf{SNR}^2 a^4}, \\ &\leq \frac{16(2^{R/2}-1)^2}{9\mathsf{SNR}^2}.\end{aligned}$$

For the Alamouti scheme, the minimum distance is given by:

$$a^2 = \frac{3}{4(2^R-1)}.$$

Thus, the probability of error is given by

$$\begin{aligned}\mathbb{P}(\mathbf{X}_A \to \mathbf{X}_B) &\leq \frac{1}{\mathsf{SNR}^4 a^8}, \\ &= \frac{256(2^R-1)^4}{81\mathsf{SNR}^4}.\end{aligned}$$

EXERCISE 3.23. 1. Let $t = \min(m, k)$. Consider the repetition scheme in which we transmit the same symbol over a different antenna at each time until either we run out of antennas (in which case we can cycle again over the same antennas) or the block of lenght k ends. We further repeat the transmission n times over different blocks of length k. The antennas are used one at a time.

If $\mathbf{X}$ is the matrix representing the transmitted codeword, with $\mathbf{X}(i,j)$ denoting the signal transmitted through antenna i at time j ($1 \leq i \leq m$, $1 \leq j \leq kn$) the codewords of the above repetition scheme are of the form:

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} x$$

where x is the scalar symbol we want to transmit and for simplicity we have assumed $m = k$.

From the channel model, we see that the same symbol is transmitted over $t \cdot n$ independent channel realizations, resulting in a diversity gain of $t \cdot n = \min(m, k) \cdot n$. We will see in the next part that this simple repetition code achieves the maximum possible diversity gain.

2. The channel model is:
$$\mathbf{y}[i]^T = \mathbf{h}[i]^T \mathbf{X}[i] + \mathbf{z}[i]$$
where i, $1 \leq i \leq n$ is the block index, $\mathbf{y}[i] \in \mathcal{C}^k$ is the received signal for block i, $\mathbf{h}[i] \in \mathcal{C}^m$ is the channel vector assumed constant for block i and independent across blocks, $\mathbf{X}[i] \in \mathcal{C}^{mxk}$ is the part of the codeword transmitted during the block i, and $\mathbf{z}[i]$ is i.i.d. white Gaussian noise.

At the receiver the effective received codeword is:

$$\mathbf{v} = \left[\mathbf{h}[1]^T \mathbf{X}[1], \mathbf{h}[2]^T \mathbf{X}[2], \cdots, \mathbf{h}[n]^T \mathbf{X}[n]\right]^T \in \mathcal{C}^{kn}$$

Conditioned on the channel realization $\mathbf{h}[1], \mathbf{h}[2], \ldots, \mathbf{h}[n]$, the pairwise error probability depends only on $\|v_1 - v_2\|^2$ through the expression:

$$P_{(\mathbf{X}_1 \to \mathbf{X}_2)|\mathbf{h}} = Q\left(\sqrt{\frac{\|\mathbf{v}_2 - \mathbf{v}_1\|^2}{2N_0}}\right)$$

which can be rewritten and upper bounded as:

$$P_{(\mathbf{x}_1 \to \mathbf{x}_2)|\mathbf{h}} = Q\left[\sqrt{\sum_{i=1}^{n} \frac{\|\mathbf{h}[i]^T (\mathbf{X}_2[i] - \mathbf{X}_1[i])\|^2}{2N_0}}\right]$$

$$\leq \prod_{i=1}^{n} \exp\left[-\frac{\|\mathbf{h}[i]^T(\mathbf{X}_2[i]-\mathbf{X}_1[i])\|^2}{4N_0}\right]$$

By taking expectations over $\mathbf{h}[i], i = 1, 2, \ldots, n$ and using the fact that different blocks experience independent fadings we obtain:

$$P_{(\mathbf{x}_1 \to \mathbf{x}_2)} \leq \prod_{i=1}^{n} E\left\{\exp\left[-\frac{\|\mathbf{h}[i]^T(\mathbf{X}_2[i]-\mathbf{X}_1[i])\|^2}{4N_0}\right]\right\}$$

Let $\mathbf{A}[i] = (\mathbf{X}_2[i] - \mathbf{X}_1[i])(\mathbf{X}_2[i] - \mathbf{X}_1[i])^*$ [2] . Since $\mathbf{A}[i]$ is positive semidefinite it can be expressed as $\mathbf{A}[i] = \mathbf{V}[i]\mathbf{\Lambda}[i]\mathbf{V}[i]^*$ where $\mathbf{V}[i]$ is unitary and $\mathbf{\Lambda}[i] = diag(\lambda_1^2[i], \ldots, \lambda_m^2[i])$.

Then we have:

$$\begin{aligned}\|\mathbf{h}[i]^T(\mathbf{X}_2[i]-\mathbf{X}_1[i])\|^2 &= \mathbf{h}[i]^T\mathbf{V}[i]\mathbf{\Lambda}[i]\mathbf{V}[i]^*\mathbf{h}[i]^{*T} = \tilde{\mathbf{h}}[i]^T\mathbf{\Lambda}[i]\tilde{\mathbf{h}}[i]^{*T} \\ &= \sum_{\ell=1}^{m} |\tilde{h}_\ell[i]|^2\lambda_\ell^2[i]\end{aligned}$$

where $\tilde{\mathbf{h}}[i]^T = \mathbf{h}[i]^T\mathbf{V}[i]$ has the same distribution as $\mathbf{h}[i]^T$ because the entries of $\mathbf{h}[i]$ are i.i.d. $\mathbf{CN}(0,1)$ and $\mathbf{V}[i]$ is unitary.

Therefore letting $t_i = rank(\mathbf{A}[i]) \leq \min(m,k)$,

$$\begin{aligned}P_{(\mathbf{x}_1 \to \mathbf{x}_2)} &\leq \prod_{i=1}^{n}\prod_{\ell=1}^{m} E\left\{\exp\left[-\frac{\lambda_\ell^2[i]|h_\ell[i]|^2}{4N_0}\right]\right\} = \prod_{i=1}^{n}\prod_{\ell=1}^{m}\frac{1}{1+\lambda_\ell^2[i]/4N_0} \\ &\leq \prod_{i=1}^{n}\prod_{\ell=1}^{t_i}\frac{4N_0}{\lambda_\ell^2[i]}\end{aligned}$$

where we have assumed $\lambda_1^2[i] \geq \lambda_2^2[i] \geq \cdots \geq \lambda_{t_i}^2[i] > 0$ for $i = 1, 2, \ldots, n$.

Finally we get the pairwise code design criterion: choose the code so that $\mathbf{A}[i]$ is full rank for all $1 \leq i \leq n$ and for all pairs of codewords (which assures a maximum diversity gain equal to $n \cdot \min(m,k)$) and among those codes, choose the one that maximizes the minimum product $\prod_{i=1}^{n}\prod_{\ell=1}^{t_i}\lambda_\ell^2[i] = \prod_{i=1}^{n}\det(\mathbf{A}[i])$ over all pairs of codewords.

Note that the repetition code proposed in a) with $x = \pm 1$ achieves full diversity gain.

[2] $\mathbf{A}[i]$ depends on the pair of codewords considered. To make the notation simpler we avoided using more indices to explicitly show this dependence.

For pure time diversity we set $m = 1$ in which case the matrices $\mathbf{A}[i]$ reduce to complex scalars with eigenvalues $|\mathbf{A}[i]| = \|\mathbf{x}_2[i] - \mathbf{x}_1[i]\|^2$ and the pairwise error probability depends on the product distance between the pair of codewords $\prod_{i=1}^n \|\mathbf{x}_2[i] - \mathbf{x}_1[i]\|^2$.

For pure spatial diversity we set $n = 1$ in which case there is only one matrix $\mathbf{A}$ for each pair of codewords. The design criterion reduces to choosing the code such that $\mathbf{A}$ is full rank for all pair of codewords, and among those codes that satisfy this, choose the one that maximizes the minimum $\det(\mathbf{A})$ for all pair of codewords. This criterion is the one obtained in class.

EXERCISE 3.24.

1. Assume uniform scattering around the mobile. In this case the fading correlation is independent of the direction of movement, so we can assume that the mobile is moving in the same direction as where the second antenna would be located. The mobile reaches the location of the second antenna after traveling a distance d, after a time interval d/v where v is the speed of movement. The correlation of the fading gains between these 2 locations is given by $R[\lfloor d/v * W \rfloor]$, where $1/W$ is the sampling interval of the discrete time model.

   The fading gains are zero mean, circularly symmetric jointly complex Gaussian, so their joint distribution is completely determined by their correlation matrix. Letting $\mathbf{h} = [h_1 h_2]^T$ be the gains at the locations of the 2 antennas at a given time, we have:
$$\mathbf{h} \sim CN(\mathbf{0}, \mathbf{K_h}) \tag{3.15}$$
   where
$$\mathbf{K_h} = \begin{bmatrix} R[0] & R[\lfloor d/v \cdot W \rfloor] \\ R[\lfloor d/v \cdot W \rfloor]^* & R[0] \end{bmatrix} \tag{3.16}$$

2. The received signal at the 2 antennas is $\mathbf{y}[n] = \mathbf{h}x[n] + \mathbf{w}[n]$ where $\mathbf{w}[n] \sim CN(\mathbf{0}, N_0\mathbf{I}_2)$. Conditioned on $\mathbf{h}$ the error probability for BPSK is $Q(\sqrt{2\|\mathbf{h}\|^2 \mathsf{SNR}})$ and the average error probability is:
$$P_e = E\left[Q(\sqrt{2\|\mathbf{h}\|^2 \mathsf{SNR}})\right] \tag{3.17}$$
   where the expectation is taken with respect to the distribution of $\mathbf{h}$ which was found in part (1).

3. The problem of directly doing a high SNR approximation in the computation of (3.17) is that the 2 components of $\mathbf{h}$ are correlated. However, noting that any nonsingular transformation of $\mathbf{y}[n]$ is a sufficient statistic, we can do a transformation that decorrelates the entries of $\mathbf{h}$.

   Let $\mathbf{K_h} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$ where $\mathbf{U}$ is unitary and $\mathbf{\Lambda}$ is diagonal with non-negative entries. We can always find this decomposition because $\mathbf{K_h}$ is positive semidefinite. Then define $\tilde{\mathbf{y}}[n] = \mathbf{U}^*\mathbf{y} = \mathbf{U}^*\mathbf{h}[n]x[n] + \mathbf{U}^*\mathbf{w}[n] = \tilde{\mathbf{h}}[n]x[n] + \tilde{\mathbf{w}}[n]$. It follows that $\tilde{\mathbf{h}}[n] \sim CN(\mathbf{0}, \mathbf{\Lambda})$ and $\tilde{\mathbf{w}}[n] \sim CN(\mathbf{0}, N_0\mathbf{I}_2)$. Now we can use the error probability expression found in b) but with uncorrelated fading gains, which in the case of circularly symmetric complex Gaussian random variables implies independence, and use a high SNR approximation for the Q function:

$$\begin{aligned} P_e &= \mathbb{E}\left[Q\left(\sqrt{2\|\tilde{\mathbf{h}}\|^2\mathsf{SNR}}\right)\right] \approx \mathbb{E}\left[e^{-\|\tilde{\mathbf{h}}\|^2\mathsf{SNR}}\right] \\ &= \mathbb{E}\left[e^{-|\tilde{h_1}|^2\mathsf{SNR}}\right]\mathbb{E}\left[e^{-|\tilde{h_2}|^2\mathsf{SNR}}\right] = \frac{1}{(\lambda_1\mathsf{SNR}+1)}\frac{1}{(\lambda_2\mathsf{SNR}+1)} \end{aligned} \tag{3.18}$$

where λ_1 and λ_2 are the diagonal elements of $\mathbf{\Lambda}$ (the eigenvalues of $\mathbf{K_h}$). These eigenvalues can be computed explicitly. Assuming $R[0] = 1$ and letting $\rho = R[\lfloor d/v \cdot W \rfloor]$ we obtain $\lambda_1 = 1 + |\rho|$ and $\lambda_2 = 1 - |\rho|$. If $|\rho| > 0$ then $\lambda_i > 0$ $(i = 1, 2)$, and we can further approximate (3.18) for high SNR:

$$P_e \approx \frac{1}{\lambda_1\lambda_2\mathsf{SNR}^2} = \frac{1}{(1-|\rho|^2)\mathsf{SNR}^2} \tag{3.19}$$

In this case we get a diversity gain of 2 and the correlation between antennas increases the error probability by the factor $1/(1 - |\rho|^2)$ as compared to the uncorrelated antenna case. If on the other hand $|\rho| = 1$ (perfect correlation between antennas) then $\lambda_2 = 0$ and $P_e \approx \frac{1}{2\mathsf{SNR}}$. In this case the diversity gain reduces to 1.

As we increase the antenna separation d the correlation $|\rho|$ decreases since the correlation function $R[m]$ is monotonically decreasing, and the probability of error decreases.

Exercise 3.25. We have

$$\mathbf{y}^t = \begin{bmatrix} h_1 & h_2 & \cdots & h_L \end{bmatrix} \begin{bmatrix} x[1] & x[2] & x[3] & \cdots \\ 0 & x[1] & x[2] & x[3] \\ 0 & 0 & \ddots & \ddots \\ 0 & 0 & 0 & x[1] \end{bmatrix} + \mathbf{w}$$

Now, we want to argue that the typical decoding error is when all the channel gains h_is are small (i.e. $|h_i|^2 < 1/\mathsf{SNR}$). Consider the following sub-optimal decoder:

- Suppose $|h_1|^2 > 1/\mathsf{SNR}^{1-\epsilon}$: then using the first received symbol alone, we can decode $x[1]$ and then using the second symbol, after subtracting of $x[1]$, decode $x[2]$ and so on.

- Now suppose $|h_1|^2 < 1/\mathsf{SNR}$ but $|h_2|^2 > 1/\mathsf{SNR}^{1-\epsilon}$, then we use the second received symbol to decode $x[1]$. Note that since the first channel tap is small the interference caused in the second symbol because of $x[2]$ is also small, and hence we can decode $x[1]$ using the second received symbol. Similarly, we can use the third symbol to decode $x[2]$.
- And so on for $|h_1|^2, |h_2|^2, \cdots, |h_m|^2 < 1/\mathsf{SNR}$ and $|h_{m+1}|^2 > 1/\mathsf{SNR}^{1-\epsilon}$: we use the $m+1$th symbol to decode $x[1]$.

For this sub-optimal decoder because of the tail behavior of the Q-function, even if one of the channels gains is larger that $1/\mathsf{SNR}^{1-\epsilon}$, then the probability of error decays exponentially in SNR. Thus, typically error happens only when *all* the channel gains are small which gives us a best possible diversity order of L.

Exercise 3.26. 1. With $x[0] = \pm 1$,

$$\det(\mathbf{X}_A - \mathbf{X}_B)(\mathbf{X}_A - \mathbf{X}_B)^* \geq 4^L$$

and

$$\mathbb{P}\{\mathbf{X}_A \to \mathbf{X}_B\} \leq \mathsf{SNR}^{-L}$$

for any pair of codewords that differ in the first component. Hence by the union bound, the probability of error p_0 on the first symbol is

$$p_0 < 2^{(L-1)}\mathsf{SNR}^{-L} \approx (\frac{1}{2\mathsf{SNR}})^L.$$

2. To get the same rate using the naive scheme, one has to use 2^L- PAM. The distance between constellation points is of the order of 2^{-L}. Hence , the error probability is of the order of

$$\left[\frac{4 \cdot 2^L}{\mathsf{SNR}}\right]^L$$

Hence, the first scheme uses a factor of $2^{-(L+1)}$ less energy than the naive scheme for the same error probability. This coding gain is exponential in L (linear in L in dB.)

3. Even if we are trying to calculate the error probability for a middle stream, the determinant is still lower bounded by 4^L (consider the first non-zero stream). Then, the probability of error is upper bounded by: $2^{(N-1)}\mathsf{SNR}^{-L}$.

Exercise 3.27. 1. The channel matrix is

$$\begin{bmatrix} h_0 & 0 \\ h_1 & h_0 \end{bmatrix}$$

2. The ZF equalizer would only look at $y[0]$ since there is interference on the $y[1]$ component. But then it makes an error whenever h_0 is in deep fade. Hence, its diversity is 1. This is less than the diversity gain of 2 for the ML approach.

EXERCISE 3.28. The DFT of the fading coefficients $\{h_\ell\}_{\ell=0}^{L-1}$ after extending the sequence adding $N-L$ zeros is:

$$\tilde{h}_n = \frac{1}{\sqrt{N}} \sum_{\ell=0}^{L-1} h_\ell e^{-j2\pi \ell n/N} \tag{3.20}$$

for $n = 0, 1, \dots, N-1$. Since $\{h_\ell\}_{\ell=0}^{L-1}$ are i.i.d. $CN(0, 1/L)$ it follows that $\{\tilde{h}_n\}_{n=0}^{N-1}$ are circularly symmetric jointly complex Gaussian, so their statistics are completely specified by the correlation function

$$\begin{aligned} R[r] &= E[\tilde{h}_{n+r}\tilde{h}_n^*] = \frac{1}{N}\sum_{\ell=0}^{L-1}\sum_{m=0}^{L-1} E[h_m h_\ell^*] e^{-j\frac{2\pi}{N}[m(n+r)-nl]} \\ &= \frac{1}{NL}\sum_{\ell=0}^{L-1} e^{-j\frac{2\pi}{N}\ell r} = \frac{1}{NL}\frac{\sin\left(\frac{\pi r L}{N}\right)}{\sin\left(\frac{\pi r}{N}\right)} e^{-j\frac{r\pi(L-1)}{N}}(1-\delta[r]) + \frac{1}{N}\delta[r] \end{aligned} \tag{3.21}$$

valid of $|r| \le (N-1)/2$.

The coherence bandwidth is given by $W_c = W/(2L)$ and the tone spacing is W/N, so the coherence bandwidth expressed in number of tones is $N/(2L)$. We expect the gains of the carriers separated by more than the coherence bandwidth to be approximately independent. To state this more precisely, we want to show that $|R[r]| \ll R[0]$ for $\alpha N/(2L) \le |r| \le (N-1)/2$, where we have assumed N to be odd for simplicity, and α is some constant, say 10. Note that the carrier gains are periodic in n so the correlation function is also periodic in r with period N, so we need to consider values of r only within one period.

To prove this statement we need to upper bound $|R[r]|$ for $\alpha N/(2L) \le |r| \le (N-1)/2$. We first note that $|\sin\left(\frac{\pi r L}{N}\right)| \le 1$ and $|\sin(x)| \ge 2|x|/\pi$ for $|x| \le \pi/2$, so $\sin\left(\frac{\pi r}{N}\right) \ge 2|r|/N \ge \alpha/L$ in the range of interest. Therefore for $\alpha N/(2L) \le |r| \le (N-1)/2$,

$$|R[r]| \le \frac{1}{NL}\frac{L}{\alpha} = \frac{1}{\alpha N} = \frac{1}{\alpha}R[0] \tag{3.22}$$

It follows that the correlation decays at least as α^{-1} for tones separated by α times the coherence bandwidth.

EXERCISE 3.29. In outdoor environments the difference in the distances travelled by different paths is of the order of a few hundred meters. For example, for a path difference of 300m the delay spread is 1μs, so we can say in general that T_d is in the order of a microsecond. On the other hand $T_c \sim 1/(4\mathcal{D}) \sim 4.5$ms for a mobile speed of 60km/h and a carrier frequency of 1GHz. We see that $T_c >> T_d$ in typical scenarios.

OFDM: OFDM assumes that the channel remains constant during the transmission of an OFDM block, which requires $N \ll T_c W$. The overhead (both in time and power) incurred by the use of the cyclic prefix is $L/(L+N)$, where $L = T_d W$ is the number of taps. To have small overhead we need $L \ll N$, and this implies $T_d \ll T_c$. Thus the underspread condition is required for a small overhead.

DSSS: The Rake receiver requires the channel to remain constant during the transmission of the spreading sequence of length n, so $n \ll T_c W$. Also we require $n \gg L$ for the ISI to be negligible. These two conditions together require that $T_d \ll T_c$.

Channel Estimation: In both systems we need to estimate the channel. We define $L_{crit} = K\mathcal{E}/N_0$ as a threshold for determining the channel estimation performance. The channel estimation error is small when $L \ll L_{crit}$, or equivalently when $T_d \ll K\mathcal{E}/(WN_0)$. Also $K \ll T_c W$ for the channel to remain constant during the measurement interval. Thus we need $T_d \ll T_c\mathcal{E}/N_0$. A large ratio T_c/T_d makes the channel estimation easier by requiring a smaller SNR.

Note that in the case of OFDM we have control over the number of carriers over which we spread our energy, so it is possible to have small estimation error even when the condition $L \ll L_{crit}$ is not met.

Exercise 3.30.

1. We know that the taps $h_\ell[i]$'s are circularly symmetric. For fixed time i and different values of ℓ the taps are independent random variables. This follows because the different taps correspond to different signal paths which experience independent reflections (attenuations and phase shifts). The complex gain of the n^{th} carrier at the i^{th} OFDM symbol, for an OFDM block length of N and L taps is given by:

$$\tilde{h}_n[i] = \frac{1}{\sqrt{N}} \sum_{\ell=0}^{L-1} h_\ell[i] e^{-j2\pi \ell n/N} \tag{3.23}$$

   The multiplication by the complex exponential does not modify the distribution of $h_\ell[i]$. Also, since the L terms in the sum are independent, the multiplication by the complex exponentials does not modify the joint distribution of the different terms, and hence does not modify the distribution of the sum. Since the dependence on n only appears through the complex exponential, and we can remove the complex exponentials without modifying the distribution of $\tilde{h}_n[i]$, it follows that the $\{\tilde{h}_n[i]\}_{n=1}^N$ are identically distributed for fixed i.

2. From the text, we know that the effect of movement of paths from tap to tap is negligible compared to the variation of the taps due to Doppler spread whenever $f_c \gg W$. Making this assumption the movement of paths between taps occurs in a time scale much larger than the variation due to Doppler shifts, and therefore

we have that the processes $\{h_\ell[m]]\}_m$ are independent across ℓ. Now we can use an argument similar to that of part (1) but using vectors.

Let $\tilde{\mathbf{h}}_n = [\tilde{h}_n[1] \dots \tilde{h}_n[r]]^T$, $n = 1, 2, \dots, N$, and $\mathbf{h}_\ell = [h_\ell[1] h_\ell[2] \dots h_\ell[r]]^T$, $\ell = 0, 1, \dots, L-1$. We can write:

$$\tilde{\mathbf{h}}_n = \frac{1}{\sqrt{N}} \sum_{n=0}^{L-1} \mathbf{h}_\ell e^{-j2\pi n\ell/N} \tag{3.24}$$

We can use the result of Exercise 2.16 to conclude that the vectors $\mathbf{h}_\ell$, $\ell = 0, 1, \dots, L-1$ are circularly symmetric. Hence the distribution of each of the terms in the above sum is not modified by the product with the complex exponential. Also the initial observation implies that the vectors $\mathbf{h}_\ell$, $\ell = 0, 1, \dots, L-1$ are independent, and therefore the sum (3.24) is not modified by the product with the complex exponentials. Since the dependence on n is only through the complex exponentials, and these can be removed without modifying the distribution of the sum, it follows that $\{\tilde{\mathbf{h}}_n\}_{n=1}^N$ are identically distributed. Since r is arbitrary the result follows.

EXERCISE 3.31. Let $\mathbf{r} = (\mathbf{r}_A^T \mathbf{r}_B^T)^T$, where $\mathbf{r}_A$ and $\mathbf{r}_B$ are as defined in Exercise 3.6. Define

$$\mathbf{C} = \left(\frac{\mathcal{E}_b}{L} + N_0\right) \mathbf{I}_L, \quad \mathbf{D} = N_0 \mathbf{I}_L, \quad \mathbf{\Sigma}_A = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}, \quad \mathbf{\Sigma}_B = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \tag{3.25}$$

Then $(\mathbf{r} \mid \mathbf{x}_A) \sim CN(0, \mathbf{\Sigma}_A)$ and $(\mathbf{r} \mid \mathbf{x}_B) \sim CN(0, \mathbf{\Sigma}_B)$, so noting that $\det \mathbf{\Sigma}_A = \det \mathbf{\Sigma}_B$ the log-likelihood ratio is given by:

$$\mathcal{L}(\mathbf{r}) = \log \left[\frac{p(\mathbf{r} \mid \mathbf{x}_A)}{p(\mathbf{r} \mid \mathbf{x}_B} \right] = \mathbf{r}^* (\mathbf{\Sigma}_B^{-1} - \mathbf{\Sigma}_A^{-1}) \mathbf{r} = (\|\mathbf{r}_A\|^2 - \|\mathbf{r}_B\|^2) \frac{\mathcal{E}_b}{N_0(\mathcal{E}_b + LN_0)} \tag{3.26}$$

By comparing the log-likelihood ratio to 0 we obtain the decision rule (3.63) of the notes.

We now compute the error probability:

$$P_e = Pr\left[\|\mathbf{r}_A\|^2 > \|\mathbf{r}_B\|^2 \mid \mathbf{x}_B\right] = Pr\left[\|\mathbf{w}_A\|^2 > \|\sqrt{\mathcal{E}_b}\mathbf{h} + \mathbf{w}_B\|^2\right] \tag{3.27}$$

$\|\mathbf{w}_A\|^2$ and $\|\sqrt{\mathcal{E}_b}\mathbf{h} + \mathbf{w}_B\|^2$ have χ^2 distributions with densities $f_1(x) = \frac{1}{N_0^L (L-1)!} x^{L-1} e^{-x/N_0}$ and $f_2(x) = \frac{1}{(N_0 + \mathcal{E}_b/L)^L (L-1)!} x^{L-1} e^{-x/(N_0 + \mathcal{E}_b/L)}$ for $x \geq 0$. Letting $T = \|\mathbf{w}_A\|^2 - \|\sqrt{\mathcal{E}_b}\mathbf{h} + \mathbf{w}_B\|^2$ with density:

$$f_T(t) = \int_{-\infty}^{\infty} f_1(l) f_2(l-t) dl \tag{3.28}$$

we can express the error probability by:

$$
\begin{aligned}
P_e &= Pr\left[T > 0\right] = \int_0^\infty \int_t^\infty \frac{1}{(L-1)!^2} \frac{l^{L-1}(l-t)^{L-1}}{[N_0(N_0+\mathcal{E}_b/L)]^L} e^{-l/N_0} e^{-\frac{l-t}{N_0+\mathcal{E}_b/L}} dl dt \\
&= \int_0^\infty \int_0^\infty \frac{1}{(L-1)!^2} \frac{(s+t)^{L-1} s^{L-1}}{[N_0(N_0+\mathcal{E}_b/L)]^L} e^{-(s+t)/N_0} e^{-\frac{s}{N_0+\mathcal{E}_b/L}} ds dt \\
&= \frac{1}{(L-1)!^2} \sum_{k=0}^{L-1} \binom{L-1}{k} \int_0^\infty \int_0^\infty \frac{s^k t^{L-1-k} s^{L-1}}{\beta^L} e^{-t/N_0} e^{-s\left(\frac{1}{N_0}+\frac{1}{N_0+\mathcal{E}_b/L}\right)} ds dt \qquad (3.29) \\
&= \frac{1}{(L-1)!^2} \sum_{k=0}^{L-1} \binom{L-1}{k} \frac{(L+k-1)!}{\alpha^{L+k}} \int_0^\infty \frac{t^{L-1-k}}{\beta^L} e^{-t/N_0} dt \qquad (3.30) \\
&= \frac{1}{(L-1)!^2} \sum_{k=0}^{L-1} \binom{L-1}{k} \frac{(L+k-1)!}{\alpha^{L+k}} \frac{(L-1-k)!}{\beta^L} N_0^{L-k} \qquad (3.31) \\
&= \left(\frac{N_0}{\alpha\beta}\right)^L \sum_{k=0}^{L-1} \binom{L+k-1}{k} \left(\frac{1}{N_0\alpha}\right)^k
\end{aligned}
$$

where we have defined $\alpha = \left(\frac{1}{N_0} + \frac{1}{N_0+\mathcal{E}_b/L}\right)$ and $\beta = N_0(N_0+\mathcal{E}_b/L)$. Here (3.29) is obtained by the binomial expansion formula, and (3.30) and (3.31) follow by integrating the χ^2 densities over their domains. This last formula is just what is required.

Exercise 3.32. Define $r_A^{(\ell)}$ and $r_B^{(\ell)}$ by:

$$
\begin{aligned}
r_A^{(\ell)} &= h[\ell]x_1 + w_A^{(\ell)}, \quad \ell = 1,2,\ldots,L \\
r_B^{(\ell)} &= h[\ell]x_2 + w_B^{(\ell)}, \quad \ell = 1,2,\ldots,L \qquad (3.32)
\end{aligned}
$$

where $x_1 = \|x_A\|$ and $x_2 = 0$ if $\mathbf{x}_A$ is transmitted, and $x_1 = 0$ and $x_2 = \|x_B\|$ if $\mathbf{x}_B$ is transmitted.

Define $r_k^{(\ell)}$ as in (3.152) in the text. Then the channel estimates are:

$$
\begin{aligned}
\hat{h}[\ell] &= \frac{\mathcal{E}/N_0}{K\mathcal{E}/N_0 + L} \frac{1}{\sqrt{\mathcal{E}}} \left(K\sqrt{\mathcal{E}}h[\ell] + \tilde{w}^{(\ell)}\right) \qquad (3.33) \\
&= ah[\ell] + \bar{w}^{(\ell)} \qquad (3.34)
\end{aligned}
$$

for $\ell = 0,1,\ldots,L-1$, where $\tilde{w}^{(\ell)} \sim CN(0, KN_0)$, i.i.d., $a = \frac{K\mathcal{E}/N_0}{K\mathcal{E}/N_0+L}$ and $\bar{w}^{(\ell)} \sim CN\left(0, \frac{K\mathcal{E}/N_0}{(K\mathcal{E}/N_0+L)^2}\right)$, i.i.d..

The coherent receiver projects $\mathbf{r}_A$ and $\mathbf{r}_B$ onto $\hat{\mathbf{h}}$ obtaining:

$$
\mathbf{y}_A = \hat{\mathbf{h}}_A^* \mathbf{r}_A = a\|\mathbf{h}\|^2 x_1 + \bar{\mathbf{w}}^* \mathbf{h} x_1 + (a\mathbf{h} + \bar{\mathbf{w}})^* \mathbf{w}_A
$$

$$\mathbf{y}_B = \hat{\mathbf{h}}_B^* \mathbf{r}_B = a\|\mathbf{h}\|^2 x_2 + \bar{\mathbf{w}}^* \mathbf{h} x_2 + (a\mathbf{h} + \bar{\mathbf{w}})^* \mathbf{w}_B \tag{3.35}$$

where $x_1 = \|x_A\|$ and $x_2 = 0$ if $\mathbf{x}_A$ is transmitted, and $x_1 = 0$ and $x_2 = \|x_B\|$ if $\mathbf{x}_B$ is transmitted. We assume $\|x_A\|^2 = \|x_B\|^2 = \mathcal{E}_b = \mathcal{E}$. Also $\mathbf{w}_A$ and $\mathbf{w}_B$ are independent $CN(0, N_0\mathbf{I}_L)$ random vectors, independent of everything else.

The probability of error is given by:

$$\begin{aligned} P_e &= Pr\left(\Re\{\mathbf{y}_A\} > \Re\{\mathbf{y}_B\} \mid \mathbf{x}_B\right) \\ &= Pr\left[\Re\{(a\mathbf{h} + \bar{\mathbf{w}})^* \mathbf{w}_A\} > \Re\{a\|\mathbf{h}\|^2 \sqrt{\mathcal{E}} + \sqrt{\mathcal{E}} \bar{\mathbf{w}}^* \mathbf{h} + (a\mathbf{h} + \bar{\mathbf{w}})^* \mathbf{w}_B\}\right] \\ &= Pr\left[\Re\left\{(a\mathbf{h} + \bar{\mathbf{w}})^* \left(\frac{\mathbf{w}_A - \mathbf{w}_B}{\sqrt{\mathcal{E}}}\right)\right\} > \Re\{a\|\mathbf{h}\|^2 + \bar{\mathbf{w}}^* \mathbf{h}\}\right] \end{aligned} \tag{3.36}$$

Since it is hard to compute the above probability explicitly, we opt for computing it approximately by simulation.

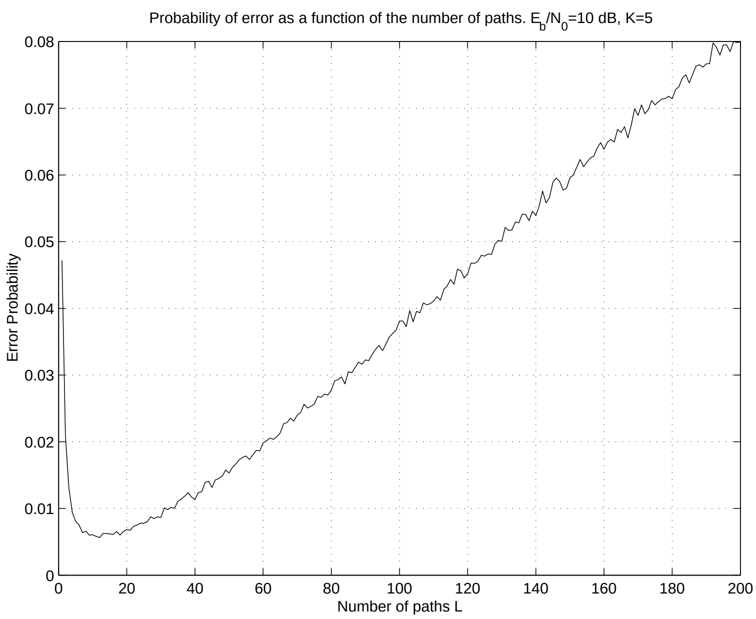


Figure 3.2: Probability of error as a function of the number of paths for $\mathcal{E}/N_0 = 10$dB and $K = 5$.

We see in Figure 3.2 that the probability of error starts degrading for $L \approx 15$. For this particular choice of the parameters $L_{cr} = K\mathcal{E}/N_0 = 50$ so for $L \gg L_{cr}$ the performance of the detector is very poor.

Exercise 3.33.

Chapter 4

Solutions to Exercises

EXERCISE 4.1. 1. For this example C=7 and M=10. The maximal allowable subsets are enumerated as below,

$$\begin{aligned} S_1 &= \{7\} \\ S_2 &= \{1,3\} \\ S_3 &= \{1,4\} \\ S_4 &= \{1,5\} \\ S_5 &= \{2,4\} \\ S_6 &= \{2,5\} \\ S_7 &= \{2,6\} \\ S_8 &= \{3,5\} \\ S_9 &= \{3,6\} \\ S_{10} &= \{4,6\} \end{aligned}$$

2. The matrix A can be represented as,

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

EXERCISE 4.2. 1. Let $T_i(r)$ be the average traffic per channel supported in the cell i. Then since the average traffic supported cannot exceed the actual traffic present in cell i we have,

$$\begin{aligned} T_i(r) &\leq \text{ Average traffic per channel in cell i} \\ &= p_i r. \end{aligned}$$

Therefore adding over all the cells we have,

$$\begin{aligned} T(r) &= \sum_{i=1}^{C} T_i(r), \\ &= \sum_{i=1}^{C} p_i r, \\ &= \frac{\sum_{i=1}^{C} p_i B}{N}, \\ &= \frac{\sum_{i=1}^{C} \sum_{l=1}^{N} I \left\{\text{Channel l used in cell i}\right\}}{N}, \end{aligned}$$

where $I_{\{\text{ Channel l used in cell i}\}}$ is the indicator function which is equal to 1 if channel l is used in cell i and zero otherwise. Exchanging the order of summation, we have,

$$T(r) \leq \frac{1}{N} \sum_{l=1}^{N} \sum_{i=1}^{C} I_{\{\text{ Channel l used in cell i}\}}$$

Now each channel l is assigned to a subset of one of the maximal sets (note that it does not have to be a strict subset). Lets assume that the channel l is assigned to a subset of the maximal set S_l which corresponds to the column j in the adjacency matrix. Therefore,

$$\begin{aligned} T(r) &\leq \frac{1}{N} \sum_{l=1}^{N} \sum_{i \in S_l} 1, \\ &= \frac{1}{N} \sum_{l=1}^{N} \sum_{i=1}^{C} a_{ij}, \\ &\leq \frac{1}{N} \sum_{l=1}^{N} \max_{j=1,\ldots,M} \sum_{i=1}^{C} a_{ij}, \\ &\leq \frac{1}{N} N \max_{j=1,\ldots,M} \sum_{i=1}^{C} a_{ij}, \end{aligned}$$

$$= \max_{j=1,\dots,M} \sum_{i=1}^{C} a_{ij}.$$

2. As noted in part (1) above: Let $T_i(r)$ be the average traffic per channel supported in the cell i. Then since the average traffic supported cannot exceed the actual traffic present in cell i we have,

$$\begin{aligned} T_i(r) &\leq \text{ Average traffic per channel in cell i} \\ &= p_i r. \end{aligned}$$

Then, summing over all the cells, we get

$$T_i(r) \leq \sum_{i=1}^{C} p_i r = r.$$

3. To combine the two upper bounds, simply choose a set of numbers $y_i \in [0, 1]$ for $i = 1, \dots, C$ and observe that

$$T_i(r) \leq p_i r = y_i p_i r + (1 - y_i) p_i r.$$

Then, summing over all $i = 1, \dots, C$ and applying the bound from part (1) to the second term and bound from part (2) to the first term, we obtain

$$T_i(r) \leq \sum_{i=1}^{C} y_i p_i r + \max_{j=1,\dots,M} \sum_{i=1}^{C} (1 - y_i) a_{ij}.$$

Exercise 4.3.

Exercise 4.4. 1.

$$\begin{aligned} s(t) &= \mathcal{R}\left[\sum_{n=0}^{\infty} x[n] \operatorname{sinc}\left(\frac{t - nT}{T}\right) \exp(j 2\pi f_c t)\right] \\ &= \mathcal{R}\left[\sum_{n=0}^{\infty} \operatorname{sinc}\left(\frac{t - nT}{T}\right) \exp(j 2\pi f_c t + \theta_n)\right] \end{aligned}$$

where $x[n] = \exp(j\theta_n)$ and θ_n is uniformly distributed on $[0, 2\pi]$ and independent across time samples n.

Now, the average power in $s(t)$ over a symbol period is given by

$$P_{\text{av}} = \mathbb{E}\left[\frac{1}{T} \int_0^T |s(t)|^2 dt\right]$$

$$\begin{aligned} |s(t)|^2 &= \left(\sum_{n=0}^{\infty} \text{sinc}\left(\frac{t-nT}{T}\right)\cos(2\pi f_c T + \theta_n)\right)^2 \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \text{sinc}\left(\frac{t-nT}{T}\right)\cos(2\pi f_c T + \theta_n)\text{sinc}\left(\frac{t-mT}{T}\right)\cos(2\pi f_c T + \theta_m) \end{aligned}$$

By independence of θ_n and θ_m for $n \neq m$ we have that

$$\begin{aligned} \mathbb{E}[|s(t)|^2] &= \sum_{n=0}^{\infty} \text{sinc}^2\left(\frac{t-nT}{T}\right)\mathbb{E}[\cos^2(2\pi f_c t + \theta_n)] \\ &= \sum_{n=0}^{\infty} \frac{1}{2}\text{sinc}^2\left(\frac{t-nT}{T}\right) \end{aligned}$$

Observe that the series above is uniformly (and absolutely) convergent on $0 \leq t \leq T$. To see this note that

$$\begin{aligned} \sum_{n=0}^{\infty} \text{sinc}^2\left(\frac{t-nT}{T}\right) &= \sum_{n=0}^{\infty} \frac{\sin^2\left(\frac{t}{T}-n\right)}{\left(\frac{t}{T}-n\right)^2} \\ &\leq \sum_{n=0}^{\infty} \frac{1}{\left(\frac{t}{T}-n\right)^2} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} \end{aligned}$$

Hence, we can interchange the summation and integration in computing the average power:

$$\begin{aligned} P_{\text{av}} &= \frac{1}{T}\int_0^T \left\{\sum_{n=0}^{\infty} \frac{1}{2}\text{sinc}^2\left(\frac{t-nT}{T}\right)\right\} dt = \frac{1}{2T}\sum_{n=0}^{\infty}\int_0^T \text{sinc}^2\left(\frac{t-nT}{T}\right) dt \\ &= \frac{1}{2}\sum_{n=0}^{\infty}\int_{u=-n}^{-n+1} \text{sinc}^2(u)du = \frac{1}{2}\int_0^1 \text{sinc}^2(u)du + \frac{1}{2}\int_0^{\infty} \text{sinc}^2(u)du \end{aligned}$$

where $u = \frac{t-nT}{T}$. Using Parseval's identity, the energy in $\text{sinc}(u)$ is the same as the energy in its Fourier transform $\text{rect}(f)$ so

$$P_{\text{av}} = \frac{1}{4} + \frac{1}{2}\int_0^1 \text{sinc}^2(u)du$$

Using MATLAB, the above quantity can be calculated to be $P_{\text{av}} \approx 0.476$.

To estimate the peak-power as a function of T, suppose that the maximum of $|s(t)|^2$ over $[0, T]$ occurs at some $t_0 \in [0, T]$. Then we can write:

$$\max_{0\leq t\leq T} |s(t)|^2 = \left(\sum_{n=0}^{\infty} \frac{\sin\left(\pi\left(\frac{t_0-nT}{T}\right)\right)}{\left(\pi\frac{t_0-nT}{T}\right)}\cos(\theta_n)\right)^2,$$

where we have used the fact that $2\pi f_c T$ is an integer. The time t_0 is a random variable since θ_n is random. However, we can obtain the following simple upper bound by expanding the above expression:

$$\begin{aligned}
\max_{0\le t\le T} |s(t)|^2 \;&\le\; \frac{1}{\pi^2}\sum_{n=2}^{\infty}\sum_{m=2}^{\infty}\frac{\cos\theta_n\cos\theta_m}{\left(\frac{t_0}{T}-n\right)\left(\frac{t_0}{T}-m\right)} \\
&+\sum_{n=0,1}^{\infty}\sum_{m=0,1}^{\infty}\frac{\sin\left(\pi\left(\frac{t_0}{T}-n\right)\right)}{\pi\left(\frac{t_0}{T}-n\right)}\frac{\sin\left(\pi\left(\frac{t_0}{T}-m\right)\right)}{\pi\left(\frac{t_0}{T}-m\right)}\cos\theta_n\cos\theta_m, \qquad (4.1)\\
&\le\; \frac{1}{\pi^2}\sum_{n=2}^{\infty}\sum_{m=2}^{\infty}\frac{\cos\theta_n\cos\theta_m}{(1-n)(1-m)} \qquad (4.2)\\
&\qquad +\sum_{n=0,1}^{\infty}\sum_{m=0,1}^{\infty}\cos\theta_n\cos\theta_m,
\end{aligned}$$

In the first inequality above, the first term was obtained by upper bounding $\sin\left(\pi\left(\frac{t}{T}-n\right)\right)$ by 1. In the second inequality, we use the observation that sinc $\left(\frac{t}{T}-n\right)$ is maximized at $t = nT$ (and this is feasible since $n = 0, 1$) where it evaluates to 1, to obtain the second term. The first term is upper bounded by noting that the minimum of $|\frac{t}{T}-n|$, for $n \neq 0, 1$, occurs when $t = T$.

Now, taking the expectation of the above upper bound over the distribution of the data symbols and observing that θ_n and θ_m are independent for $n \neq m$, we get that:

$$\mathbb{E}\left[\max_{0\le t\le T}|s(t)|^2\right] \;\le\; \frac{1}{2\pi^2}\sum_{n=2}^{\infty}\frac{1}{(n-1)^2}+1.$$

Hence

$$PP := \mathbb{E}\left[\max_{0\le t\le T}|s(t)|^2\right] \;\le\; \frac{1}{2\pi^2}\frac{\pi^2}{6}+1 \approx 1.083.$$

The Peak-to-average-power ratio $\frac{PP}{P_{\text{av}}}$ is then approximately 2.77. Also, from (4.1), we see that as $T \to \infty$, the peak-power approaches this value.

Exercise 4.5. From equation (4.31), for each $k = 1, \ldots, K$, we get,

$$\begin{aligned}
GP_k g_{k,c_k} \;&\ge\; \beta_k \sum_{n\neq k} P_n g_{n,c_k} + N_0 w \beta_k \\
P_k - \frac{1}{G}\sum_{n\neq k} P_n \frac{\beta_k g_{n,c_k}}{g_{k,c_k}} \;&\ge\; \frac{N_0 w \beta_k}{g_{k,c_k}}
\end{aligned}$$

Using the notation $\mathbf{p}$ to denote the vector $(P_1, \ldots, P_K)$, we can rewrite the above K inequalities as,

$$(\mathbf{I}_k - \mathbf{F})\,\mathbf{p} \;=\; \mathbf{b}$$

where,

$$\mathbf{b} \;=\; N_0 W \left(\frac{\beta_1}{g_{1,c_1}}, \ldots \frac{\beta_K}{g_{K,c_K}}\right)$$

and $\mathbf{F}$ is a $K \times K$ matrix given by,

$$f_{ij} \;=\; \begin{cases} 0 & \text{if } i = j \\ \frac{g_{j,c_i}\beta_i}{G g_{i,c_i}} & \text{if } i \neq j \end{cases}$$

Exercise 4.6. 1. To see that $\mathbf{F}$ is irreducible, simply take $m = 2$ and observe that, for $K \geq 3$, we have that $\mathbf{F}^m$ is a matrix with strictly positive entries. Hence, by definition, $\mathbf{F}$ is irreducible.

2. The reference [106] gives an in-depth treatment of non-negative matrices.

3. The following proof is taken from [106]. First we show that (a) implies (b). Suppose that, for a given strictly positive constraint $\mathbf{b} > 0$, there exists a strictly positive solution $\mathbf{p} > 0$ to (4.32), i.e.,

$$(\mathbf{I} - \mathbf{F})\mathbf{p} \geq \mathbf{b}.$$

Then, $\mathbf{b} + \mathbf{F}\mathbf{p} \leq \mathbf{p}$, and hence $\mathbf{F}\mathbf{p} < \mathbf{p}$. Now, since $\mathbf{F}$ is irreducible and non-negative it has a unique left (and right) strictly positive eigenvector $\mathbf{x}^T > 0$ associated with the Perron-Frobenius eigenvalue r. Pre-multiplying both sides of the inequality $\mathbf{p} > \mathbf{F}\mathbf{p}$ by this eigenvector, we get that

$$\mathbf{x}^T\mathbf{p} > \mathbf{x}^T\mathbf{F}\mathbf{p} = r\ \mathbf{x}^T\mathbf{p},$$

and hence we conclude that $r < 1$.

Now, we show that (b) implies (c). Suppose that the Perron-Frobineus eigenvalue $r < 1$. Hence, all other eigenvalues λ_i satisfy $|\lambda_i| < 1$ for $i = 1, \ldots, K-1$. Here, for simplicity, we assume that all the eigenvalues are distinct. The arguments are only slightly more technical if this condition is violated. We can express $\mathbf{F} = \mathbf{B}^{-1}\mathbf{\Lambda}\mathbf{B}^{-1}$, for some matrix $\mathbf{B}$, where $\Lambda = \text{diag}(r, \lambda_1, \ldots, \lambda_K)$. Hence, we can write

$$\mathbf{F}^n = \mathbf{B}^{-1}\text{diag}(r^n, \lambda_1^n, \ldots, \lambda_K^n)\mathbf{B},$$

for $n \geq 1$ integer. Thus, we see that $r^n \to 0$, and $\lambda_i^n \to 0$, for all i, as $n \to \infty$. Hence, $\mathbf{F}^n \to 0$ elementwise. Recall the following matrix lemma: If $\mathbf{F}^n \to 0$ elementwise, then $(\mathbf{I}-\mathbf{F})^{-1}$ exists and

$$(\mathbf{I}-\mathbf{F})^{-1} = \sum_{n=0}^{\infty} \mathbf{F}^n, \tag{4.3}$$

where the convergence takes place elementwise. Also observe that, by irreducibility, for any pair (i,j), $(\mathbf{F}^n)_{ij} > 0$ for some n as a function of (i,j). Hence, the right-hand side of (4.3) is a strictly positive matrix. Hence we conclude that $(\mathbf{I}-\mathbf{F})^{-1} > 0$. Hence we have established (c).

To prove that (c) implies (a), note that if $(\mathbf{I}-\mathbf{F})^{-1} > 0$, then $(\mathbf{I}-\mathbf{F})^{-1}\mathbf{b} > 0$, for any $\mathbf{b} > 0$. Let $\mathbf{p} = (\mathbf{I}-\mathbf{F})^{-1}\mathbf{b}$ and hence $\mathbf{p} > 0$. Thus, we have proved the equivalence of the three statements $(a),(b)$ and (c).

Exercise 4.7. 1. Given that $\left(c, \mathbf{t}^{(1)}\right)$ is feasible, we know that there exists a vector $\mathbf{p}^{(1)}$ of powers such that user k's ε_b/I_0 meets the target level of $\beta_k^{(1)}$. If $\beta_k^{(1)} \geq \beta_k^{(2)}$ then the same vector of powers, $\mathbf{p}^{(1)}$ will satisfy the threshold requirements for each mobile as the cell allocation remains unchanged. Hence $\left(c, \mathbf{t}^{(2)}\right)$ is also feasible.

2. The feasibility of $\left(c^{(3)}, \mathbf{t}^{(3)}\right)$ is evident from the definition of $\beta_k^{(3)}$. Now it remains to be shown that $\beta_k^{(3)} \geq \beta_k$. Let us consider a particular user r. Without loss of generality, assume that $p_r^{(1)*} \leq p_r^{(2)*}$. Then the new assignment will be $c_r^{(3)} = c_r^{(1)}$. Therefore,

$$\begin{aligned}
\beta_r^{(3)} &= \frac{g_{r,c_r^{(3)}} p_r^{(1)*}}{N_0 W + \sum_{k \neq r} g_{k,c_k^{(3)}} p_k^{(3)*}} \\
&= \frac{g_{r,c_r^{(3)}} p_r^{(1)*}}{N_0 W + \sum_{k \neq r \text{ and } k \in I_1} g_{k,c_k^{(3)}} p_k^{(3)*} + \sum_{k \neq r \text{ and } k \in I_2} g_{k,c_k^{(3)}} p_k^{(3)*}}
\end{aligned}$$

where $I_1 = \{k | p_k^{(1)*} \leq p_k^{(2)*}\}$ and $I_2 = \{k | p_k^{(2)*} \leq p_k^{(1)*}\}$. Therefore with the corresponding cell assignment we have,

$$\beta_r^{(3)} = \frac{g_{r,c_r^{(3)}} p_r^{(1)*}}{N_0 W + \sum_{k \neq r \text{ and } k \in I_1} g_{k,c_k^{(1)}} p_k^{(1)*} + \sum_{k \neq r \text{ and } k \in I_2} g_{k,c_k^{(2)}} p_k^{(2)*}}$$

On comparing the above expression for $\beta_r^{(3)}$ with β_r we see that the numerator remains the same in both the cases but the denominator in $\beta_r^{(3)}$ is reduced due to the new cell assignment. Hence it follows that $\beta_r^{(3)} \geq \beta_r$ for all mobiles $1 \leq r \leq K$

3. Given that uplink communication is feasible we know that there exists at least one vector of powers, $\mathbf{p}^{(1)*}$ such that the threshold requirements are made. If this is the only solution possible then its trivially the unique coordinate-wise minimum vector of power that allows successful communication. On the other hand if there is another vector of powers which is feasible, say $\mathbf{p}^{(2)*}$, then using the cell and power allotment in part (2) we get a new power vector $\mathbf{p}^{(3)*}$ each coordinate of which is less than or equal to the coordinates of $\mathbf{p}^{(2)*}$ and $\mathbf{p}^{(2)*}$. Now applying the algorithm in part (2) to $\mathbf{p}^{(3)*}$ and the other feasible vectors and proceeding in a similar fashion as above, we can conclude that there is a coordinate-wise minimum vector of powers.

EXERCISE 4.8. 1. (a) From the definitions, $g_{nl} \geq 0$ for all n and $l \in S_n$, and $I_{nl}^{(m)} \geq 0$ for all $n,l \in S_n$ if $p_n^{(m)} \geq 0$ for all n. Hence

$$p_n^{(m+1)} = \min_{l \in S_n} \frac{\beta_n I_{nl}^{(m)}}{G g_{nl}} \geq 0$$

Thus $I(\mathbf{p}) \geq 0$ for every $\mathbf{p} \geq 0$.

(b) Suppose $\mathbf{p} \geq \tilde{\mathbf{p}}$ (component-wise dominance). Then we have that

$$I_{nl}^{(m)} = \sum_{k \neq n} g_{kl} p_k^{(m)} + N_0 W \geq \sum_{k \neq n} g_{kl} \tilde{p}_k^{(m)} + N_0 W := \tilde{I}_{nl}^{(m)}$$

Hence

$$I(\mathbf{p}) = \min_{l \in S_n} \frac{\beta_n I_{nl}^{(m)}}{G g_{nl}} \geq \min_{l \in S_n} \frac{\beta_n \tilde{I}_{nl}^{(m)}}{G g_{nl}} = I(\tilde{\mathbf{p}}).$$

(c) Assume $\alpha > 1$

$$I(\alpha \mathbf{p}) = \min_{l \in S_n} \frac{\beta_n I_{nl}^{(m)}(\alpha \mathbf{p})}{G g_{nl}}$$

$$I_{nl}^{(m)}(\alpha \mathbf{p}) = \sum_{k \neq n} g_{kl} \alpha p_k^{(m)} + N_0 W < \sum_{k \neq n} g_{kl} \alpha p_k^{(m)} + \alpha N_0 W = \alpha I_{nl}^{(m)}(\mathbf{p})$$

Hence

$$I(\alpha \mathbf{p}) < \alpha I(\mathbf{p}).$$

2. Suppose I has a fixed point $\mathbf{p}_1$, i.e., $I(\mathbf{p}_1) = \mathbf{p}_1$. Assume that there exists another fixed point $\mathbf{p}_2 \geq 0$ such that $\mathbf{p}_1 \neq \mathbf{p}_2$. Then, without loss of generality, we can assume that there exists at least one $j \in \{1, \dots, K\}$ such that $p_1(j) < p_2(j)$. Now, observe that $\alpha := \max_{j \in \{1,\dots,K\}} \frac{p_2(j)}{p_1(j)} > 1$. Let $i = \arg\max_{j \in \{1,\dots,K\}} \frac{p_2(j)}{p_1(j)}$ and so $p_2(i) = \alpha p_1(i)$. Consequently $\alpha \mathbf{p}_1 \geq \mathbf{p}_2$. Hence, we have that

$$\alpha \mathbf{p}_1 = \alpha I(\mathbf{p}_1) > I(\alpha \mathbf{p}_1) \geq I(\mathbf{p}_2) = \mathbf{p}_2.$$

   But, by definition, α is such that $\alpha p_1(i) = p_2(i)$. Hence we have contradiction and it must be that $\mathbf{p}_1 = \mathbf{p}_2$.

3. Suppose $\mathbf{p}^*$ is the unique fixed point of I. Since $p^*(j) > 0$ for all j, given any initial $\mathbf{p}$, we can find $\alpha \geq 1$ such that $\alpha \mathbf{p}^* \geq \mathbf{p}$. Now, from question part (1), (c), $I(\alpha \mathbf{p}^*) < \alpha I(\mathbf{p}^*) = \alpha \mathbf{p}^*$. Now take $\mathbf{z}$ to be the all-zero vector. It is clear that $\mathbf{z} \leq \mathbf{p} \leq \mathbf{p}^*$. By part (*b*) of question 1, we know that $I^{(n)}(\mathbf{z}) \leq I^{(n)}(\mathbf{p}) \leq I^{(n)}(\alpha \mathbf{p}^*)$.

   Claim 1: $I^{(n)}(\alpha \mathbf{p}^*) \to \mathbf{p}^*$ as $n \to \infty$.

   Proof: From above we have $I(\alpha \mathbf{p}^*) < \alpha \mathbf{p}^*$. Assume that $I^{(n)}(\alpha \mathbf{p}^*) < I^{(n-1)}(\alpha \mathbf{p}^*)$. Then by part (*b*) of question 1, we have that $I(I^{(n)}(\alpha \mathbf{p}^*)) \leq I(I^{(n-1)}(\alpha \mathbf{p}^*))$, i.e., $I^{(n+1)}(\alpha \mathbf{p}^*) \leq I^{(n)}(\alpha \mathbf{p}^*)$. Hence $I^{(n)}(\alpha \mathbf{p}^*)$ is a decreasing sequence (component-wise). It is also bounded away from zero so it must converge to the unique fixed point $\mathbf{p}^*$.

   Claim 2: $I^{(n)}(\mathbf{z}) \to \mathbf{p}^*$ as $n \to \infty$.

   Proof: Clearly $\mathbf{z} < \mathbf{p}^*$ and $I(\mathbf{z}) \geq \mathbf{z}$. Suppose that $\mathbf{z} \leq I(\mathbf{z}) \leq \dots I^{(n)}(\mathbf{z}) \leq \mathbf{p}^*$. Then by part (*b*) of question 1, we have $\mathbf{p}^* = I(\mathbf{p}^*) \geq I(I^{(n)}(\mathbf{z})) \geq I(I^{(n-1)}(\mathbf{z})) = I^{(n)}(\mathbf{z})$. In other words we have shown that $\mathbf{p}^* \geq I^{(n+1)}(\mathbf{z}) \geq \mathbf{z}$. Therefore, the sequence $I^{(n)}(\mathbf{z})$ is nondecreasing and bounded above by $\mathbf{p}^*$ hence it must converge to $\mathbf{p}^*$.

   Combining Claim 1 and Claim 2 together, we get that $I^{(n)}(\mathbf{p}) \to \mathbf{p}^*$ for any initial vector $\mathbf{p}$.

4. Using the identical argument as in the proof of Claim 1, we observe that $I^{(n)}(\mathbf{p})$ is a decreasing sequence that converges to the fixed point $\mathbf{p}^*$ for any feasible vector $\mathbf{p}$. Hence it follows that $\mathbf{p} \geq \mathbf{p}^*$ for any feasible vector $\mathbf{p}$. In other words, the fixed point $\mathbf{p}^*$ is the solution to $\mathbf{p} \geq I(\mathbf{p})$ corresponding to the minimum total transmit power.

EXERCISE 4.9.

EXERCISE 4.10. 1. We assume that there is only 1 base station. The set of feasible power vectors is given by:

$$\mathcal{A} = \left\{ (P_1, P_2, \ldots, P_K) : \frac{GP_k g_{k,1}}{\sum_{j \neq k} P_j g_{j,1} + N_0 W} \geq \beta \wedge 0 \leq P_k \leq \hat{P}, k = 1, 2, \ldots, K \right\}$$

2. For K=2 we need:

$$\begin{array}{l} 0 \leq P_1 \leq \hat{P} \quad , \quad 0 \leq P_2 \leq \hat{P}, \\ \frac{GP_1 g_{1,1}}{P_2 g_{2,1} + N_0 W} \geq \beta \Rightarrow \frac{G}{\beta} \frac{g_{1,1}}{g_{2,1}} P_1 - \frac{\beta}{G} \frac{N_0 W}{g_{2,1}} \leq P_2 \\ \frac{GP_2 g_{2,1}}{P_1 g_{1,1} + N_0 W} \geq \beta \Rightarrow \frac{\beta}{G} \frac{g_{1,1}}{g_{2,1}} P_1 - \frac{N_0 W}{g_{2,1}} \geq P_2 \end{array}$$

For example let $\frac{N_0 W}{g_{2,1}} = 1$, $g_{1,1} = g_{2,1}$, $\hat{P} = 5$, for $\frac{G}{\beta} = 2$ the set $\mathcal{A}$ of feasible power vectors is shown in figure 4.1.

For this example the set of feasible power vectors is non-empty, with componentwise minimum solution $(P_1^*, P_2^*) = (1, 1)$. Clearly if $\hat{P} < 1$, $\mathcal{A} = \emptyset$. For a more interesting example of an outage situation, let $\frac{G}{\beta} = 1$ with the above choices for the other parameters. As shown in Figure 4.2 the feasible set of power vectors is empty.

3. For the parameters $g_{1,1} = g_{2,1} = 1$, $G = 2$, $N_0 W = 1$ and $\beta = 1$ we have that the component-wise minimum feasible power vector is $(P_1^*, P_2^*) = (1, 1)$. We plot in Figure 4.3 the evolution of P_1 and P_2 over time when the power control algorithm is run with a probability of error of 10^{-3} in the power control bit for each user.

4. We first redo part (1). Let K be the number of users, M the number of base stations, and $g_{k,m}$ the gain from user k to base station m. Then the set $\mathcal{A}$ of feasible power vectors to support a given $\mathcal{E}_b/N_0 = \beta$ requirement under soft handoff in the uplink can be expressed as:

$$\mathcal{A} = \{(P_1, P_2, \ldots, P_K) : \frac{GP_k g_{k,m}}{\sum_{\substack{j=1 \\ j \neq k}}^{N} P_j g_{j,m} + N_0 W} \geq \beta, \text{for some } m, 1 \leq m \leq M, \quad (4.4)$$

$$0 \leq P_k \leq \hat{P}, k = 1, 2, \ldots, K\}$$

which can be specialized for the case of 2 base stations by choosing $M = 2$.

We now redo part (2).

Let $A = \left\{ (P_1, P_2) : P_2 \leq \frac{1}{g_{2,1}} \left(\frac{GP_1 g_{1,1}}{\beta} - N_0 W \right) \right\}$, $B = \left\{ (P_1, P_2) : P_2 \leq \frac{1}{g_{2,2}} \left(\frac{GP_1 g_{1,2}}{\beta} - N_0 W \right) \right\}$, $C = \left\{ (P_1, P_2) : P_2 \geq \frac{\beta}{G g_{2,1}} (P_1 g_{1,1} + N_0 W) \right\}$, $D = \left\{ (P_1, P_2) : P_2 \geq \frac{\beta}{G g_{2,2}} (P_1 g_{1,2} + N_0 W) \right\}$,

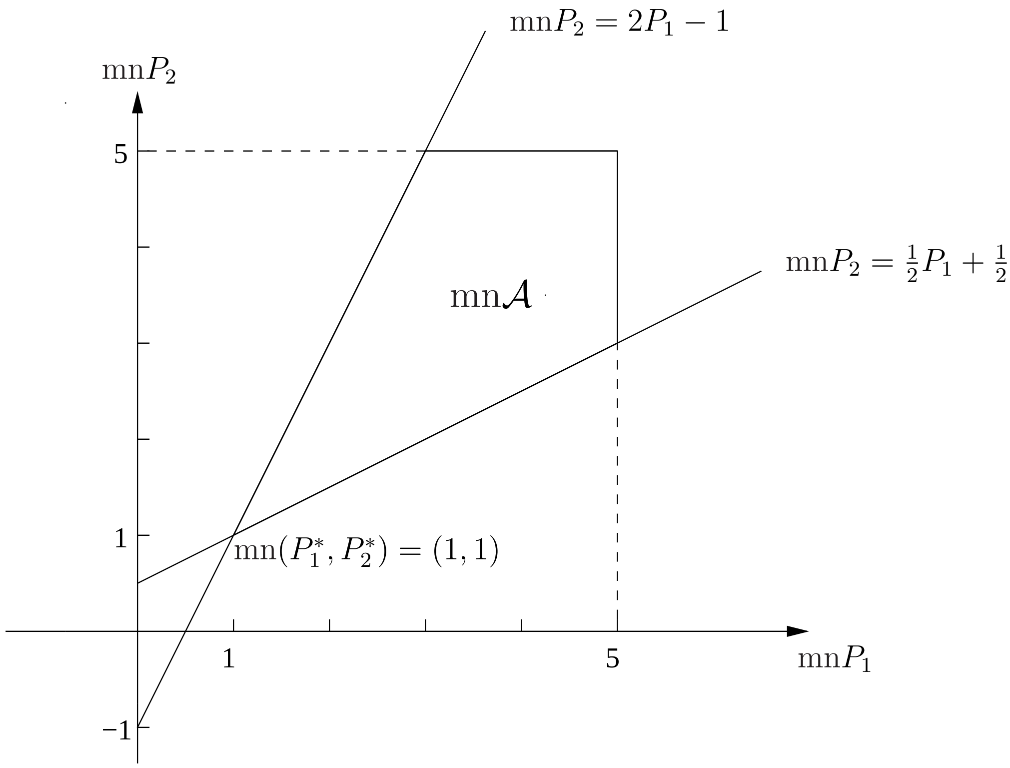


Figure 4.1: Set $\mathcal{A}$ of feasible power vectors for M=1 and K=2. The figure corresponds to $\frac{N_0 W}{g_{2,1}} = 1$, $g_{1,1} = g_{2,1}$, $\hat{P} = 5$, and $\frac{G}{\beta} = 2$. Here (P_1^*, P_2^*) is the componentwise minimum power vector in the feasible set.

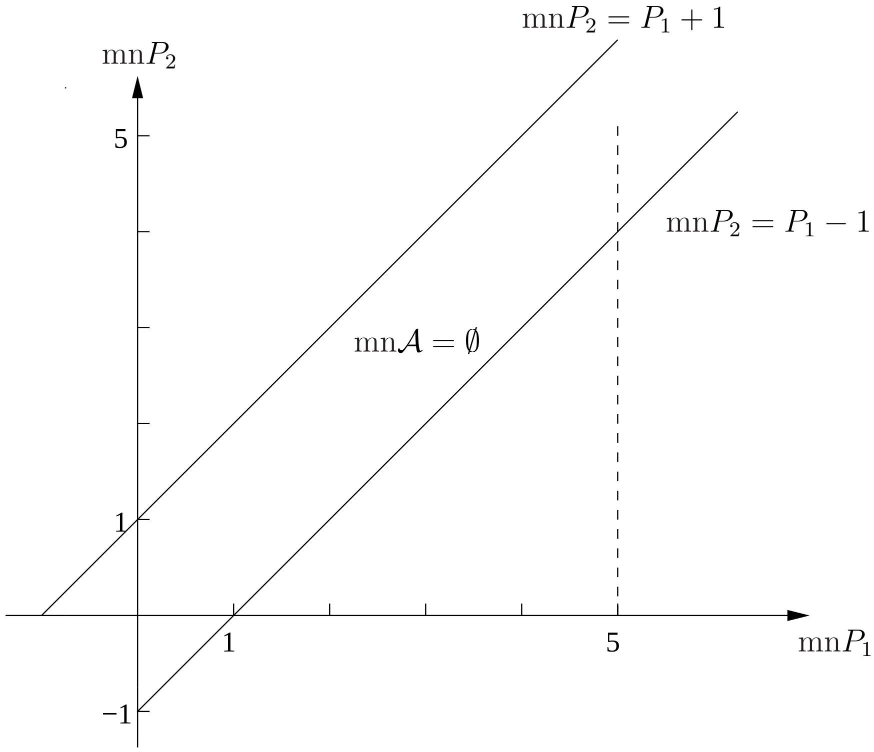


Figure 4.2: Empty set $\mathcal{A}$ of feasible power vectors for M=1 and K=2. The figure corresponds to $\frac{N_0 W}{g_{2,1}} = 1$, $g_{1,1} = g_{2,1}$, $\hat{P} = 5$, and $\frac{G}{\beta} = 1$.

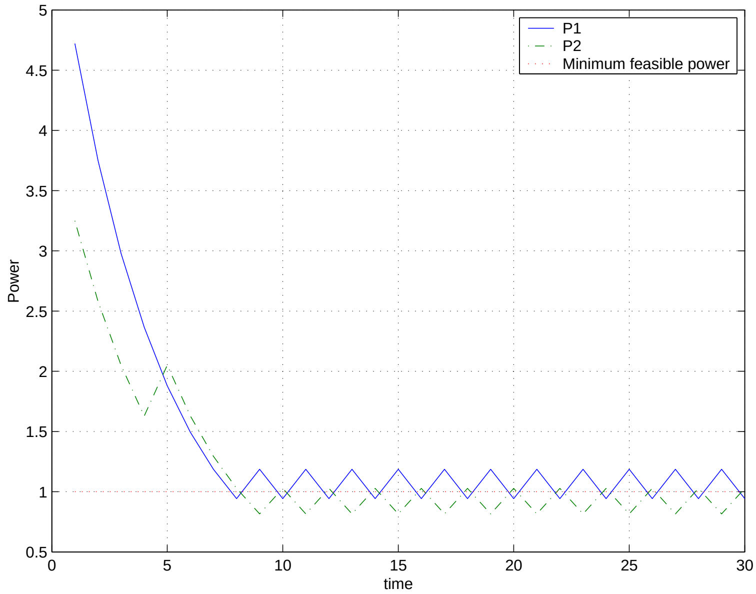


Figure 4.3: Trajectory of power updates over time for a power control bit error probability of 10^{-3}.

$E = \left\{(P_1, P_2) : 0 \leq P_1 \leq \hat{P}, 0 \leq P_2 \leq \hat{P}\right\}$. Then for $K = M = 2$ we can express $\mathcal{A}$ as $\mathcal{A} = ((A \cup B) \cap (C \cup D)) \cap E$. For the case of $G/\beta = 2$, $g_{1,1} = g_{1,2} = g_{2,1} = 1/4$, $g_{2,2} = 1/2$, $\hat{P} = 5$, and $N_0 W = 1$, the corresponding set of feasible power vectors is shown in Figure 4.4. We see that there is a componentwise minimum power vector $(P_1, P_2)^* = (20/7, 12/7)$ in the feasible set.

Also if $G/\beta < \min\left(\sqrt{\frac{g_{1,1}g_{2,2}}{g_{1,2}g_{2,1}}}, \sqrt{\frac{g_{1,2}g_{2,1}}{g_{1,1}g_{2,2}}}\right)$ then $\mathcal{A} = \emptyset$.

5. With the parameter choices of part (4) we plot in Figure 4.5 the evolution of P_1 and P_2 over time when the power control algorithm is run with a probability of error of 10^{-3} in the power control bit for each user.

6. This is not true. What determines the choice of base station is the signal to interference plus noise ratio (SINR), and not the channel gain. If the channel gain from a given user to a particular base station is large but that base station is supporting many users, then the SINR may be small, even smaller than β, in which case the given base station cannot be chosen. On the other hand there can be another base station for which the channel gain is not so large, but if the interference seen by this user in the base station is small, then the SINR can be large, possibly exceeding β, in which case the base station can be chosen to serve the user.

Exercise 4.11. 4. We analyze the CDMA system first. The cell model for this

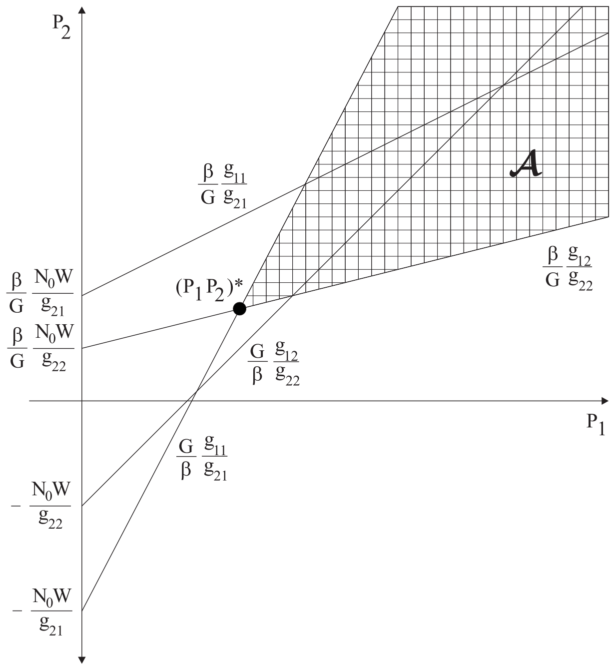


Figure 4.4: Set $\mathcal{A}$ of feasible power vectors for $M = K = 2$. The figure corresponds to $G/\beta = 2$, $g_{1,1} = g_{1,2} = g_{2,1} = 1/4$, $g_{2,2} = 1/2$, and $N_0W = 1$. Here $(P_1, P_2)^*$ is the componentwise minimum power vector in the feasible set.

problem is shown in figure 4.6.

Here it is assume that users are perfectly power controlled in their respective cells. Therefore the received power at each base station from the users in the cell is equal to some constant Q, independent of the user. Neglecting noise we have

$$\frac{GQ}{(k-1)Q + \sum_{k=1}^{K} Q(\frac{r_{2,k}}{r_{1,k}})^{\alpha}} \geq \beta \tag{4.5}$$

for any user of cell $\sharp 1$ that it is not in the outage. Then

$$G \geq \beta \left[k - 1 + \sum_{k=1}^{K} (\frac{r_{2,k}}{r_{1,k}})^{\alpha} \right]$$
$$\Rightarrow \sum_{k=1}^{K} (\frac{r_{2,k}}{r_{1,k}})^{\alpha} \leq \frac{G}{\beta} - (K-1)$$

If this condition is not met there is an outage. Then

$$P_{out} = Pr\{\sum_{k=1}^{K} (\frac{r_{2,k}}{r_{1,k}})^{\alpha} > \frac{G}{\beta} - (K-1)\}$$

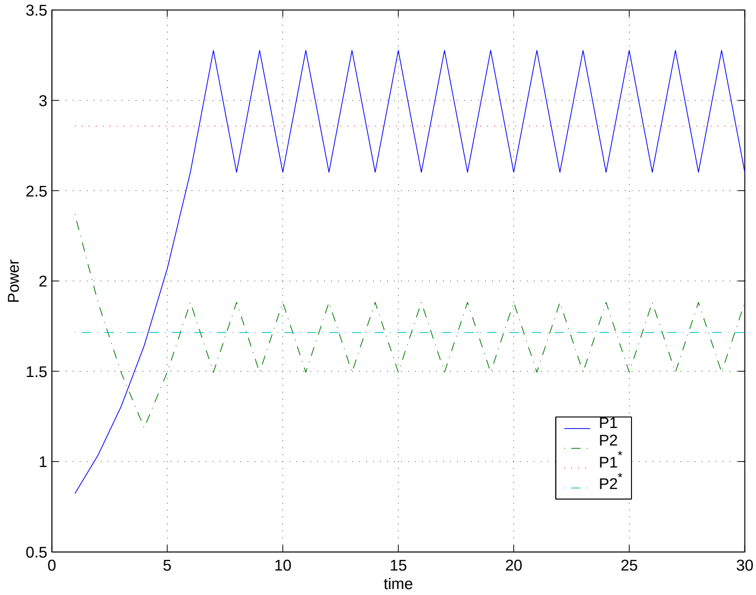


Figure 4.5: Trajectory of power updates over time for 2 base stations, for a power control bit error probability of 10^{-3}.

Let $X = \sum_{k=1}^{K} (\frac{r_{2,k}}{r_{1,k}})^{\alpha}$. For large K we approximate X by a Gaussian random variable with mean $\mu = E(X)$ and standard deviation $\sigma = \sqrt{var(X)}$. We first note that $\{\frac{r_{2,k}}{r_{1,k}}\}_k$ are iid. Then $\mu = KE\left[(\frac{r_{2,k}}{r_{1,k}})^{\alpha}\right]$ and $\sigma = \sqrt{Kvar\left((\frac{r_{2,k}}{r_{1,k}})^{\alpha}\right)}$. It is not possible to obtain a closed form expression for μ and σ for general α. We will approximate $r_{1,1}$ by $E(r_{1,1}) = d$ then

$$
\begin{aligned}
\mu &\approx KE\left[(\frac{r_{2,k}}{r_{1,k}})^{\alpha}\right] = \frac{2K}{d^{\alpha}} \int_0^{\frac{d}{2}} \frac{1}{d} r_{2,1}^{\alpha} dr_{2.1} \\
&= \frac{2K}{d^{\alpha}} \frac{1}{d} \left(\frac{d}{2}\right)^{\alpha+1} \frac{1}{\alpha+1} = \frac{K}{2^{\alpha}} \frac{1}{\alpha+1}
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma^2 &\approx \frac{K}{d^{2\alpha}} \left[2 \int_0^{\frac{d}{2}} \frac{1}{d} r_{2,1}^{2\alpha}\, dr_{2,1} - \left(\frac{d}{2}\right)^{2\alpha} \frac{1}{(\alpha+1)^2} \right] \\
&= \frac{K}{d^{2\alpha}} \left[2\frac{1}{d} \left(\frac{d}{2}\right)^{2\alpha+1} \frac{1}{2\alpha+1} - \left(\frac{d}{2}\right)^{2\alpha} \frac{1}{(\alpha+1)^2} \right] \\
&= \frac{K}{2^{2\alpha}} \left[\frac{1}{2\alpha+1} - \frac{1}{(\alpha+1)^2} \right]
\end{aligned}
$$

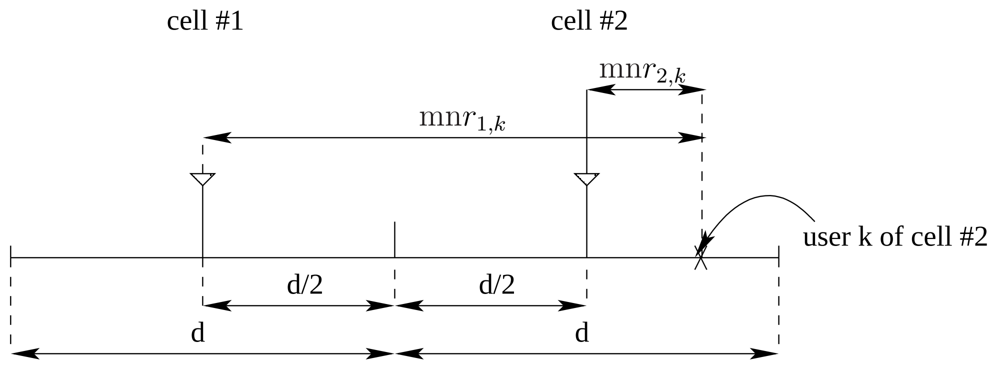


Figure 4.6: Cell model for Exercise 4.11

Then

$$P_{out} \approx Q\left\{\left[\frac{G}{\beta} - (K-1) - \mu\right]\frac{1}{\sigma}\right\}$$
$$\Rightarrow \left[\frac{G}{\beta} - (K-1) - \mu\right]\frac{1}{\sigma} \approx Q^{-1}(P_{out})$$

Replacing $G = \frac{W}{R}$, μ and σ we have

$$\left[\frac{W}{R\beta} - (K-1) - K\frac{2^{-\alpha}}{\alpha+1}\right]\frac{2^{\alpha}}{\sqrt{K}}\left[\frac{1}{2\alpha+1} - \frac{1}{(\alpha+1)^2}\right]^{-\frac{1}{2}} = Q^{-1}(P_{out})$$

or

$$\frac{W}{RK} = \left[Q^{-1}(P_{pout})\frac{1}{\sqrt{K}2^{\alpha}}\left[\frac{1}{2\alpha+1} - \frac{1}{(\alpha+1)^2}\right]^{\frac{1}{2}} + \frac{2^{-\alpha}}{\alpha+1} + 1 - \frac{1}{K}\right]\beta$$

Therefore the spectral efficiency $f = \frac{RK}{W}$ is given by

$$f = \frac{RK}{W} = \frac{1}{\beta}\left[Q^{-1}(P_{pout})\frac{1}{\sqrt{K}2^{\alpha}}\left[\frac{1}{2\alpha+1} - \frac{1}{(\alpha+1)^2}\right]^{\frac{1}{2}} + \frac{2^{-\alpha}}{\alpha+1} + 1 - \frac{1}{K}\right]^{-1} \quad (4.6)$$

Now as K and W go to ∞ we obtain

$$\lim_{K,W\to\infty} = \frac{1}{\beta}\left[1 + \frac{1}{2^{\alpha}(\alpha+1)}\right]^{-1}$$

Also as α increases, f increases to $\frac{1}{\beta}$.

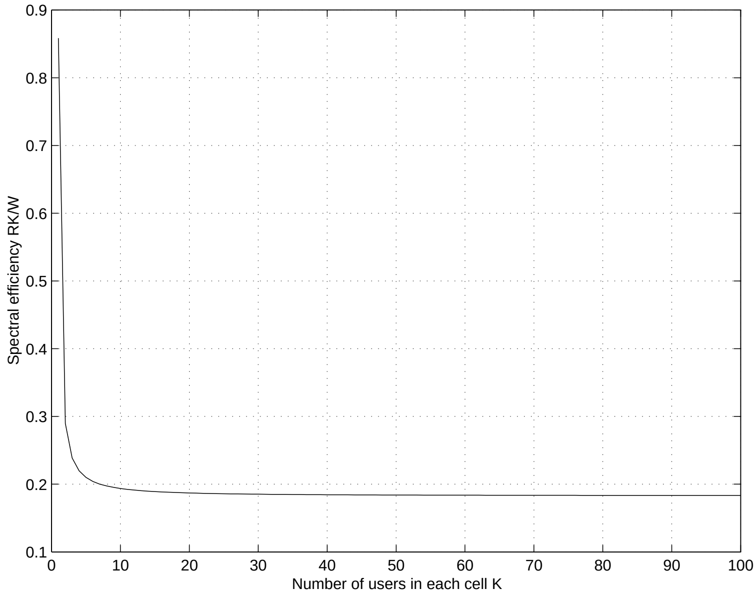


Figure 4.7: Spectral efficiency as a function of the number of users in each cell for $Q^{-1}(P_{out}) = 2$, $\alpha = 2$ and $\beta = 7$dB.

For the case that $\alpha = 2$ and $\beta = 7dB$ the resulting spectral efficiency is plotted in the figure 4.7 and we have

$$\lim_{K,W\to\infty} f(\alpha = 2, \beta = 7dB) = 0.1833$$

1. Now consider the orthogonal case. Since the users are orthogonal within the cell there is no intra-cell interference. Also we will assume that the users are power controlled at each base station so that the received power from all the users is the same. The out-of-cell interference is averaged over many OFDM symbols, so that each base station observes an interference that is the average over all the users of the neighboring cell. We will reuse most of the derivation of part (4). We have

$$\frac{GQ}{\sum_{k=1}^{K} Q\left(\frac{r_{2,k}}{r_{1,k}}\right)^{\alpha}} \geq \beta$$

for any user of cell 1 that is not in outage, where G is the processing gain, Q is the received power of a user to its base station, and $r_{i,k}$ is the distance of user k to base station i. Note that this is exactly the same expression found in part (4) without the term (K-1)Q in the denominator.

Defining μ and σ as before, we can reuse all the expressions found in part (4) eliminating the term $(K-1)$. We obtain

$$P_{out} \approx \left\{\left[\frac{G}{\beta} - \mu\right]\frac{1}{\sigma}\right\}$$

and

$$f = \frac{RK}{W} = \frac{1}{\beta}\left[\frac{Q^{-1}(P_{out})}{2^{\alpha}\sqrt{K}}\sqrt{\frac{1}{2\alpha+1} - \frac{1}{(1+\alpha)^2}} + \frac{1}{2^{\alpha}(\alpha+1)}\right]^{-1}$$

as the spectral efficiency.

2. As K and W go to ∞ we obtain:

$$\lim_{K,W\to\infty} f = \frac{2^{\alpha}(\alpha+1)}{\beta}$$

3. We see in Figure 4.8 that the spectral efficiency increases as the bandwidth and the number of users grow.

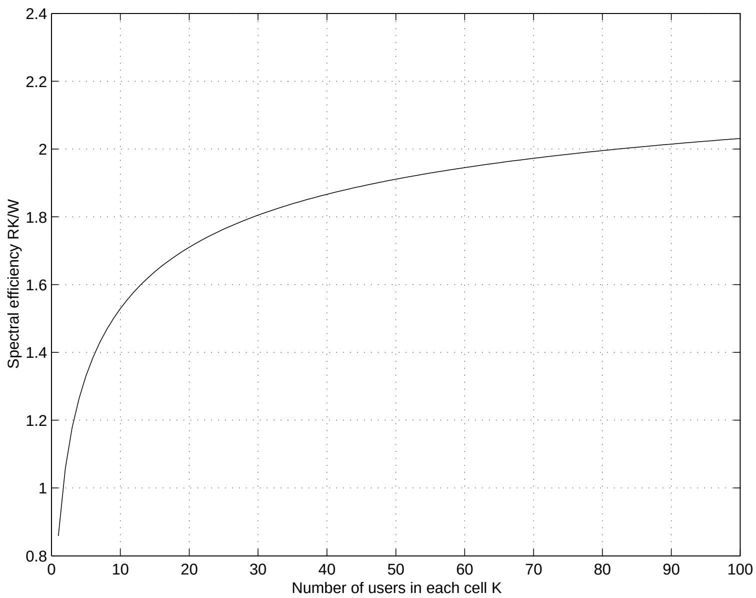


Figure 4.8: Spectral efficiency as a function of the number of users in each cell for $Q^{-1}(P_{out}) = 2$, $\alpha = 2$ and $\beta = 7$dB.

4. From the previous point we observe that removing the intra-cell interference made the spectral efficiency grow with the bandwidth and the number of users. The intra-cell interference contributes with a term $(K-1)Q$ to the total interference. When this interference is normalized by dividing it by the number of users we get $(1-1/K)Q$ which increases with K. We see that the intra-cell interference per user increases with K. For example for $K = 1$ there is no other user in the cell, and hence there is no intra-cell interference in the CDMA case. As K increases, the normalized intra-cell interference also increases reducing the spectral efficiency. This effect turns out to be more important than the interference averaging, dominating the dependence of the spectral efficiency with K and W as we observed in the figure in part (4).

EXERCISE 4.12. 1. The outage probability for user 1 is given by:

$$P_{out} = Pr\left[\frac{G\epsilon_1 P}{\sum_{j=2}^{N}\epsilon_j P + N_0 W} < \beta\right] = Pr\left[\frac{\epsilon_1}{\sum_{j=2}^{N}\epsilon_j + N_0 W/P} < \beta/G\right] \quad (4.7)$$

Since we don't have any power constraint we can let $P \to \infty$. Also for large N we can use the CLT to approximate $\sum_{j=2}^{N}\epsilon_j \sim N((N-1)\tilde{\mu}, (N-1)\tilde{\sigma}^2)$, where $\tilde{\mu} = E(\epsilon_j)$ and $\tilde{\sigma}^2 = Var(\epsilon_j)$.

The pdf of ϵ can be obtained by the transformation $\epsilon = e^X$ where $X \sim N(\mu, \sigma^2)$. It is given by:

$$f_\epsilon(\epsilon) = \frac{1}{\sqrt{2\pi}\sigma\epsilon} e^{-\frac{1}{2}\left(\frac{\log\epsilon - \mu}{\sigma}\right)^2} \quad (4.8)$$

for $\epsilon > 0$. Using this density we can compute $\tilde{\mu} = e^{\mu+\sigma^2/2}$ and $\tilde{\sigma}^2 = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$.

Therefore we can write:

$$\begin{aligned} P_{out} &= E\left\{P\left[\left.\frac{\epsilon_1}{\sum_{j=2}^{N}\epsilon_j} < \beta/G\right| \epsilon_1\right]\right\} \\ &= \int_0^\infty Q\left(\frac{\epsilon G/\beta - (N-1)\tilde{\mu}}{\sqrt{N-1}\tilde{\sigma}}\right)\frac{1}{\sqrt{2\pi}\sigma\epsilon} e^{-\frac{1}{2}\left(\frac{\log\epsilon-\mu}{\sigma}\right)^2} d\epsilon \quad (4.9) \end{aligned}$$

We would like to compute the spectral efficiency $\eta = NR/W = N/G$ as a function of the number of users N for a given outage probability P_{out}. For this we need to solve numerically the implicit function $G(N)$ defined by equation (4.9).

2. We show in Figure 4.9 a plot of the spectral efficiency as a function of the number of users, for the parameter choices $\beta = 7dB$, $\mu = 0$, $\sigma^2 = 0.053019 = 1/(10\log_{10} e)^2$ (which corresponds to a standard deviation of 1dB in ϵ). As N increases the spectral efficiency always decreases. There is an averaging effect but it is masked by the fact that $(N-1)/N$ increases with N.

3. In the other examples considered in the text the only randomness in the SINR was due to the interference, which was averaged out as N increased, converging to a constant. However in this problem the power control error for the given user remains random even for large N. There is an averaging of the interference, but this is not enough to make the SINR converge to a constant. This randomness in the SINR results in a degraded spectral efficiency. As a basis for comparison we have plotted in Figure 4.9 the spectral efficiency that results when the user of interest has perfect power control, but the interference has the same power control error ϵ considered before. In this case we see how the interference averaging effect

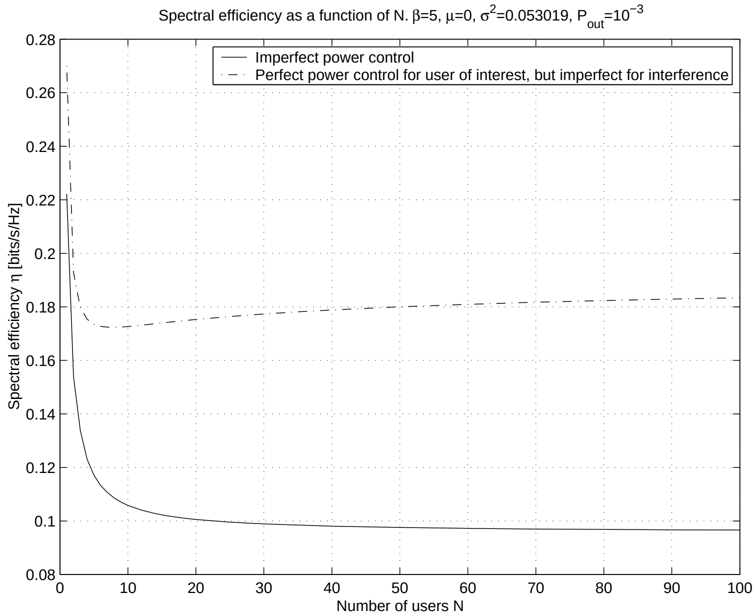


Figure 4.9: Spectral efficiency as a function of the number of users.

eventually takes over resulting in a larger spectral efficiency for large N. By comparing the two curves we see that imperfect power control produces a large performance degradation.

EXERCISE 4.13. 1. Let $h_i[l]$ be the channel's lth tap from base station i ($i = A, B$) to the user of interest. Assume for simplicity that both channels have L taps, and assume that the received signals from the two base station are chip and symbol synchronous. Assuming no ISI and using the notation of section 3.4.2 of the notes we can write the received signal vector as:

$$\mathbf{y} = \sum_{l=0}^{L-1} h_A[l]\mathbf{x}_{1,A}^{(l)} + \sum_{m=2}^{M}\sum_{l=0}^{L-1} h_A[l]\mathbf{x}_{m,A}^{(l)} + \sum_{l=0}^{L-1} h_B[l]\mathbf{x}_{1,B}^{(l)} + \sum_{m=2}^{N}\sum_{l=0}^{L-1} h_B[l]\mathbf{x}_{m,B}^{(l)} + \mathbf{w} \tag{4.10}$$

where M and N are the number of users in cells A and B respectively. Since the user is in soft handoff we can assume that the signals received from the two base stations have comparable power, and we can use a Gaussian approximation for the interference plus noise term. Letting $\tilde{\mathbf{w}} \sim CN(\mathbf{0}, (\sum_{m=2}^{M} \|\mathbf{h}_A\|^2\mathcal{E}_A^c + \sum_{m=2}^{N} \|\mathbf{h}_B\|^2\mathcal{E}_B^c + N_0)\mathbf{I}_{n+L-1})$, where $\mathcal{E}_i^c$ is the chip energy of the transmitted signal from base station i, and n is the symbol length. Then we can write:

$$\mathbf{y} = \sum_{l=0}^{L-1} h_A[l]\mathbf{x}_{1,A}^{(l)} + \sum_{l=0}^{L-1} h_B[l]\mathbf{x}_{1,B}^{(l)} + \tilde{\mathbf{w}} \tag{4.11}$$

The received signal without the noise lies in the span of the vectors $\left\{ \frac{\mathbf{u}_A^{(l)}}{\|\mathbf{u}_A\|}, \frac{\mathbf{u}_B^{(l)}}{\|\mathbf{u}_B\|} \right\}_{l=1}^{L-1}$. Assuming that the spreading sequences are orthogonal, and that their shifts are also orthogonal, we have that the previous set of vectors is an orthonormal set, and we can project onto these vectors to obtain $2L$ sufficient statistics:

$$r_i^{(l)} = h_i[l]\|\mathbf{u}_i\|x + w_i^{(l)} \tag{4.12}$$

where $i = A, B$ and $l = 0, 1, \ldots, L-1$. We can further project onto the direction of

$[(\|\mathbf{u}_A\|h_A[0])(\|\mathbf{u}_A\|h_A[1])\cdots(\|\mathbf{u}_A\|h_A[L-1])(\|\mathbf{u}_B\|h_B[0])(\|\mathbf{u}_B\|h_B[1])\cdots(\|\mathbf{u}_B\|h_B[L-1])]^T$ to obtain the sufficient statistic:

$$r = \sqrt{(\|\mathbf{h}_A\|^2\|\mathbf{u}_A\|^2 + \|\mathbf{h}_B\|^2\|\mathbf{u}_B\|^2)}x + w \tag{4.13}$$

where $w \sim CN(0, (\sum_{m=2}^{M} \|\mathbf{h}_A\|^2\mathcal{E}_A^c + \sum_{m=2}^{N} \|\mathbf{h}_B\|^2\mathcal{E}_B^c + N_0))$. Letting $\|\mathbf{u}_i\|^2 = G\mathcal{E}_i^c$, $i = A, B$ and assuming $x \in \{-1, 1\}$ we can write the error probability as:

$$P_e = Q\left(\sqrt{\frac{2G(\|\mathbf{h}_A\|^2\mathcal{E}_A^c + \|\mathbf{h}_B\|^2\mathcal{E}_B^c)}{\sum_{m=2}^{M} \|\mathbf{h}_A\|^2\mathcal{E}_A^c + \sum_{m=2}^{N} \|\mathbf{h}_B\|^2\mathcal{E}_B^c + N_0}}\right) \tag{4.14}$$

2. Let N be the total number of base stations, K the total number of users, S the set of all base stations, S_k the active set of user k (i.e. the set of base stations with which user k is in soft handoff), A_i the set of users that have base station i in their active sets, $g_{k,i}$ the channel gain from base station i to user k, and $P_{k,i}$ the power of the signal transmitted to user k from base station i. Then the SINR seen by user k is given by:

$$\mathsf{SINR}_k = \frac{G\sum_{i\in S_k} g_{k,i}P_{k,i}}{\sum_{i\in S_k} g_{k,i}\sum_{j\in A_i, j\neq k} P_{j,i} + \sum_{i\in S\setminus S_k} g_{k,i}\sum_{j\in A_i} P_{j,i} + N_0 W} \tag{4.15}$$

Assuming that there is a minimum SINR requirement β for reliable communication, the set of feasible power vectors is given by:

$$\mathcal{A} = \{(P_{1,1}P_{1,2}\cdots P_{1,N}P_{2,1}\cdots P_{2,N}\cdots P_{K,1}\cdots P_{K,N}) : \mathsf{SINR}_k \geq \beta, P_{k,i} = 0 \text{ if } i \notin S_k, k = 1, 2, \ldots \tag{4.16}$$

The power control problem consists of finding a power vector in $\mathcal{A}$.

Exercise 4.14. 1. The 2 latin squares have entries $R_{i,j}^a = (ai + j) \mod N$ and $R_{i,j}^b = (bi + j) \mod N$, where N is prime and $a \neq b$.

Consider the pair of ordered pairs:

$$(k_a, k_b)_{i,j} = ((ai + j) \mod N, (bi + j) \mod N)$$

$$(k_a, k_b)_{l,m} = ((al+m) \mod N, (bl+m) \mod N)$$

We want to show $(i,j) \neq (l,m) \Rightarrow (k_a,k_b)_{i,j} \neq (k_a,k_b)_{l,m}$, that is, any 2 ordered pairs must be different. Let $(d_1,d_2) = (k_a,k_b)_{i,j} - (k_a,k_b)_{l,m}$. We have to show that $(d_1,d_2) \neq (0,0)$.

$$(d_1,d_2) = ([a(i-l)+(j-m)] \mod N, [b(i-l)+(j-m)] \mod N) \quad (4.17)$$

If $i=l$, then we must have $j \neq m$ and $(d_1,d_2) = ((j-m) \mod N, (j-m) \mod N)$. Since $-(N-1) \leq (j-m) \leq (N-1)$ and $(j-m) \neq 0$ it follows that $(j-m) \mod N \neq 0$, therefore $(d_1,d_2) \neq (0,0)$.

If on the other hand $i \neq l$ then $d_1 - d_2 = (a-b)(i-l) \mod N$. $d_1 - d_2 = 0$ requires that N divides $(a-b)$ or $(i-l)$ (note that here we use the fact that N is prime). But both are in $[-(N-1),1] \cup [1,(N-1)]$ so they are not divisible by N. It follows that $d_1 - d_2 \neq 0$ and hence $(d_1,d_2) \neq (0,0)$.

2. Adapted from J. van Lint, R. Wilson, "A course in Combinatorics," Second Ed., Cambridge University Press, 2001.

   Consider a set of M mutually orthogonal latin squares. The entries in each latin square correspond to virtual channel numbers. We are free to rename the channels so that the first row of each latin square is $(1,2,\ldots,N)$. Then the pairs $(k_l,k_m)_{1,j}$ for any pair of matrices (l,m) ($l \neq m$, $l,m \in [1,\ldots,M]$) are (j,j). Now consider the $(2,1)$ entry of each latin square. It can't be 1 because 1 already appears in the position $(1,1)$. Also these elements must be different in all the matrices, because the pairs (k_l,k_m) with repeated entries have already appeared. Thus we have $M \leq (N-1)$. Note that N need not be prime for this result to hold.

Exercise 4.15. 1. Let $M = N/n$.

$$\begin{aligned}
\bar{P} &= \frac{1}{T}\int_0^T s(t)^2 dt = \frac{1}{T}\int_0^T \frac{1}{2N}\left[\sum_{i=0}^{n-1}\left(D[i]e^{j2\pi(f_c+iM/T)t} + D[i]^* e^{-j2\pi(f_c+iM/T)t}\right)\right]^2 dt \\
&= \frac{1}{2N}\sum_{i=0}^{n-1}\sum_{k=0}^{n-1}\frac{1}{T}\left\{\int_0^T D[i]D[k]e^{j2\pi(2f_c+(i+k)M/T)t}dt + \int_0^T D[i]D[k]^* e^{j2\pi((i-k)M/T)t}dt\right. \\
&+ \left.\int_0^T D[i]^* D[k]e^{j2\pi((-i+k)M/T)t}dt + \int_0^T D[i]^* D[k]^* e^{-j2\pi(2f_c+(i+k)M/T)t}dt\right\} \quad (4.18)
\end{aligned}$$

The magnitudes of the first and last integrals of each term can be shown to be upper bounded by $T|D[i]|^2(2\pi\zeta)^{-1}$, so when divided by T they are negligible.

The other 2 integrals evaluate to $T|D[i]|^2\delta_{i,k}$. Therefore we can write:

$$\bar{P} = \frac{1}{2N}\sum_{i=0}^{n-1}|D[i]|^2 \tag{4.19}$$

When the symbols are chosen from an equal energy constellation with $|D[i]| = 1$ we get $P_{av} = \frac{n}{2N}$.

2.

$$d[i] = \frac{1}{\sqrt{N}}\sum_{k=0}^{n-1} D[kN/n]e^{j\frac{2\pi}{N}\frac{ikN}{n}} \tag{4.20}$$

The symbols $D[i]$ are chosen uniformly on the unit circle, so their distribution is invariant to rotations in the complex plane. Therefore their distribution is circularly symmetric, and so is the distribution of their sum. Also the rotations induced by the complex exponential in the IDFT do not change the resulting distribution, so $d[0], \dots, d[N-1]$ are identically distributed.

As $N \to \infty$ with the ratio $n/N = \alpha$ kept constant, we can apply a version of the CLT for circularly symmetric random variables to conclude that the distribution of $d[i]$ converges to a circularly symmetric complex Gaussian distribution. Since $E[|D[k]|^2] = 1$ we get that $d[i] \sim CN(0, \alpha)$ for large N.

c) Since $|d[0]|^2/\alpha \sim Exp(1)$ and $P_{av}/\alpha = 1/2$ we can write

$$Pr\left[\frac{|d[0]|^2}{P_{av}} < \theta(\eta)\right] = Pr\left[\frac{|d[0]|^2}{\alpha} < \theta(\eta)P_{av}/\alpha\right] = 1 - e^{-\theta(\eta)/2} = 1 - \eta \tag{4.21}$$

Thus $\theta(\eta) = -2\log_e \eta$. For the special case of $\eta = 0.05$ we obtain $\theta(0.05) = 5.99$.

EXERCISE 4.16.

Chapter 5

Solutions to Exercises

EXERCISE 5.1. See handwritten solutions in file s04.h6sol_1.pdf (ex 3, part b))

EXERCISE 5.2. The received SNR at the base-station of the user at the edge of the cell (at distance d from it) is given by:

$$\mathsf{SNR} = \frac{P}{N_0 W d^\alpha},$$

where it is assumed that the bandwidth allocated to the user is W Hz, $\alpha > 2$ is the path-loss exponent and N_0 is noise variance. Using a reuse ratio of $0 < \rho \leq 1$, the closest base-station reusing the same frequency as the given base-station is at distance $2d/\rho$. Since we are dealing with a linear arrangement, there are only 2 such base-stations (the interference due to the others are significantly smaller and are going to be ignored). Thus, the received interference is given by

$$I = 2P\left(\frac{\rho}{d}\right)^\alpha,$$

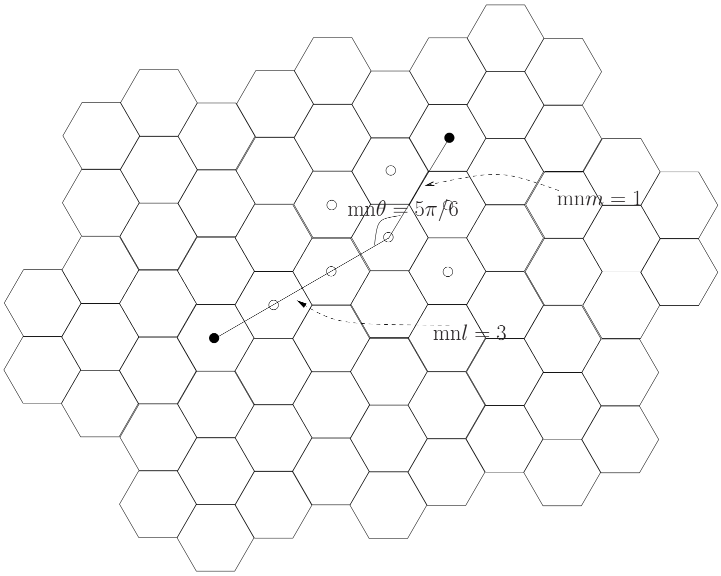
hence the received SINR is:

$$\mathsf{SINR} = \frac{\frac{P}{d^\alpha}}{\rho N_0 W + 2\frac{P}{\left(\frac{2d}{\rho}\right)^\alpha}} = \frac{\mathsf{SNR}}{\rho + 2\left(\frac{\rho}{2}\right)^\alpha \mathsf{SNR}},$$

and so $f_\rho := 2\left(\frac{\rho}{2}\right)^\alpha$.

Solution 5.2:

First, we observe the following simple method of evaluating the distance between



(centers of) any two cells in a hexagonal packing of the plane:

$$r(l, m, \theta) = 2d\sqrt{3m^2 + l^2 + 2\sqrt{3}lm\cos\theta},$$

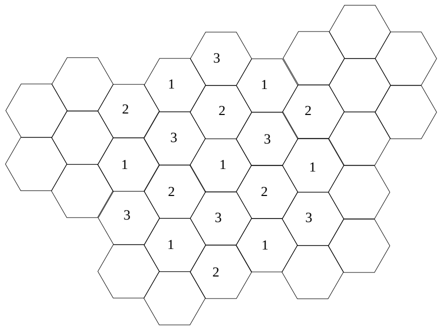
where $2d$ is the distance between the centers of two adjecent cells and the triple (l, m, θ) uniquely specifies the relative positions of the two cells w.r.t. one another in the following way (also see the diagram above): To get from one cell to another, we go m steps along the "offset axis" (the one that runs through the vertices of the cells) for a distance of $2\sqrt{3}dm$, and l steps along the "principal axes" (the one that bisects the sides of the cell) for a distance of $2ld$. The angle between the two axes is given by θ.

In the diagrams below, we illustrate optimal reuse patterns for $\rho = \frac{1}{3}, \frac{1}{4}, \frac{1}{7}, \frac{1}{9}$. For these reuse ratios we can see that identical distances separate the cells using the same frequency band, for all frequency bands. This is not true in general and only holds for specific reuse ratios (we only show up to $\rho = \frac{1}{9}$, but there are other ones like $\rho = \frac{1}{16}$, etc. For instance:

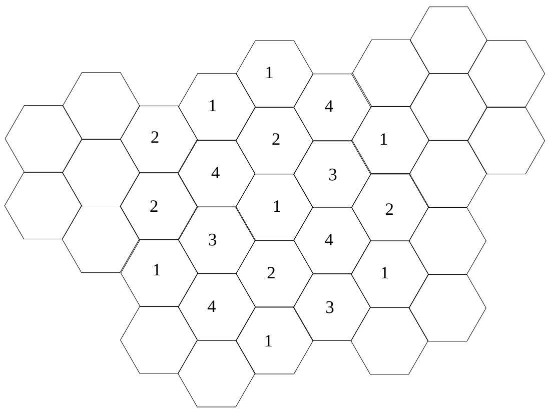
1. For $\rho = \frac{1}{3}$, there are 6 nearest cells at distance $2d\sqrt{3}$ interfering. Hence, the received SINR at the base-station due to the user at the edge of the cell (at distance d, i.e., on the side of the cell not on the vertex) is given by

$$\mathsf{SINR} = \frac{\frac{P}{2^\alpha}}{\frac{N_0 W}{3} + 6\frac{P}{(2d\sqrt{3})^\alpha}} = \frac{\mathsf{SNR}}{\rho + \frac{6}{(2\sqrt{3})^\alpha}\mathsf{SNR}},$$

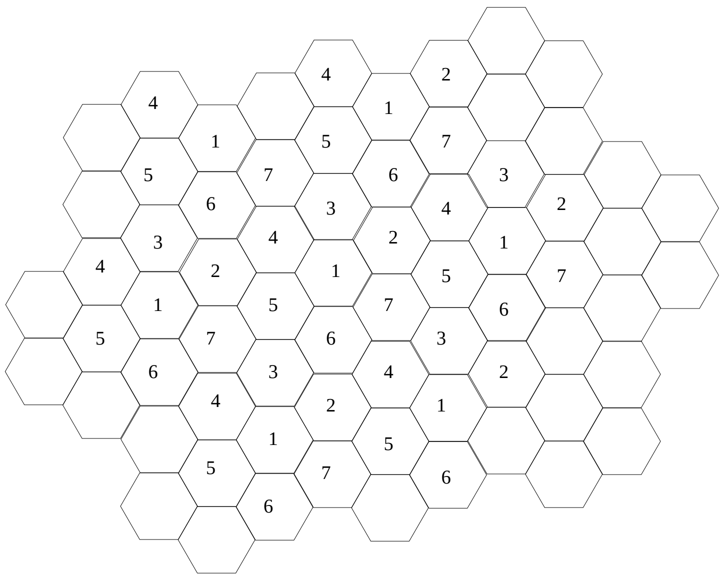
hence, we can write $f_\rho = \frac{6}{(2\sqrt{3})^\alpha}$,



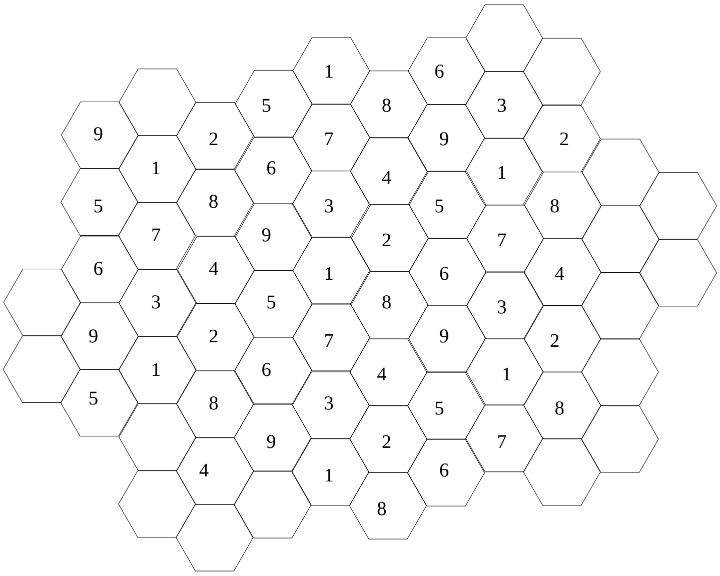
2. For $\rho = \frac{1}{4}$, there are 6 cells at distance $2d\sqrt{4}$ and so $f_\rho = \frac{6}{(2\sqrt{4})^\alpha}$,



3. For $\rho = \frac{1}{7}$, there are 6 cells at distance $2d\sqrt{7}$ and so $f_\rho = \frac{6}{(2\sqrt{7})^\alpha}$,



4. For $\rho = \frac{1}{9}$, there are 6 cells at distance $2d\sqrt{9}$ and so $f_\rho = \frac{6}{(2\sqrt{9})^\alpha}$.



Thus we see that, for these reuse ratios, the approximation $f_\rho = \frac{6}{(2\sqrt{\frac{1}{\rho}})^\alpha}$ is a good one.

In the plot below, we show the high-SNR approximation for the rate:

$$\rho W \log_2(1 + \frac{1}{f_p}),$$

for $\alpha = 2, 4, 6$ in the hexagonal packing of the plane. Note that the universal reuse ratio $\rho = 1$ yields the largest rate.

Exercise 5.3.

Exercise 5.4.

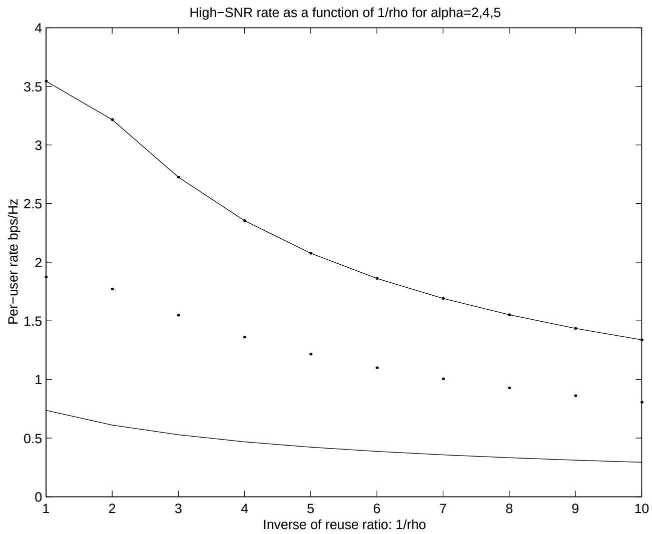
Exercise 5.5. In the figure below, we see that the optimal reuse ratio is $\rho = 1/2$ when studying the high-SNR approximation for the per-user rate in a linear network:

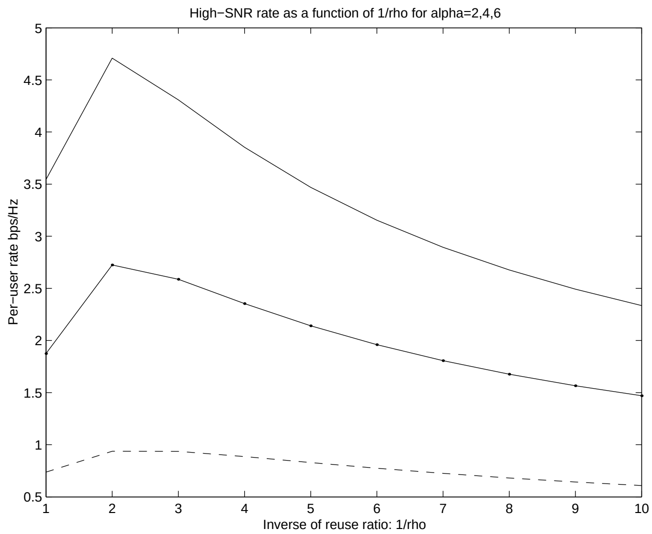
$$\rho W \log_2(1 + \frac{1}{f_p}),$$

where $f_p = 2(\frac{\rho}{2})^\alpha$. The plots were generated using $\alpha = 2, 4, 6$ as sample values.

Exercise 5.6. 1. This strategy achieves a rate:

$$R = \alpha \log\left(1 + \frac{P_1}{N_0}\right) + (1 - \alpha) \log\left(1 + \frac{P_2}{N_0}\right) \tag{5.1}$$





where $\alpha P_1 + (1-\alpha)P_2 = P$. By Jensen's inequality and the concavity of $\log(\cdot)$ we have:

$$\begin{aligned} & \alpha \log\left(1+\frac{P_1}{N_0}\right) + (1-\alpha)\log\left(1+\frac{P_2}{N_0}\right) \\ \leq \; & \log\left(1+\frac{\alpha P_1 + (1-\alpha)P_2}{N_0}\right) = \log\left(1+\frac{P}{N_0}\right) = C_{AWGN} \end{aligned}$$

If $P_1 \neq P_2$ we have a strict inequality and it follows that this strategy is suboptimal.

2. As in a) assume we use a strategy of transmitting with power constraint P_1 a fraction α of the time, and with power constraint P_2 the remaining time. Also let

$P = \alpha P_1 + (1-\alpha)P_2$ be the total power constraint. Since this is just a particular strategy that satisfies the power constraint, the achievable rate cannot exceed the capacity of the channel, which is the supremum of all achievable rates for strategies that satisfy the power constraint. Thus we can write:

$$R = \alpha C(P_1) + (1-\alpha)C(P_2) \le C(P) = C(\alpha P_1 + (1-\alpha)P_2) \tag{5.2}$$

valid for any P_1 and P_2 and $\alpha \in [0,1]$. Therefore $C(P)$ is a concave function of P.

Exercise 5.7. For QAM with 2^k points, the average error probability can be expressed as (for some constant α):

$$P_e = \alpha Q\left(\sqrt{\frac{2a^2}{N_0}}\right), \tag{5.3}$$

where the distance between two consecutive points on each of the axes is $2a$ and N_0 is the background noise. This expression is obtained by applying the union bound to the pairwise errors between the nearest neighbors to some central point. However, from Exercise 3.4, we know that the average SNR per symbol is given by $\mathrm{SNR} = \frac{E_{av}}{N_0}$, where

$$E_{av} = \frac{2a^2}{3}(2^k - 1). \tag{5.4}$$

Hence, we have that

$$P_e = \alpha Q\left(\sqrt{\frac{3\ \mathrm{SNR}}{2^k - 1}}\right) \approx \alpha \exp\left(-\frac{3\ \mathrm{SNR}}{2(2^k-1)}\right), \tag{5.5}$$

where the last inequality holds by the high-SNR approximation of $Q(\mathrm{SNR}) \approx \exp(-\mathrm{SNR}^2/2)$. Thus we have that the number of bits

$$k \approx \log_2\left(\frac{3\ \mathrm{SNR}}{2\ln(\alpha/P_e)}\right) = \log_2 \mathrm{SNR} + \text{constant}, \tag{5.6}$$

and the rate of QAM has the optimal order of growth with SNR on the AWGN channel.

Exercise 5.8.

Exercise 5.9.

Exercise 5.10. First, we compute the high-SNR approximation of the mean of $\log(1+ |h|^2\ \mathsf{SNR})$:

$$\mu \ := \ \mathbb{E}[\log(1+|h|^2\mathsf{SNR})] = \int_0^\infty \log(1+x\mathsf{SNR})f(x)dx,$$

$$
\begin{aligned}
&= \int_0^{\frac{1}{\mathsf{SNR}}} \log(1 + x\mathsf{SNR}) f(x)dx + \int_{\frac{1}{\mathsf{SNR}}}^{\infty} \log(1 + x\mathsf{SNR}) f(x)dx, \\
&\approx \mathbb{P}\left(|h|^2 \le \frac{1}{\mathsf{SNR}}\right) + \log \mathsf{SNR} \int_{\frac{1}{\mathsf{SNR}}}^{\infty} f(x)dx + \int_{\frac{1}{\mathsf{SNR}}}^{\infty} f(x) \log x dx.
\end{aligned}
$$

Where in the last line we have used the high-SNR approximation $\log(1 + \mathsf{SNR}) \approx \log \mathsf{SNR}$. Observe that, in the high-SNR regime, $\mathbb{P}\left(|h|^2 \le \frac{1}{\mathsf{SNR}}\right) \approx \frac{1}{SNR}$ and the last two terms are approximately $\log \mathsf{SNR}$ and $\mathbb{E}[\log |h|^2]$, respectively. So we have that

$$
\mu \approx \frac{1}{\mathsf{SNR}} + \log \mathsf{SNR} + \mathbb{E}[\log |h|^2] \approx \log \mathsf{SNR} + \mathbb{E}[\log |h|^2].
$$

Similarly we can define $\sigma^2 := \mathbb{E}[\log^2(1 + |h|^2\mathsf{SNR})]$ and use the same method to obtain the following high-SNR approximation:

$$
\sigma^2 \approx \log^2 \mathsf{SNR} + \log \mathsf{SNR}\mathbb{E}[\log |h|^2] + \mathbb{E}[\log^2 |h|^2].
$$

Finally, the standard deviation is defined as STD $:= \sqrt{\sigma^2 - \mu^2}$ and its high-SNR approximation is computed using the above expressions:

$$
\text{STD} \approx \sqrt{\mathbb{E}[\log^2 |h|^2] - \mathbb{E}[\log |h|^2]^2},
$$

which is a constant as a function of SNR. Hence $\frac{\text{STD}}{\mu}$ goes to zero as SNR increases.

On the other hand, in the low-SNR regime we use the approximation $\log(1+\mathsf{SNR}) \approx \mathsf{SNR} \log_2 e$ to get

$$
\begin{aligned}
\mu &\approx \mathsf{SNR}\mathbb{E}[|h|^2] \log_2 e, \\
\sigma^2 &\approx \mathsf{SNR}^2\mathbb{E}[|h|^4] \log_2^2 e,
\end{aligned}
$$

Hence, at low-SNR, we have that,

$$
\frac{\text{STD}}{\mu} \approx \frac{\sqrt{\mathbb{E}[|h|^4] - \mathbb{E}[|h|^2]^2}}{\mathbb{E}[|h|^2]} = \text{constant}.
$$

This makes sense because at high-SNR, the capacity formula is degree-of-freedom limited and changes in $|h|^2$ have a diminishing marginal effect, whereas in the low-SNR regime, the capacity formula is very sensitive to changes in the overall received SNR and hence even the smallest changes in $|h|^2$ affect the performance.

Exercise 5.11. The received SNR is given by $\frac{\|\mathbf{h}^\dagger \mathbf{x}\|^2}{N_0}$. Hence we need to maximize this quantity over all $\mathbf{x} \in \mathbb{C}^L$ such that $\|\mathbf{x}\|^2 \le P$ for fixed N_0 and some constant $P > 0$. By the Cauchy-Schwarz inequality, we have that

$$
\|\mathbf{h}^\dagger \mathbf{x}\|^2 \le \|\mathbf{h}^\dagger\|^2 \|\mathbf{x}\|^2,
$$

with equality if and only if $\mathbf{x} = \alpha \mathbf{h}$ for some constant scalar α. Hence, the optimal choice is

$$\mathbf{x} = \frac{P\mathbf{h}}{\|\mathbf{h}\|^2},$$

which is exactly the transmit beam-forming strategy.

EXERCISE 5.12.

EXERCISE 5.13. 1. Let $\mathbf{1} = [1, 1, \cdots, 1]^T \in \mathcal{R}^L$. Then the channel equation is:

$$\mathbf{y} = \mathbf{1}x + \mathbf{z} \tag{5.7}$$

where $\mathbf{z} \sim CN(0, N_0 \mathbf{I}_L)$ and x must satisfy the power constraint $E[x] \leq P$.

We note that we can project the received signal onto the direction of $\mathbf{1}$ obtaining the sufficient statistic:

$$r = \frac{\mathbf{1}^*}{\sqrt{L}}\mathbf{y} = \sqrt{L}x + \tilde{z} \tag{5.8}$$

where $\tilde{z} \sim CN(0, N_0)$. Defining $\tilde{x} = \sqrt{L}x$ we see that we have an AWGN channel with power constraint LP and noise variance N_0. Therefore $C = \log\left(1 + \frac{LP}{N_0}\right)$. We see that there is a power gain of L with respect to the single receive antenna system.

2. Let $\mathbf{h} = [h_1, h_2, \cdots, h_L]^T \in \mathcal{C}^L$. Then the channel equation is:

$$\mathbf{y} = \mathbf{h}x + \mathbf{z} \tag{5.9}$$

where $\mathbf{z} \sim CN(0, N_0 \mathbf{I}_L)$, $\mathbf{h}$ is known at the receiver and x must satisfy the power constraint $E[x] \leq P$.

Since the receiver knows the channel, it can project the received signal onto the direction of $\mathbf{h}$ obtaining the sufficient statistic:

$$r = \frac{\mathbf{h}^*}{\|\mathbf{h}\|}\mathbf{y} = \|\mathbf{h}\|x + \tilde{z} \tag{5.10}$$

where $\tilde{z} \sim CN(0, N_0)$. Then the problem reduces to computing the capacity of a scalar fading channel, with fading coefficient given by $\|\mathbf{h}\|$. It follows that:

$$C = E\left[\log\left(1 + \frac{\|\mathbf{h}\|^2 P}{N_0}\right)\right] = E\left[\log\left(1 + \frac{LP}{N_0}\frac{\|\mathbf{h}\|^2}{L}\right)\right] \tag{5.11}$$

In contrast, the single receive antenna system has a capacity $C = E\left[\log\left(1 + \frac{|h|^2 P}{N_0}\right)\right]$. The capacity is increased by having multiple receive antennas for two reasons:

first there is a power gain L, and second $\frac{\|\mathbf{h}\|^2}{L}$ has the same mean but less variance than $|h|^2$, and we get a diversity gain. Note that $Var\left[\frac{\|\mathbf{h}\|^2}{L}\right] = 1/L$ whereas $Var\left[|h|^2\right] = 1$.

As $L \to \infty$, $\frac{\|\mathbf{h}\|^2}{L} \to_{a.s.} 1$, so it follows that $C \approx \log\left(1 + \frac{LP}{N_0}\right)$ for large L.

3. With full CSI, the transmitter knows the channel, and for a given realization of the fading process $\{\mathbf{h}[n]\}_{n=1}^N$ the channel supports a rate:

$$R = \frac{1}{N}\sum_{n=1}^{N}\log\left(1 + \frac{\|\mathbf{h}[n]\|^2 P[n]}{N_0}\right) \tag{5.12}$$

and the problem becomes that of finding the optimal power allocation strategy. We note that the problem is the same as the one corresponding to the case of a single receive antenna, replacing $|h[n]|^2$ by $\|\mathbf{h}[n]\|^2$. It follows that the optimal solution is also obtained by waterfilling:

$$P^*(\|\mathbf{h}\|^2) = \left(\frac{1}{\lambda} - \frac{N_0}{\|\mathbf{h}\|^2}\right)^+ \tag{5.13}$$

where λ is chosen so that the power constraint is satisfied, i.e. $E[P^*(\|\mathbf{h}\|^2)] = P$. The resulting capacity is:

$$C = E\left[\log\left(1 + \frac{\|\mathbf{h}\|^2 P^*}{N_0}\right)\right] \tag{5.14}$$

At low SNR, when the system is power limited, the benefit of having CSI at the transmitter comes from the fact that we can transmit only when the channel is good, saving power (which is the limiting resource) when the channel is bad. The larger the fluctuation in the channel gain, the larger the benefit. If the channel gain is constant, then the waterfilling strategy reduces to transmitting with constant power, and there is no benefit in having CSI at the transmitter. When there are multiple receive antennas, there is diversity and $\|\mathbf{h}\|^2/L$ does not fluctuate much. In the limit as $L \to \infty$ we have seen that this random variable converges to a constant with probability one. Then, as L increases, the benefit of having CSI at the transmitter is reduced.

4.

$$P_{out} = Pr\left[\log\left(1 + \frac{\|\mathbf{h}\|^2 P}{N_0}\right) < R\right] = Pr\left[\|\mathbf{h}\|^2 < (2^R - 1)\frac{N_0}{P}\right] \tag{5.15}$$

We know that we can approximate the pdf of $\|\mathbf{h}\|^2$ around 0 by:

$$f(x) \approx \frac{1}{(L-1)!} x^{L-1} \tag{5.16}$$

where Rayleigh fading was assumed, and hence the distribution function of $\|\mathbf{h}\|^2$ evaluated at x is approximately given by:

$$F(x) \approx \frac{1}{L!} x^{L} \tag{5.17}$$

for x small. Thus, for large SNR we get the following approximation for the outage probability:

$$P_{out} \approx \frac{1}{L!}\left[(2^R - 1)\frac{N_0}{P}\right]^L \tag{5.18}$$

We see that having multiple antennas reduces the outage probability by a factor of $(2^R - 1)^L/L!$ and also increases the exponent of SNR^{-1} by a factor of L.

EXERCISE 5.14. 1. The Alamouti scheme transmits two independent symbols u_1, u_2 over the two antennas in two channel uses as follows:

$$\mathbf{X} = \begin{bmatrix} u_1 & -u_2^* \\ u_2 & u_1^* \end{bmatrix}.$$

To show that the scheme radiates energy in an isotropic manner, we need to show that the energy in the projection of this codeword in any direction $\mathbf{d} \in \mathbb{C}^2$ depends only on the magnitude of $\mathbf{d}$ and not its direction. Let $\mathbb{E}[u_1 u_2^*] = 0$ and $\mathbb{E}[|u_1|^2] = \mathbb{E}[|u_2|^2] = P/2$. We then have:

$$\mathbf{d}^\dagger \mathbb{E}[\mathbf{X}\mathbf{X}^\dagger]\mathbf{d} = [d_1^* \ d_2^*] \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = P\|\mathbf{d}\|^2.$$

2. Suppose that the transmitted vector $\mathbf{x} = [x_1 \ x_2]^T$ is such that $\mathbb{E}[x_1 x_2^*] = 0$ and $\mathbb{E}[|x_1|^2] = \mathbb{E}[|x_2|^2] = P$. Then, for any $\mathbf{d} = [d_1 \ d_2]^T$,

$$\mathbf{d}^\dagger \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]\mathbf{d} = \mathbf{d}^\dagger \mathbb{E}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} x_1^* & x_2^* \end{bmatrix} \mathbf{d} = \mathbf{d}^\dagger P \mathbf{I} \mathbf{d} = P\|\mathbf{d}\|^2,$$

hence the scheme radiates energy isotropically.

To prove the converse, assume that the scheme $\mathbf{x} = [x_1 \ x_2]^T$ is isotropic, i.e., for any two vectors $\mathbf{d}_a$ and $\mathbf{d}_b$ such that $\|\mathbf{d}_a\|^2 = \|\mathbf{d}_b\|^2 = 1$, we have that

$$\mathbf{d}_a^\dagger \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]\mathbf{d}_a = \mathbf{d}_b^\dagger \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]\mathbf{d}_b. \tag{5.19}$$

Then we must prove that $\mathbb{E}[x_1 x_2^*] = 0$ and $\mathbb{E}[|x_1|^2] = \mathbb{E}[|x_2|^2]$. To see that this must be so, first choose $\mathbf{d}_a = [1\ 0]^T$ and $\mathbf{d}_b = [0\ 1]^T$. Substituting this into (5.19) we obtain that $\mathbb{E}[|x_1|^2] = \mathbb{E}[|x_2|^2]$.

Now, choose $\mathbf{d}_a = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$, and $\mathbf{d}_b = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$. Then, (5.19) yields

$$\mathbb{E}[|x_1|^2 + x_1^* x_2 + x_1 x_2^* + |x_2|^2] = \mathbb{E}[|x_1|^2 - x_1^* x_2 - x_1 x_2^* + |x_2|^2],$$

Hence we get that $\mathbb{E}[x_1^* x_2 + x_1 x_2^*] = 0$ which implies that $\text{Real}(\mathbb{E}[x_1^* x_2]) = 0$.

Now, choose $\mathbf{d}_a = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{j\sqrt{2}} \end{bmatrix}$, and $\mathbf{d}_b = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{j\sqrt{2}} \end{bmatrix}$. Then, (5.19) yields

$$\mathbb{E}\left[|x_1|^2 + \frac{x_1^* x_2}{j} - \frac{x_1 x_2^*}{j} + |x_2|^2\right] = \mathbb{E}\left[|x_1|^2 - \frac{x_1^* x_2}{j} + \frac{x_1 x_2^*}{j} + |x_2|^2\right].$$

Hence we get that $\mathbb{E}\left[\frac{x_1^* x_2}{j} - \frac{x_1 x_2^*}{j}\right] = 0$, which implies that $\text{Imag}(\mathbb{E}[x_1^* x_2]) = 0$. Thus we conclude that $\mathbb{E}[x_1^* x_2] = 0$ and we have established the converse.

Exercise 5.15. 1. A MISO channel is given by the following input-output relation:

$$y[m] = \mathbf{h}^\dagger \mathbf{x}[m] + z[m],$$

with the total power constraint $\mathbb{E}[\|\mathbf{x}\|^2] \leq P$. The received SNR is given by

$$\mathsf{SNR} = \frac{\mathbb{E}[|\mathbf{h}^\dagger \mathbf{x}|^2]}{N_0} = \frac{\mathbb{E}[\mathbf{h}^\dagger \mathbf{x}\mathbf{x}^\dagger \mathbf{h}]}{N_0} = \frac{\mathbf{h}^\dagger \mathbf{K}_x \mathbf{h}}{N_0}.$$

Thus this channel is equivalent to a scalar channel with the same received SNR. Hence, the maximal rate of reliable communication on this channel is given by

$$C = \log(1 + \mathsf{SNR}) = \log\left(1 + \frac{\mathbf{h}^\dagger \mathbf{K}_x \mathbf{h}}{N_0}\right).$$

2. Since the covariance matrix $\mathbf{K}_x$ is positive semi-definite, it admits the decomposition $\mathbf{K}_x = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger$, where $\mathbf{\Lambda}$ is a diagonal matrix and $\mathbf{U}$ is a unitary matrix. Since the channel is i.i.d. Rayleigh, the vector $\mathbf{h}$ is isotropically distributed, i.e., $\mathbf{h}^\dagger \mathbf{U}$ has the same distribution as $\mathbf{h}^\dagger$. Thus the quadratic form

$$\mathbf{h}^\dagger \mathbf{K}_x \mathbf{h} = \mathbf{h}^\dagger \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger \mathbf{h} = (\mathbf{h}^\dagger \mathbf{U})\mathbf{\Lambda}(\mathbf{h}^\dagger \mathbf{U})^\dagger$$

has the same distribution as $\mathbf{h}^\dagger \mathbf{\Lambda} \mathbf{h}$. Therefore, we can restrict $\mathbf{K}_x$ to be diagonal without sacrificing outage performance.

Exercise 5.16.

Exercise 5.17.

Exercise 5.18. The outage probability of a parallel channel with L i.i.d. Rayleigh branches is given by the following expression:

$$P_{\text{out}}^{\text{parallel}} := \mathbb{P}\left(\sum_{l=1}^{L} \log(1 + |h_l|^2 \mathsf{SNR}) < LR\right).$$

Observe that the following inclusion holds:

$$\left\{\sum_{l=1}^{L} \log(1 + |h_l|^2 \mathsf{SNR}) < LR\right\} \supseteq \bigcap_{l=1}^{L} \{\log(1 + |h_l|^2 \mathsf{SNR}) < R\}.$$

Hence, since h_l's are i.i.d.,

$$P_{\text{out}}^{\text{parallel}} \geq \mathbb{P}\left(\log(1 + |h_l|^2 \mathsf{SNR}) < R\right)^L.$$

Since the branches are Rayleigh distributed, we have that, at high-SNR,

$$\mathbb{P}\left(\log(1 + |h_l|^2 \mathsf{SNR}\right) < R) = \mathbb{P}\left(|h_l|^2 < \frac{2^R - 1}{\mathsf{SNR}}\right) \approx \frac{2^R - 1}{\mathsf{SNR}}.$$

Hence, the outage of the parallel channel at high-SNR satisfies

$$P_{\text{out}}^{\text{parallel}} \geq \left(\frac{2^R - 1}{\mathsf{SNR}}\right)^L.$$

A simple upper bound that exhibits identical scaling with SNR is obtained by observing that

$$\left\{\sum_{l=1}^{L} \log(1 + |h_l|^2 \mathsf{SNR}) < LR\right\} \subseteq \bigcap_{l=1}^{L} \{\log(1 + |h_l|^2 \mathsf{SNR}) < LR\},$$

hence yielding

$$P_{\text{out}}^{\text{parallel}} \leq \left(\frac{2^{LR} - 1}{\mathsf{SNR}}\right)^L.$$

The SNR scaling of the lower and upper bounds is identical, though the pre-constants are slightly different. However, a more precise analysis can be done (see Section 9.1.3, equation (9.19) and Exercise 9.1)) which shows that the lower bound is actually tight, i.e., the outage probability scales as

$$P_{\text{out}}^{\text{parallel}} \approx \left(\frac{2^R - 1}{\mathsf{SNR}}\right)^L.$$

when the rate is given by $R = r \log \mathsf{SNR}$ for $0 \le r \le 1$. We now give a heuristic argument as to why this is true. For the complete proof, see [156], Theorem 4.

Let $R = r \log \mathsf{SNR}$ and let $|h_l|^2 = \mathsf{SNR}^{-\alpha_l}$, for $\alpha_l \in \mathbb{R}$ and $l = 1, \ldots, L$ (observe that we can always do this since $|h_l|^2$ is a non-negative random variable). The $|h_l|^2$ are independent and exponentially distributed with mean 1, i.e., the joint density p_h is given by

$$p_h(|h_1|^2, \ldots, |h_L|^2) = e^{-\sum_{l=1}^L |h_l|^2}.$$

Applying the change of variable to the above density, we obtain the joint density of α_l's, p_α:

$$p_\alpha(\alpha_1, \ldots, \alpha_L) = (\log \mathsf{SNR})^L e^{-\sum_{l=1}^L \mathsf{SNR}^{-\alpha_l}} \mathsf{SNR}^{-\sum_{l=1}^L \alpha_l}.$$

Now, we can express the outage probability as

$$\begin{aligned} P_{\text{out}}^{\text{parrallel}} &= \mathbb{P}\left(\prod_{l=1}^L (1 + \mathsf{SNR}|h_l|^2) < 2^{LR}\right), \\ &\approx \mathbb{P}\left(\mathsf{SNR}^{\sum_{l=1}^L (1-\alpha_l)^+} < 2^{LR}\right), \\ &= \mathbb{P}\left(\sum_{l=1}^L (1-\alpha_l)^+ < Lr\right), \end{aligned}$$

where, we've used the high-SNR approximation $(1 + \mathsf{SNR}|h_l|^2) \approx \mathsf{SNR}^{(1-\alpha_l)^+}$ (the function $(x)^+$, denotes $\max\{0, x\}$). Hence, the outage probability is given by the following integral:

$$P_{\text{out}}^{\text{parrallel}} = \int_{\mathcal{A}} (\log \mathsf{SNR})^L e^{-\sum_{l=1}^L \mathsf{SNR}^{-\alpha_l}} \mathsf{SNR}^{-\sum_{l=1}^L \alpha_l} d\alpha_1 \ldots \alpha_L,$$

where $\mathcal{A} = \left\{\alpha_1, \ldots, \alpha_L \in \mathbb{R} : \sum_{l=1}^L (1-\alpha_l)^+ < Lr\right\}$. Since we are considering the high-SNR regime, the term $(\log \mathsf{SNR})^n$ has no effect on the SNR exponent. Furthermore, the term $e^{-\sum_{l=1}^L \mathsf{SNR}^{-\alpha_l}}$ decays exponentially with SNR for $\alpha_l < 0$, so we can concentrate only on $\alpha_l > 0$. Moreover, the exponential terms approach 1 for $\alpha_l > 0$ and e for

$\alpha_l = 0$. Hence, the exponential terms have no effect on the SNR exponent. Thus, we can approximate the outage probability as

$$P_{\text{out}}^{\text{parrallel}} \approx \int_{\mathcal{A}^+} \mathsf{SNR}^{-\sum_{l=1}^{L}\alpha_l} d\alpha_1 \ldots \alpha_L,$$

where $\mathcal{A}^+ = \left\{\alpha_1, \ldots, \alpha_L > 0 : \sum_{l=1}^{L}(1-\alpha_l)^+ < Lr\right\}$. By Laplace's principle of large deviations, we have that the integral above is dominated by the term with the largest SNR exponent. Thus,

$$P_{\text{out}}^{\text{parrallel}} \approx \mathsf{SNR}^{-\inf_{\mathcal{A}^+}\sum_{l=1}^{L}\alpha_l}.$$

It can be verified that $\inf_{\mathcal{A}^+}\sum_{l=1}^{L}\alpha_l = (1-r)L$. Hence, we have that

$$P_{\text{out}}^{\text{parallel}} \approx \left(\frac{2^R}{\mathsf{SNR}}\right)^L.$$

when the rate is given by $R = r\log\mathsf{SNR}$ for $0 \le r \le 1$.

EXERCISE 5.19. 1. If we transmit the same signal $x[m]$ on each of the parallel channels, the received signal can be written as:

$$\mathbf{y}[m] = \mathbf{h}x[m] + \mathbf{z}[m].$$

The optimal receiver performs maximal ratio combining and hence this channel becomes equivalent to a scalar AWGN channel with received signal-to-noise ratio given by $\|\mathbf{h}\|^2\mathsf{SNR}$, where SNR is the per-channel signal-to-noise ratio on the original parallel channel. Now, suppose that the rate requirement is R bits/sec/Hz per channel. Then, this *scheme* has outage probability given by

$$\begin{aligned} P_{\text{out}}^{\text{repetition}} := \mathbb{P}(\log(1+\|\mathbf{h}\|^2\mathsf{SNR}) < LR) &= \mathbb{P}\left(\|\mathbf{h}\|^2 < \frac{2^{LR}-1}{\mathsf{SNR}}\right), \\ &\approx \frac{1}{L!}\left(\frac{2^{LR}-1}{\mathsf{SNR}}\right)^L, \end{aligned}$$

where the last line comes from the high-SNR approximation of the distribution of $\mathbf{h}$ (chi-square with $2L$ degrees of freedom).

2. Using the result of Exercise 5.18, we see that in order to guarantee the same outage probability, we require a larger SNR in the repetition scheme then the minimal required SNR dictated by the outage performance of the channel. In particular, let $\mathsf{SNR}_{\text{parallel}}$ and $\mathsf{SNR}_{\text{repetition}}$ be the minimum required SNR and the

SNR required under repetition coding, respectively. Then, in order to have the same outage probabilities, we need that (at high-SNR),

$$\left(\frac{2^R}{\mathsf{SNR}_{\text{parallel}}}\right)^L = \frac{1}{L!}\left(\frac{2^{LR}}{\mathsf{SNR}_{\text{repetition}}}\right)^L,$$

and consequently we obtain that

$$\frac{\mathsf{SNR}_{\text{repetition}}}{\mathsf{SNR}_{\text{parallel}}} = \frac{2^{R(L-1)}}{(L!)^{1/L}}.$$

For instance, with $R = 1$ bps/Hz, and $L = 5$, the repetition scheme requires roughly 18 dB more power over the minimal power requirement to achieve the required outage performance.

3. For small $x > 0$, $\log(1+x) \approx x \log_2 e$. We use this for the low-SNR approximation for both the outage probability of the repetition scheme as well as that of the parallel channel itself. Hence we have that

$$P_{\text{out}}^{\text{parallel}} := \mathbb{P}\left(\sum_{l=1}^{L} \log(1 + |h_l|^2 \mathsf{SNR}) < LR\right) \approx \mathbb{P}\left(\|\mathbf{h}\|^2 < \frac{LR}{\mathsf{SNR}\log_2 e}\right),$$

$$P_{\text{out}}^{\text{repetition}} := \mathbb{P}(\log(1 + \|\mathbf{h}\|^2 \mathsf{SNR}) < LR) \approx \mathbb{P}\left(\|\mathbf{h}\|^2 < \frac{LR}{\mathsf{SNR}\log_2 e}\right),$$

hence the outage performance of the repetition scheme is approximately optimal in the low-SNR regime.

To conclude, in the high-SNR regime, the AWGN parallel channel is degree-of-freedom limited and the repetition scheme performs poorly in this regime since it is wasteful of the available degrees of freedom (independent channels) by virtue of sending the same information on all of them at any given time-slot. However, at low SNR, the channel is SNR limited and this shortcoming of the repetition scheme is not evident since the scheme does reap the receive beamforming (coherent combining) benefit and hence match the power gain achievable by any other scheme.

Exercise 5.20. 1. The low-SNR ϵ-outage probability approximation of the parallel channel is given by (see Exercise 5.19, part 3):

$$\mathbb{P}\left(\|\mathbf{h}\|^2 < \frac{LC_\epsilon}{\mathsf{SNR}\log_2 e}\right) = \epsilon,$$

where C_ϵ denotes the per-channel ϵ-outage capacity, i.e., the largest rate achievable while maintaining outage probability below ϵ. Let $F(x) = \mathbb{P}(\|\mathbf{h}\|^2 > x)$ be the complementary CDF of $\mathbf{h}$. Then we have that

$$C_\epsilon = \frac{1}{L} F^{-1}(1-\epsilon) \mathsf{SNR}\log_2 e.$$

2. For Rayleigh i.i.d. fading branches, $F^{-1}(1-\epsilon) \approx (L!)^{\frac{1}{L}}\epsilon^{\frac{1}{L}}$, and so

$$C_\epsilon = \frac{1}{L}(L!)^{\frac{1}{L}}\epsilon^{\frac{1}{L}}\mathsf{SNR}\log_2 e,$$

is the per-channel outage capacity.

3. The delay-spread of the channel is 1μs. Hence, from equation (2.47), page 33, we know that the coherence bandwidth is $\frac{1}{2\times 10^{-6}} = 0.5$ MHz. But the available bandwidth is 1.25 MHz. Hence, if we exploit the frequency coherence, we can have two independent, parallel channels in frequency. Also, since our time constraint is 100ms and the coherence time is 50ms, we have two parallel channels in time. This makes a total of four parallel channels that we can exploit. Consequently, we let $L = 4$ in our calculations.

   Since the SINR per chip is -17 dB and the processing gain is $G = W/R = 1.25\text{MHz} \times 100\text{ms} = 125000$, the SINR per bit per user is roughly 34 dB. Plugging in these values into the formula given in part (2) of this question, we get that $C_{0.01}$ is roughly 631 bps/Hz/user.

   On the other hand, the capacity of the unfaded AWGN channel with the same SNR is roughly 3607 bps/Hz. Thus the 1%-outage capacity of this parallel channel is roughly 17.5% of the unfaded AWGN channel with the same received SNR.

Solution 5.21:

1. Using only one antenna at a time, we convert the MISO channel into a parallel channel. The maximal rate achievable with this strategy is given by:

$$C^{\text{parallel}} = \frac{1}{L}\sum_{l=1}^{L}\log(1+|h_l|^2\mathsf{SNR}),$$

compared by the capacity of this MISO channel (observe that the channel gain is constant and known to both the transmitter and receiver):

$$C^{\text{MISO}} = \log(1+\|\mathbf{h}\|^2\mathsf{SNR}).$$

At high-SNR, we can approximate the two rates as follows:

$$C^{\text{parallel}} \approx \log\mathsf{SNR} + \frac{1}{L}\sum_{l=1}^{L}\log|h_l|^2,$$
$$C^{\text{MISO}} \approx \log\mathsf{SNR} + \log\|\mathbf{h}\|^2.$$

Hence, at high-SNR, the ratio of the two rates goes to 1.

2. At low-SNR, we can make the following approximations:

$$C^{\text{parallel}} \approx \frac{1}{L}\sum_{l=1}^{L}|h_l|^2\mathsf{SNR}\log_2 e = \frac{1}{L}\|\mathbf{h}\|^2\mathsf{SNR}\log_2 e,$$
$$C^{\text{MISO}} \approx \|\mathbf{h}\|^2\mathsf{SNR}\log_2 e.$$

   Thus, the loss from capacity goes to $\frac{1}{L}$ as $\mathsf{SNR} \to 0$.

   The parallelization scheme is degree-of-freedom efficient so at high-SNR its performance is close to the optimal performance on the MISO channel due to the fact that the AWGN MISO channel is degree-of-freedom limited at high-SNR. However, the optimal strategy is for the transmitter to do beamforming (having knowledge of the channel) and hence harness the power gain afforded in this way. The parallelization scheme does not perform beamforming and hence suffers a loss from capacity in the SNR-limited low-SNR regime.

3. The outage probability expressions of the MISO channel and the scheme which turns it into a parallel channel are given by:

$$P_{\text{out}}^{\text{parallel}} := \mathbb{P}\left(\sum_{l=1}^{L}\log(1+|h_l|^2\mathsf{SNR}) < LR\right),$$
$$P_{\text{out}}^{\text{MISO}} := \mathbb{P}\left(\log(1+\|\mathbf{h}\|^2\mathsf{SNR}) < R\right).$$

   Assuming i.i.d. Rayleigh fading, we can use the result of Exercise 5.18 to obtain the high-SNR approximations:

$$P_{\text{out}}^{\text{parallel}} \approx \left(\frac{2^R-1}{\mathsf{SNR}}\right)^L,$$
$$P_{\text{out}}^{\text{MISO}} \approx \frac{1}{L!}\left(\frac{2^R-1}{\mathsf{SNR}}\right)^L,$$

   hence, the outage probability of the scheme which converts the MISO channel to a parallel channel is $L!$ times larger than the actual outage probability of the MISO channel at high-SNR.

   At low-SNR, we have

$$C_\epsilon^{\text{MISO}} \approx F^{-1}(1-\epsilon)\mathsf{SNR}\log_2 e,$$
$$C_\epsilon^{\text{parallel}} \approx \frac{1}{L}F^{-1}(1-\epsilon)\mathsf{SNR}\log_2 e,$$

   hence, the outage capacity of the scheme is L times smaller than the outage capacity of the MISO channel at low-SNR.

EXERCISE 5.21.

EXERCISE 5.22.

EXERCISE 5.23. 1. For the AWGN channel the maximum achievable rate is given by:

$$R = W \log\left(1 + \frac{\bar{P}}{N_0 W}\right) = W \log\left(1 + \frac{\mathcal{E}_b R}{N_0 W}\right) \tag{5.20}$$

where we used $\bar{P}/R = \mathcal{E}_b$.

Then, the minimum required $\mathcal{E}_b/N_0$ for reliable communication is:

$$\left(\frac{\mathcal{E}_b}{N_0}\right)_{req} = \frac{W}{R}\left(2^{R/W} - 1\right) \tag{5.21}$$

For the IS-95 system we get $\left(\frac{\mathcal{E}_b}{N_0}\right)_{req} = 0.695 = -1.58\text{dB}$.

At low SNR R/W is small, and we can approximate $2^{R/W} = \exp[(R/W)\ln 2] \approx 1 + (R/W)\ln 2$, to get

$$\left(\frac{\mathcal{E}_b}{N_0}\right)_{req} \approx \ln 2 = -1.59\text{dB} \tag{5.22}$$

and we see that as the SNR goes to zero, the minimum $\mathcal{E}_b/N_0$ requirement is -1.59dB.

2. Since we are forced to repeat each transmitted symbol 4 times, we consider the received signal in a block of length 4:

$$\mathbf{y} = \mathbf{1}x + \mathbf{z} \tag{5.23}$$

and use 3.a) to conclude that $I(x; \mathbf{y}) \leq \log(1 + 4P/N_0)$ where the upper bound can be achieved by choosing the input distribution to be $CN(0, P)$ i.i.d.. Then the maximum achievable rate (in bits/s/Hz) of this strategy is:

$$R_{max} = \frac{1}{4}\log\left(1 + \frac{4P}{N_0}\right) \tag{5.24}$$

which is strictly smaller than $\log(1 + P/N_0)$, the capacity of the AWGN channel. The loss is due to the concavity of the $\log(\cdot)$ function. For small x, $\log(x)$ is approximately linear and the loss due to concavity is small for low SNR. On the other hand, repetition coding has a large loss for high SNR.

3. Loss is greater at high SNR where the loss of d.o.f. is felt more.

4. For repetition coding the minimum $\mathcal{E}_b/N_0$ required for reliable communication is given by:

$$\left(\frac{\mathcal{E}_b}{N_0}\right)_{req} = \frac{W}{R}\left(\frac{2^{4R/W}-1}{4}\right) \tag{5.25}$$

The increase in $\mathcal{E}_b/N_0$ requirement is:

$$\frac{\left(\frac{\mathcal{E}_b}{N_0}\right)_{req(rep)}}{\left(\frac{\mathcal{E}_b}{N_0}\right)_{req(AWGN)}} = \frac{2^{4R/W}-1}{4(2^{R/W}-1)} \tag{5.26}$$

For the IS-95 system this loss is only 0.035dB.

EXERCISE 5.24. 1. The channel model is

$$y[m] = h[m]x[m] + w[m],$$

and the rate it can support when channel state is $h[m]$ is

$$R = \log\left(1 + \frac{|h[m]|^2 P(h[m])}{N_0}\right).$$

Using channel inversion to keep a constant rate R, we need

$$P(h[m]) = \frac{(2^R-1)N_0}{|h[m]|^2}.$$

Thus the average power needed is

$$\begin{aligned}\mathbb{E}[P] &= (2^R-1)N_0\mathbb{E}\left(\frac{1}{|h[m]|^2}\right)\\ &= (2^R-1)N_0\int_0^\infty \frac{1}{x}e^{-x}dx\\ &> (2^R-1)N_0\int_0^M \frac{1}{x}e^{-x}dx\\ &> (2^R-1)N_0e^{-M}\int_0^M \frac{1}{x}dx = \infty\end{aligned}$$

2. The Channel model is

$$y_l[m] = h_l[m]x[m] + w_l[m], \quad l = 1, \ldots, L$$

and the rate it can support when channel state is $\mathbf{h}[m] = (h_1[m], \dots, h_L[m])$ is

$$R = \frac{1}{2}\log\left(1 + \frac{|\mathbf{h}[m]|^2 P(\mathbf{h}[m])}{N_0}\right).$$

Using channel inversion to keep a constant rate R, we need

$$P(\mathbf{h}[m]) = \frac{(2^R - 1)N_0}{|\mathbf{h}[m]|^2}.$$

Since $|\mathbf{h}[m]|^2$ is a χ^2 distribution with pdf

$$f(x) = \frac{x^{L-1}}{(L-1)!}e^{-x},$$

the average power needed is

$$\begin{aligned}\mathbb{E}[P] &= (2^R - 1)N_0\mathbb{E}\left(\frac{1}{|\mathbf{h}[m]|^2}\right)\\ &= (2^R - 1)N_0 \int_0^\infty \frac{1}{x}\frac{x^{L-1}}{(L-1)!}e^{-x}dx\end{aligned}$$

$$= \frac{(2^R - 1)N_0}{L-1}.$$

3. Assume the noise $w \sim \mathcal{CN}(0,1)$, for different target rate and L, the average power is plotted in the following figure.

   We can see that the power needed is decreasing with increasing number of receiver antennas (actually inversely proportional to $L-1$).

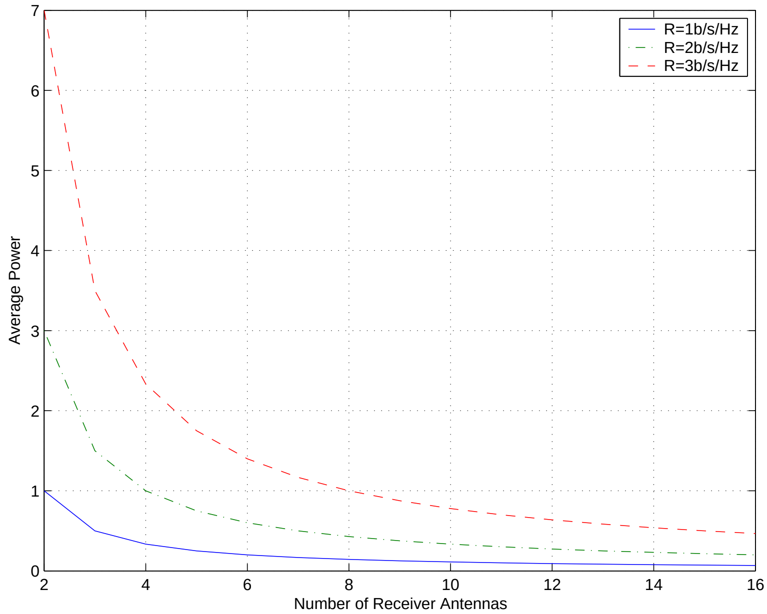
EXERCISE 5.25. 1. Using optimal scheme, the capacity is

$$C = W\log(1 + \text{SINR}),$$

where $W = 1.25$MHz. Hence the SINR threshold for using capacity achieving codes is

$$\text{SINR} = 2^{\frac{C}{W}} - 1$$

The following table compares the SINR threshold of using capacity achieving codes to that of IS-856. The differences are always larger than 3dB. So the codes in IS-856 are not close to optimal.



Requared rate(kb/s)	Optimal SINR threshold(dB)	SINR threshold using IS-856(dB)
38.4	-16.7	-11.5
76.8	-13.6	-9.2
153.6	-10.5	-6.5
307.2	-7.3	-3.5
614.4	-3.9	-0.5
921.6	-1.8	2.2
1228.8	-0.1	3.9
1843.2	2.5	8.0
2457.6	4.6	10.3

2. When repeated L times, the capacity is

$$C = \frac{W}{L} \log(1 + L \times \text{SINR}).$$

So the threshold SINR is

$$\text{SINR} = \frac{2^{\frac{LC}{W}} - 1}{L}.$$

When $C = 38.4kb/s$, $W = 1.25MHz$, and $L = 16$, the SINR threshold is -16.0 dB. Compared to -16.7 dB computed in part (1), there is not much performance loss from the repetition.

Exercise 5.26. 1. Given

$$h[m+1] = \sqrt{1-\delta}h[m] + \sqrt{\delta}w[m+1]$$

the auto-correlation function of the channel process is

$$\begin{aligned}
R[n] &= \mathbb{E}\left[h^*[m]h[m+n]\right] \\
&= \mathbb{E}\left[h^*[m](\sqrt{1-\delta}h[m+n-1] + \sqrt{\delta}w[m+n-1])\right] \\
&= \sqrt{1-\delta}\mathbb{E}\left[h^*[m]h[m+n-1]\right] \\
&= \sqrt{1-\delta}R[n-1].
\end{aligned}$$

Thus

$$R[n] = (\sqrt{1-\delta})^n R[0] = (\sqrt{1-\delta})^n.$$

2. If coherence time is T_c and sampling rate is $W = 2 \times 1.25MHz$, then

$$(\sqrt{1-\delta})^{WT_c} = 0.05$$

leads to

$$\delta = 1 - (0.05)^{\frac{2}{WT_c}}$$

So for $T_c = 25ms$ (walking), $\delta = 0.0000958$; for $T_c = 2.5ms$ (driving), $\delta = 0.000958$.

3. Since $h[0]$ and $h[n]$ are jointly Gaussian, the optimal estimator is MMSE.

$$\begin{aligned}
\hat{h}[n] &= \mathbb{E}[h[n]|h[0]] \\
&= \frac{\mathbb{E}[h^*[0]h[n]]}{\mathbb{E}[|h[0]|^2]}h[0] \\
&= (\sqrt{1-\delta})^n h[0].
\end{aligned}$$

4. From the property of MMSE for jointly Gaussian random variables, we can write

$$h[n] = \hat{h}[n] + h_e[n],$$

where the estimation error $h_e[n]$ is independent of $h[n]$, with variance

$$\begin{aligned}
\sigma_e^2 &= \mathbb{E}[|h_e[n]|^2] \\
&= \mathbb{E}[|h[n]|^2] - \frac{\mathbb{E}[h^*[0]h[n]]\mathbb{E}[h[0]h^*[n]]}{\mathbb{E}[|h[0]|^2]} \\
&= 1 - (1-\delta)^n
\end{aligned}$$

For IS-856 with 2-slot delay in the feed back,

$$\sigma_e^2 = 1 - (1-\delta)^n = 1 - (0.05)^{\frac{2n}{WT_c}},$$

where $n \sim 4000$. For $T_c = 25ms$ (walking), $\sigma_e^2 = 0.318$; for $T_c = 2.5ms$ (driving), $\sigma_e^2 = 0.978$.

Exercise 5.27.

Exercise 5.28.

Exercise 5.29.

Exercise 5.30.

Chapter 6

Solutions to Exercises

EXERCISE 6.1. Channel model is

$$y[m] = x_1[m] + x_2[m] + w[m].$$

The signal power at receiver is

$$P = \mathbb{E}[(x_1[m] + x_2[m])^2] = P_1 + P_2 + 2\mathbb{E}[x_1[m]x_2[m]].$$

When the two users can cooperate, they can choose the correlation of $x_1[m]$ and $x_2[m]$ to be one and thus get the largest total power

$$P = P_1 + P_2 + 2\sqrt{P_1 P_2},$$

and hence the maximum sum rate they can achieve is

$$C_{coop} = \log\left(1 + \frac{P_1 + P_2 + 2\sqrt{P_1 P_2}}{N_0}\right).$$

In the case of $P_1 = P_2 = P$,

$$C_{coop} = \log\left(\frac{1 + 4P}{N_0}\right),$$

whereas the sum rate without cooperation is

$$C_{no_coop} = \log\left(\frac{1 + 2P}{N_0}\right).$$

At high SNR,

$$\frac{C_{coop}}{C_{no_coop}} \simeq \frac{\log(4P/N_0)}{\log(2P/N_0)} \to 1 \quad \text{as } P \to \infty.$$

At low SNR,

$$\frac{C_{coop}}{C_{no_coop}} \simeq \frac{4P/N_0}{2P/N_0} = 2 \quad \text{as } P \to 0.$$

Thus in lower SNR region the cooperative gain is more effective.

EXERCISE 6.2. For orthogonal multiple access channel the rates of the two users satisfy

$$R_1 < \alpha \log\left(1 + \frac{P_1}{\alpha N_0}\right)$$
$$R_2 < (1-\alpha) \log\left(1 + \frac{P_2}{(1-\alpha) N_0}\right)$$

When the degrees of freedom are split proportional to the powers of the users, we have

$$\alpha = \frac{P_1}{P_1 + P_2}.$$

Thus the sum rate satisfy

$$R_1+R_2 < \frac{P_1}{P_1+P_2} \log\left(1 + \frac{P_1}{\frac{P_1}{P_1+P_2} N_0}\right) + \frac{P_2}{P_1+P_2} \log\left(1 + \frac{P_2}{\frac{P_2}{P_1+P_2} N_0}\right) = \log\left(1 + \frac{P_1+P_2}{N_0}\right),$$

which is the optimal sum rate.

For arbitrary split of degrees of freedom, from the strictly concavity property of $log(1+x)$, we have

$$\begin{aligned} R_1 + R_2 &< \alpha \log\left(1 + \frac{P_1}{\alpha N_0}\right) + (1-\alpha) \log\left(1 + \frac{P_2}{(1-\alpha) N_0}\right) \\ &\leq \log\left(1 + \left(\alpha \frac{P_1}{\alpha N_0} + (1-\alpha) \frac{P_2}{(1-\alpha) N_0}\right)\right) \\ &= \log\left(1 + \frac{P_1+P_2}{N_0}\right), \end{aligned}$$

and equality holds only when

$$\frac{P_1}{\alpha} = \frac{P_2}{(1-\alpha)},$$

that is when the degrees of freedom are split proportional to the powers of the users. Any other split of degrees of freedom are strictly sub-optimal.

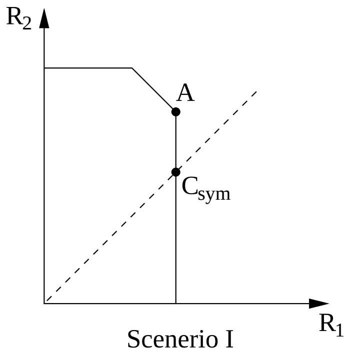
EXERCISE 6.3. The symmetric capacity is

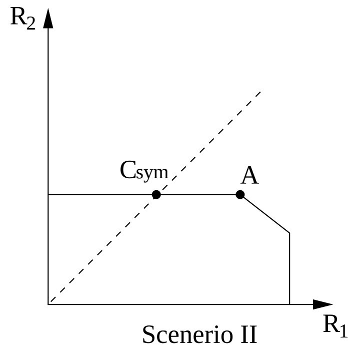
$$C_{sym} = \frac{1}{2} \log\left(1 + \frac{P_1+P_2}{N_0}\right).$$

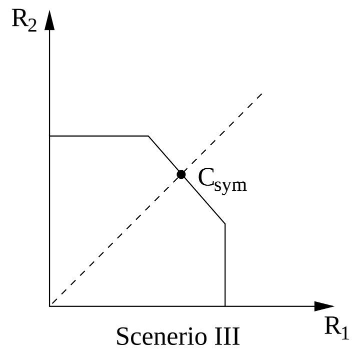
There are three scenarios of capacity region shown in the following Figure.

In scenario I and II, the point A is superior to the symmetric rate point, in scenario III we do not have a superior point.

EXERCISE 6.4.







Exercise 6.5.

Exercise 6.6.

Exercise 6.7.

Exercise 6.8.

Exercise 6.9.

Exercise 6.10.

Exercise 6.11.

Exercise 6.12. Let e_i means the event of decoding incorrectly at stage i, and e_i^c means the event of decoding correctly at stage i, then the probability of error for the kth user under SIC satisfies

$$\begin{aligned} p_e &= \mathbb{P}\left(e_1 \bigcup (e_2|e_1^c) \bigcup (e_3|e_1^c, e_2^c) \bigcup \cdots \bigcup (e_k|e_1^c \dots e_{k-1}^c)\right) \\ &\leq \mathbb{P}(e_1) + \mathbb{P}(e_2|e_1^c) + \cdots + \mathbb{P}(e_k|e_1^c \dots e_{k-1}^c) \\ &= \sum_{i=1}^{k} p_e^{(i)}, \end{aligned}$$

where $p_e^{(i)} = \mathbb{P}(e_i | e_1^c \dots e_{i-1}^c)$ is the probability of decoding the ith user incorrectly assuming that all the previously users are decoded correctly.

EXERCISE 6.13. In the following we denote $\text{SNR}_1 = \frac{P_1 T_c}{N_0}$ and $\text{SNR}_2 = |h_2|^2 \frac{P_2 T_c}{N_0}$.

1. By neglecting user 2 and using training signal $x_t[m]$, we have

$$y[m] = h_1[m] x_t[m] + \omega[m].$$

Let $\hat{h}_1[m]$ be the estimation of channel state $h_1[m]$, then the MMSE of $h_1[m]$ from $y[m]$ can be calculated using (A.85) in Appendix A, and it is

$$\begin{aligned} \mathbb{E}\left[(h_1[m] - \hat{h}_1[m])^2\right] &= \frac{\mathbb{E}\left[|h_1|^2\right] N_0}{\mathbb{E}\left[|h_1|^2\right] \|x_t\|^2 + N_0} \\ &= \frac{N_0}{0.2 P_1 T_c + N_0} \\ &= \frac{1}{0.2\text{SNR}_1 + 1}. \end{aligned}$$

2. The channel can be written as

$$y[m] = \hat{h}_1[m] x_1[m] + h_2[m] x_2[m] + (h_1[m] - \hat{h}_1[m]) x_1[m] + \omega[m],$$

where the SIC decoder can subtract $\hat{h}_1[m] x_1[m]$ from channel estimation and user 1's signal. The term $(h_1[m] - \hat{h}_1[m]) x_1[m] + \omega[m]$ is the noise plus the interference from inaccurate estimation of the channel. Thus

$$\mathbb{E}\left[((h_1[m] - \hat{h}_1[m]) x_1[m] + \omega[m])^2\right] = \frac{P_1 T_c}{0.2\text{SNR}_1 + 1} + N_0,$$

and the SINR of user 2 is

$$\begin{aligned} \text{SINR}_2 &= \frac{|h_2|^2 P_2 T_c}{\left(\frac{P_1 T_c}{0.2\text{SNR}_1 + 1} + N_0\right)} \\ &= \frac{\text{SNR}_2}{\left(\frac{\text{SNR}_1}{0.2\text{SNR}_1 + 1} + 1\right)} \\ &= \frac{\text{SNR}_2 (0.2\text{SNR}_1 + 1)}{1.2\text{SNR}_1 + 1}. \end{aligned}$$

The numerical calculation is shown in the following figure. We can see that the degradation is worse if the power of user 1 increases. This is because user 1's signal is the interference to user 2 due to inaccurate estimation of the channel, and with the increase of the power of user 1, the interference also increases, hence the SINR for user 2 decreases.

3. If user 1 is decoded correctly, then we can estimate the channel state from both the training symbol and user 1's signal. That is, we estimate $h_1[m]$ from y_1 and y_2 where

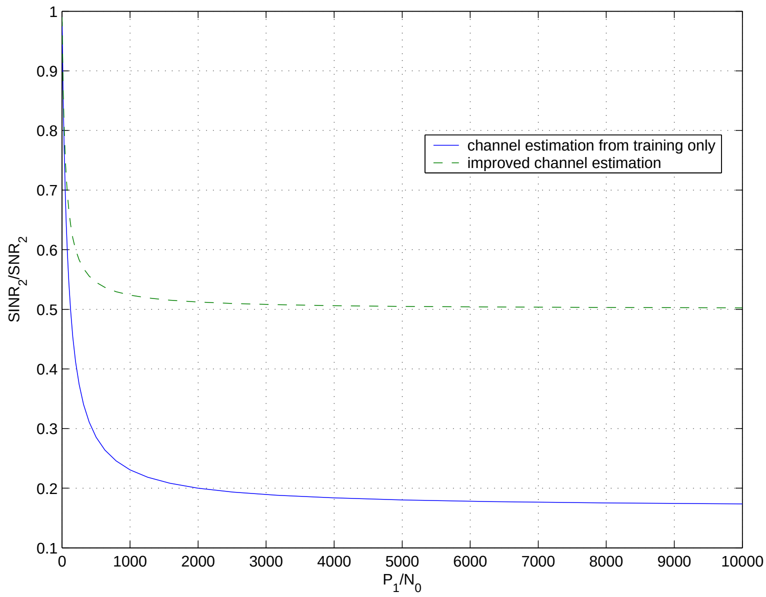
$$y_1 = h_1 x_t + \omega_1,$$
$$y_2 = h_1 x_1 + \omega_2.$$

Using (A.85) in Appendix A, we have

$$\begin{aligned}\mathbb{E}\left[(h_1 - \hat{h}_1)^2\right] &= \frac{\mathbb{E}\left[|h_1|^2\right] N_0}{\mathbb{E}\left[|h_1|^2\right]\left(\|x_t\|^2 + \|x_1\|^2\right) + N_0} \\ &= \frac{N_0}{P_1 T_c + N_0} \\ &= \frac{1}{\mathrm{SNR}_1 + 1}.\end{aligned}$$

Using the above estimation error to redo part(2), we get

$$\begin{aligned}\mathrm{SINR}_2 &= \frac{\mathrm{SNR}_2}{\left(\frac{\mathrm{SNR}_1}{\mathrm{SNR}_1+1} + 1\right)} \\ &= \frac{\mathrm{SNR}_2(\mathrm{SNR}_1 + 1)}{2\mathrm{SNR}_1 + 1}.\end{aligned}$$



From the figure we can see that the SINR for user 2 improves, especially at high P_1/N_0.

EXERCISE 6.14. 1. At very high SNR, we have

$$\log\left(1+\frac{\sum_{i=1}^{k} P|h_i|^2}{N_0}\right) \simeq \log\frac{P}{N_0}$$

for $k = 1, 2, \ldots, K$.

Thus for any $S \subset \{1, \ldots, K\}$, we have that

$$|S|R > \log\left(1+\frac{\sum_{i\in S} P|h_i|^2}{N_0}\right)$$

is approximately equivalent to

$$R > \frac{1}{|S|}\log\frac{P}{N_0}$$

at very high SNR, and hence the dominating event for $p_{\text{out}}^{\text{ul}}$ is when $|S| = K$, that is, the one on sum rate.

2. At high SNR,

$$\begin{aligned} p_{out} &\simeq \mathbb{P}\left\{KR > \log(1+\text{SNR}\sum_{k=1}^{K}|h_k|^2)\right\} \\ &= \mathbb{P}\left\{\sum_{k=1}^{K}|h_k|^2 < \frac{2^{KR}-1}{\text{SNR}}\right\} \\ &\simeq \frac{1}{K!}\left(\frac{2^{KR}-1}{\text{SNR}}\right)^K. \end{aligned}$$

Let

$$\frac{1}{K!}\left(\frac{2^{KC_\epsilon^{\text{sym}}}-1}{\text{SNR}}\right)^K = \epsilon,$$

we get

$$C_\epsilon^{\text{sym}} = \frac{1}{K}\log\left(1+\text{SNR}(K!\epsilon)^{1/K}\right).$$

3. At very high SNR,

$$\begin{aligned} \frac{C_\epsilon^{\text{sym}}}{C_\epsilon} &= \frac{\frac{1}{K}\log\left(1+\text{SNR}(K!\epsilon)^{1/K}\right)}{\log(1+\epsilon\text{SNR})} \\ &\simeq \frac{1}{K}\frac{\log(\text{SNR})+\log(K!\epsilon)^{1/K}}{\log(\text{SNR})+\log\epsilon} \\ &\simeq \frac{1}{K}. \end{aligned}$$

EXERCISE 6.15.

EXERCISE 6.16.

EXERCISE 6.17.

EXERCISE 6.18.

EXERCISE 6.19.

EXERCISE 6.20.

EXERCISE 6.21.

EXERCISE 6.22. The probability that no one user sends a request rate is the probability that all users' channel is less than γ, that is

$$\begin{aligned}\mathbb{P}\{\text{no one sends a request rate}\} &= \prod_{k=1}^{K} \mathbb{P}\{\text{user k's channel is less than}\gamma\} \\ &= \prod_{k=1}^{K} \mathbb{P}\{|h_k|^2 \text{SNR} < \gamma\} \\ &= \left(1 - e^{-\gamma/\text{SNR}}\right)^K \\ &= (1 - e^{-\gamma})^K,\end{aligned}$$

where we used $\text{SNR} = 0dB = 1$. We need this probability to be ϵ, thus

$$(1 - e^{-\gamma})^K = \epsilon,$$

and the solution for Γ is

$$\gamma = -\ln\left(1 - \epsilon^{1/K}\right).$$

Now, the probability that any user sends a request is

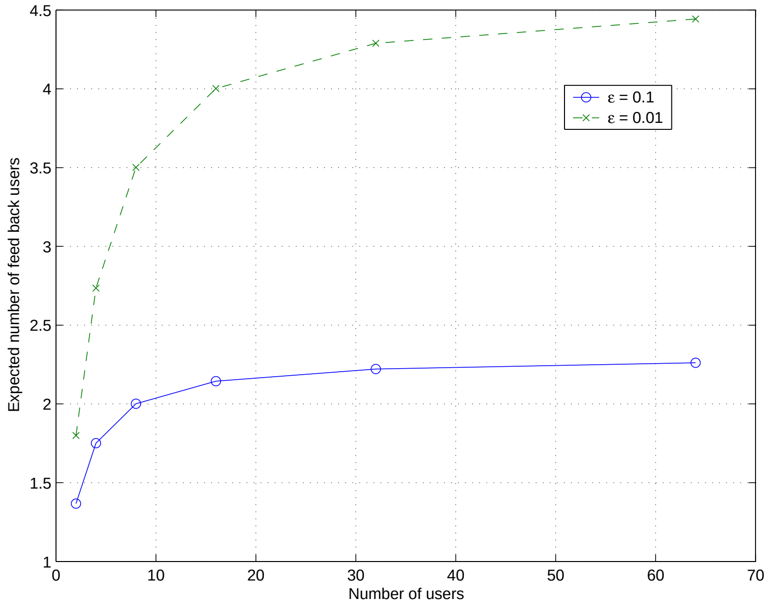
$$p = e^{-\gamma},$$

and the number of users that sends in a requested rate is a Binomial random variable with parameter p, hence

$$\mathbb{E}[\text{number of users that sends in a requested rate}] = Kp = Ke^{-\gamma} = K\left(1 - \epsilon^{1/K}\right).$$

The expected number of users that sends in a requested rate for different K and ϵ is shown in the following figure.

EXERCISE 6.23.



EXERCISE 6.24.

EXERCISE 6.25.

EXERCISE 6.26.

EXERCISE 6.27.

EXERCISE 6.28.

EXERCISE 6.29.

EXERCISE 6.30.

EXERCISE 6.31. 1. Under Alamouti scheme, the effective SNR is

$$u_1 = (|h_1|^2 + |h_2|^2)\frac{\text{SNR}}{2},$$

for the single antenna case, the effective SNR is

$$u_2 = |h_1|^2\text{SNR},$$

where h_1 and h_2 are i.i.d $\mathcal{CN}(0,1)$. Hence the distributions are

$$f_1(u_1) = \frac{2}{\text{SNR}}\left(\frac{2u_1}{\text{SNR}}\right)e^{-\frac{2u_1}{\text{SNR}}},$$
$$f_2(u_2) = \frac{1}{\text{SNR}}e^{-\frac{u_2}{\text{SNR}}}.$$

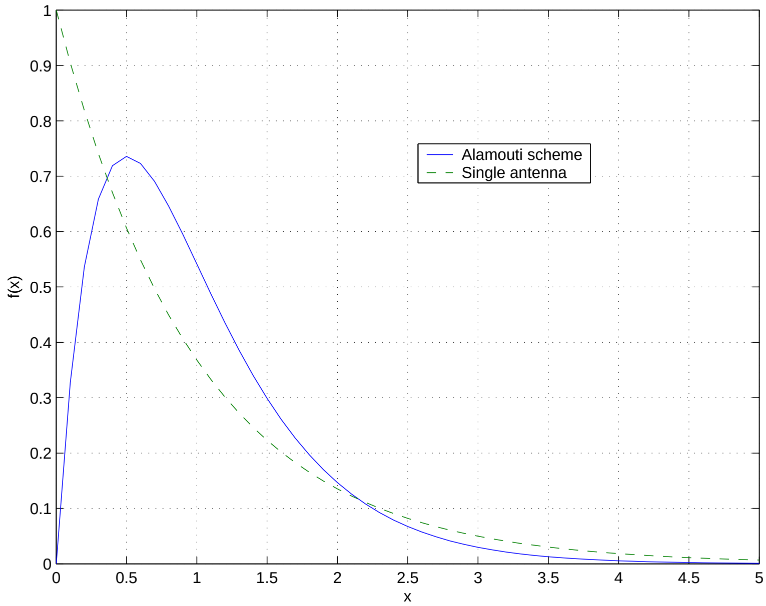


Figure 6.1: Probability distribution functions

2. From the plot in part (1), we can see the probability that effective SNR is low is smaller for the Alamouti scheme. So we expect that dual antenna can provide some gain. Rigorous proof is as follows.

   Let C_A be the capacity for Alamouti scheme, and C_S be the capacity for the single antenna case, we have

$$
\begin{aligned}
C_S &= \mathbb{E}\left[\log(1+|h_1|^2\mathrm{SNR})\right] \\
&= \frac{1}{2}\mathbb{E}\left[\log(1+|h_1|^2\mathrm{SNR})\right] + \frac{1}{2}\mathbb{E}\left[\log(1+|h_2|^2\mathrm{SNR})\right] \\
&\le \mathbb{E}\left[\log\left(1+\frac{|h_1|^2+|h_2|^2)}{2}\mathrm{SNR}\right)\right] \\
&= C_A,
\end{aligned}
$$

   where in the first step we used the fact that $|h_1|^2$ and $|h_2|^2$ are i.i.d, and in the second step we used the Jensen's inequality.

3. When there are K users, assume SNR $= 1$, we have

   For the Alamouti scheme, the effective SNR is $u_1 = \frac{1}{2}\max_{1,\ldots,K}(|h_{1k}|^2 + |h_{2k}|^2)$, and

$$
\mathbb{P}\left(\frac{1}{2}\max_{k=1,\ldots,K}(|h_{1k}|^2+|h_{2k}|^2) < x\right) = \left(1-(1+2x)e^{-2x}\right)^K,
$$

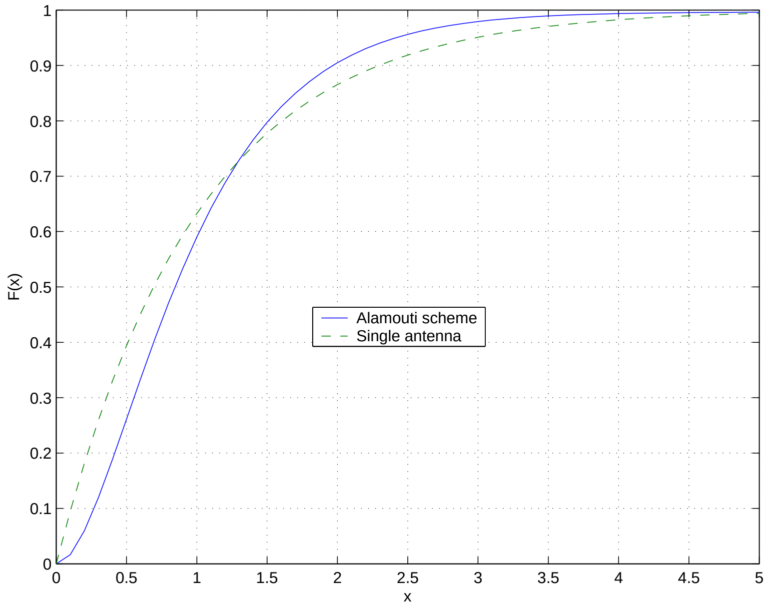


Figure 6.2: Cumulative distribution functions

so

$$f_1(u_1) = 4Ku_1e^{-2u_1}\left(1-(1+2u_1)e^{-2u_1}\right)^{K-1}.$$

Thus the capacity under Alamouti scheme is

$$C_A = \int_0^{\infty} 4Ku_1e^{-2u_1}\left(1-(1+2u_1)e^{-2u_1}\right)^{K-1}\log(1+u_1)du_1.$$

For the single antenna case, the effective SNR is $u_2 = \max_{k=1,\ldots,K} |h_k|^2$, and

$$\mathbb{P}\left(\max_{k=1,\ldots,K} |h_k|^2 < x\right) = \left(1-e^{-x}\right)^K,$$

so

$$f_2(u_2) = Ke^{-u_2}\left(1-e^{-u_2}\right)^{K-1}.$$

Thus the capacity for the single antenna case is

$$C_S = \int_0^{\infty} Ke^{-u_2}\left(1-e^{-u_2}\right)^{K-1}\log(1+u_2)du_2.$$

The achievable throughput under both schemes at SNR=0dB for different number of users is shown in Figure 6.3.

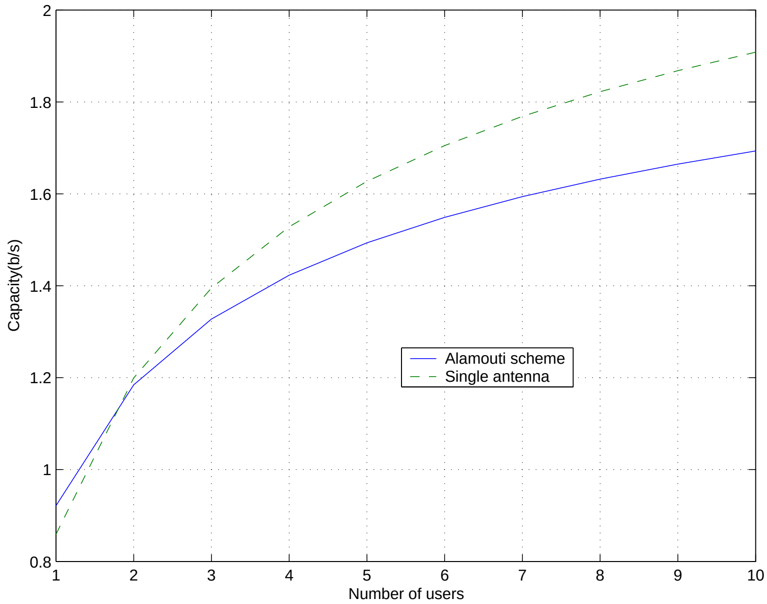


Figure 6.3: Throught vs. number of users

4. We can see from the plot in part (3) that single antenna scheme performs better than the Alamouti scheme when $K \geq 2$. So in this case we do not need use dual transmit antenna. This is because the probability of getting one good channel is larger than getting two good channels.

Chapter 7

Solutions to Exercises

EXERCISE 7.1. 1. Refer to Figure 7.3(a) in the textbook, we have

$$d_i = \left(d^2 + ((i-1)\Delta_r\lambda_c)^2 + 2(i-1)d\Delta_r\lambda_c cos\phi\right)^{1/2}$$
$$= d\left(1 + \frac{((i-1)\Delta_r\lambda_c)^2}{d^2} + \frac{2(i-1)\Delta_r\lambda_c cos\phi}{d}\right)^{1/2}.$$

For large d (large compare to the size of receiver antenna array), we can expand the above equation to the first order term and get

$$d_i \simeq d\left(1 + \frac{1}{2}\frac{2(i-1)\Delta_r\lambda_c cos\phi}{d}\right) = d + (i-1)\Delta_r\lambda_c cos\phi,$$

which is equation (7.19).

2. From Figure 7.1 we have

$$d_{ik}^2 = [d - (k-1)\Delta_t\lambda_c\cos\phi_t + (i-1)\Delta_r\lambda_c\cos\phi_r]^2$$
$$+ [(k-1)\Delta_t\lambda_c\sin\phi_t - (i-1)\Delta_r\lambda_c\sin\phi_r]^2$$
$$\simeq d^2 - 2d(k-1)\Delta_t\lambda_c\cos\phi_t + 2d(i-1)\Delta_r\lambda_c\cos\phi_r$$

for large d. Hence

$$d_{ik} \simeq d\left[1 - \frac{1}{d}(k-1)\Delta_t\lambda_c\cos\phi_t + \frac{1}{d}(i-1)\Delta_r\lambda_c\cos\phi_r\right]^{1/2}$$
$$= d - (k-1)\Delta_t\lambda_c\cos\phi_t + (i-1)\Delta_r\lambda_c\cos\phi_r,$$

which is equation (7.27).

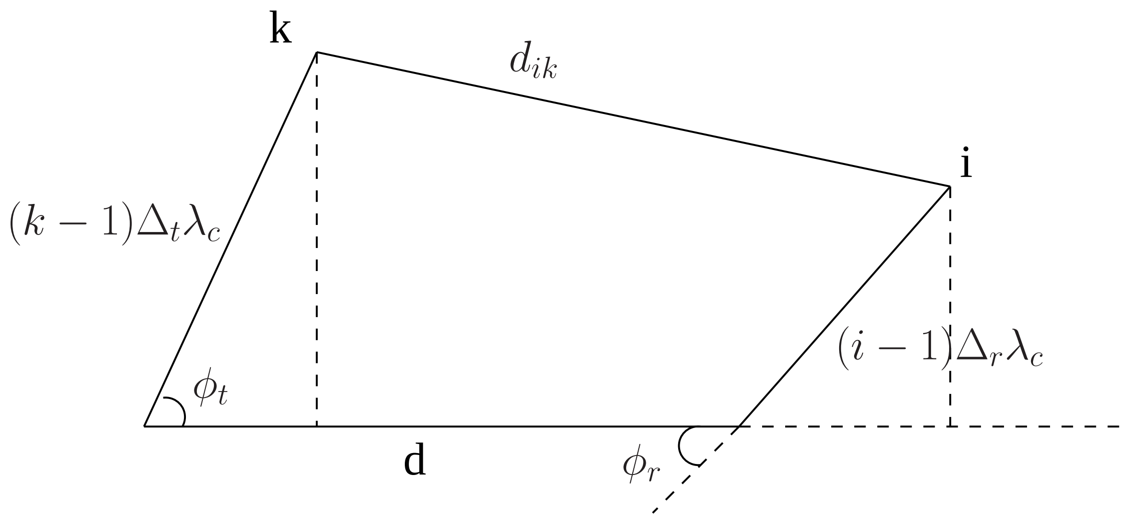


Figure 7.1: MIMO case

Exercise 7.2.

$$\mathbf{e_r}(\Omega) = \frac{1}{\sqrt{n_r}} \begin{bmatrix} 1 \\ \exp(-j2\pi\Delta_r\Omega) \\ \exp(-j2\pi 2\Delta_r\Omega) \\ \vdots \\ \exp(-j2\pi(n_r-1)\Delta_r\Omega) \end{bmatrix}.$$

We have

$$\mathbf{e_r}\left(\Omega + \frac{1}{\Delta_r}\right) = \frac{1}{\sqrt{n_r}} \begin{bmatrix} 1 \\ \exp(-j2\pi\Delta_r\Omega)\exp(-j2\pi) \\ \exp(-j2\pi 2\Delta_r\Omega)\exp(-j4\pi) \\ \vdots \\ \exp(-j2\pi(n_r-1)\Delta_r\Omega)\exp(-j2(n_r-1)\pi) \end{bmatrix} = \mathbf{e_r}(\Omega).$$

Hence $\mathbf{e_r}(\Omega)$ is periodic with period $\frac{1}{\Delta_r}$.

Next suppose there exists a $\theta \in \left(0, \frac{1}{\Delta_r}\right)$ such that

$$\mathbf{e_r}(\Omega + \theta) = \frac{1}{\sqrt{n_r}} \begin{bmatrix} 1 \\ \exp(-j2\pi\Delta_r\Omega)\exp(-j2\pi\Delta_r\theta) \\ \exp(-j2\pi 2\Delta_r\Omega)\exp(-j4\pi\Delta_r\theta) \\ \vdots \\ \exp(-j2\pi(n_r-1)\Delta_r\Omega)\exp(-j2(n_r-1)\pi\Delta_r\theta) \end{bmatrix} = \mathbf{e_r}(\Omega),$$

and we must have $\exp(-j2\pi\Delta_r\theta) = 1$. Clearly there is no $\theta \in \left(0, \frac{1}{\Delta_r}\right)$ such that $\exp(-j2\pi\Delta_r\theta) = 1$. So the smallest period is $\frac{1}{\Delta_r}$.

EXERCISE 7.3.

$$
\begin{aligned}
f_r(\Omega_r) &= f_r(\Omega_{r2} - \Omega_{r1}) = e_r(\Omega_{r1})^* e_r(\Omega_{r2}) \\
&= \frac{1}{n_r} [1 \ \exp(j2\pi\Delta_r\Omega_{r1}) \ \exp(j2\pi 2\Delta_r\Omega_{r1}) \ \ldots \ \exp(j2\pi(n_r-1)\Delta_r\Omega_{r1})] \\
&\quad \times \begin{bmatrix} 1 \\ \exp(-j2\pi\Delta_r\Omega_{r2}) \\ \exp(-j2\pi 2\Delta_r\Omega_{r2}) \\ \vdots \\ \exp(-j2\pi(n_r-1)\Delta_r\Omega_{r2}) \end{bmatrix} \\
&= \frac{1}{n_r} [1 + \exp(-j2\pi\Delta_r\Omega_r) + \exp(-j2\pi 2\Delta_r\Omega_r) + \cdots + \exp(-j2\pi(n_r-1)\Delta_r\Omega_r)] \\
&= \frac{1}{n_r} \frac{1 - \exp(-j2\pi n_r\Delta_r\Omega_r)}{1 - \exp(-j2\pi\Delta_r\Omega_r)} \\
&= \frac{\exp(-j\pi n_r\Delta_r\Omega_r)}{n_r \exp(-j\pi\Delta_r\Omega_r)} \times \frac{\exp(j\pi n_r\Delta_r\Omega_r) - \exp(-j\pi n_r\Delta_r\Omega_r)}{\exp(j\pi\Delta_r\Omega_r) - \exp(-j\pi\Delta_r\Omega_r)} \\
&= \frac{\exp(-j\pi(n_r-1)\Delta_r\Omega_r)}{n_r} \times \frac{\sin(\pi n_r\Delta_r\Omega_r)}{\sin(\pi\Delta_r\Omega_r)} \\
&= \frac{\exp(-j\pi(n_r-1)\Delta_r\Omega_r)}{n_r} \times \frac{\sin(\pi L_r\Omega_r)}{\sin(\pi L_r\Omega_r/n_r)},
\end{aligned}
$$

which is equation (7.35).

EXERCISE 7.4. The degree of freedom of MIMO channel depends on the angular spread of the scatters/reflectors, and depends on the antenna array length. There are two problems of (7.82).

1. The degree of freedom depends on the antenna array length, which determines the resolution of the channel. Simply increasing number of antenna does not necessarily increase channel resolution.

2. The number of multipath does not have direct impact on the degree of freedom, it is the angular spread that influences the degree of freedom.

EXERCISE 7.5.

EXERCISE 7.6.

EXERCISE 7.7.

Exercise 7.8.

Exercise 7.9.

Exercise 7.10.

Exercise 7.11.

Chapter 8

Solutions to Exercises

EXERCISE 8.1. Consider the singular value of decomposition:

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}.$$

Then, the channel model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w},$$

can be rewritten as

$$\tilde{\mathbf{y}} = \tilde{\mathbf{\Lambda}}\tilde{\mathbf{x}} + \mathbf{w},$$

where $\tilde{\mathbf{y}} = \mathbf{U}^*\mathbf{y}$ and $\tilde{\mathbf{x}} = \mathbf{V}\mathbf{x}$. Moreover, $\tilde{\mathbf{x}}$ has the same total power constraint as $\mathbf{x}$ since $\mathbf{V}$ is a unitary matrix. Thus, the channel capacity depends only on the singular values of $\mathbf{H}$ for a total power constraint. But, $\mathbf{H}^*$ has the same non-zero singular values as $\mathbf{H}$ and hence the capacity of the reciprocal channel is same as the original channel.

EXERCISE 8.2.

EXERCISE 8.3.

EXERCISE 8.4. 1. Let $\mathbf{A} = \mathbf{H}\mathbf{U}$. Then, $\mathbf{A}$ is zero mean. Let $\mathbf{a}_i$ and $\mathbf{h}_i$ denote the ith columns of $\mathbf{A}$ and $\mathbf{H}$ respectively. Then,

$$\begin{aligned} E[\mathbf{a}_j^*\mathbf{a}_i] &= \mathbf{U}^*E[\mathbf{h}_j^*\mathbf{h}_i]\mathbf{U}, \\ &= \mathbf{U}^*\delta_{ij}\mathbf{I}\mathbf{U}, \\ &= \delta_{ij}\mathbf{I}. \end{aligned}$$

Also,

$$\begin{aligned} E[\mathbf{a}_j^t\mathbf{a}_i] &= \mathbf{U}^tE[\mathbf{h}_j^t\mathbf{h}_i]\mathbf{U}, \\ &= \mathbf{0}. \end{aligned}$$

Thus, $\mathbf{A}$ has i.i.d. $\mathcal{CN}(0,1)$ entries.

2. Consider the eigenvalue decomposition of $\mathbf{K}_x$:

$$\mathbf{K}_x = \mathbf{U}\mathbf{D}\mathbf{U}^*,$$

where $\mathbf{D}$ is a positive diagonal matrix and $\mathbf{U}$ is a unitary matrix. Then, the mutual information can be written as:

$$\begin{aligned} E_{\mathbf{H}}\left[\log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{K}_x\mathbf{H}^*\right)\right] &= E_{\mathbf{H}}\left[\log\det\left(\mathbf{I} + \frac{1}{N_0}(\mathbf{H}\mathbf{U})\mathbf{D}(\mathbf{H}\mathbf{U})^*\right)\right], \\ &= E_{\mathbf{H}\mathbf{U}}\left[\log\det\left(\mathbf{I} + \frac{1}{N_0}(\mathbf{H}\mathbf{U})\mathbf{D}(\mathbf{H}\mathbf{U})^*\right)\right], \\ &= E_{\mathbf{H}}\left[\log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{D}\mathbf{H}^*\right)\right], \end{aligned}$$

where the last step follows from the first part. Moreover, we also have $\text{Trace}(\mathbf{K}_x) = \text{Trace}(\mathbf{D})$. Thus, the input covariance matrix can be restricted to be a diagonal matrix.

3. Note that $\log\det(\mathbf{X})$ is a concave function of $\mathbf{X}$ for positive definite $\mathbf{X}$ (see Page 74 of *Convex Optimization* by Stephen Boyd and L. Vandenberghe). Also $\mathbf{I} + \mathbf{H}\mathbf{D}\mathbf{H}^*$ is linear in $\mathbf{D}$. Thus, the function $\log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{D}\mathbf{H}^*\right)$ is concave in $\mathbf{D}$. Moreover, the function is symmetric in the diagonal entries of $\mathbf{D}$: this follows from the fact that reordering rows of $\mathbf{H}$ does not change the distribution of $\mathbf{H}$. Thus, for a trace constraint on $\mathbf{D}$ the mutual information is maximized when $\mathbf{D}$ is a multiple of the identity matrix.

EXERCISE 8.5. 1. Multiple receive antennas provide a power gain which is very crucial for a wideband CDMA system which works at low SNR. However, the gain is not very crucial for a narrowband GSM which works at high SNR. The gain can be crucial for a wideband OFDM system if the mobile device is on the boundary of a cell and hence is working at low SNR.

2. Multiple transmit antennas can provide a degree of freedom gain for the high SNR narrowband GSM system. However for a wideband CDMA system, the transmit antennas will not be useful as the optimal coding strategy will be to use a single transmit antenna at low SNR. Similar arguments work for the wideband OFDM system depending which SNR regime it is working in.

EXERCISE 8.6.

EXERCISE 8.7.

EXERCISE 8.8. 1. The pilot signal with transmit power SNR will estimate the channel $\mathbf{H}$ with an mean squared estimation error of order of $\frac{1}{\mathsf{SNR}}$. This estimation error in the channel multiplied by the transmitted signal will act as an additive noise during rest of the communication. Thus, while communicating, the effective noise seen by the receiver is the additive Gaussian noise and the noise due to estimation error of $\mathbf{H}$ which is of the order $\mathsf{SNR}\frac{1}{\mathsf{SNR}} = 1$. Thus, for the effective coherent channel, the total additive noise has a bounded variance for any SNR. Thus, for $k \times n_r$ channel, assuming that effective noise is Gaussian (worst-case assumption) the channel capacity for the effective coherent channel is lower bounded by $\min(k, n_r) \log \mathsf{SNR}$. But since k time slots were used for the pilot scheme, the effective rate of communication is given by at least

$$\frac{T_c - k}{T_c} \min(k, n_r) \log \mathsf{SNR} \quad \text{bits/s/Hz}$$

.

2. With n_t transmit antennas, we only need at most n_t time slots for training. Additional training time can only hurt the overall rate at high SNR. Thus, we have $k \leq n_t$. Now, if $n_r \leq n_t$, then we only use n_r *transmit* antennas. As other antennas will not provide a degree of freedom gain. Thus we also have $k \leq \min(n_r, n_t)$. Now, if $k \leq \min(n_r, n_t)$, $\frac{T_c - k}{T_c} k$ is increasing in k for $k < T_c/2$ and decreasing for $k > T_c/2$. Thus, the optimal value of k is given by:

$$k^* = \min(n_r, n_t, T_c/2).$$

EXERCISE 8.9. Consider the channel from a particular transmit antenna i to a particular receive antenna j. This channel can be modeled as a simple scalar ISI channel with tap coefficients $\mathbf{H}_l(i,j)$. For this channel, the usual scalar OFDM scheme will yield N_c tones define by:

$$\tilde{\mathbf{H}}_n(i,j) = \sum_{l=0}^{L-1} \mathbf{H}_l(i,j) e^{-\frac{j2\pi nl}{N_c}}.$$

Now, for the original MIMO channel, we can look at each receive antenna separately. For each receive antenna, signals transmitted from different transmit antennas add linearly at the receiver. Since the OFDM scheme is a linear operation, the overall effective OFDM channel for one particular receive antenna can be written as the sum of the individual OFDM channels. Now, for the original MIMO channel, since for each receive antenna, the OFDM scheme is the same, we get that the overall OFDM scheme can be written as:

$$\tilde{\mathbf{H}}_n = \sum_{l=0}^{L-1} \mathbf{H}_l e^{-\frac{j2\pi nl}{N_c}}.$$

EXERCISE 8.10. For a fixed physical environment (i.e., fixed $\mathbf{H}$), the capacity for a total power constraint P is given by

$$\max_{\text{Trace}(\mathbf{K}_x)\leq P} W \log\det\left(\mathbf{I} + \frac{1}{N_o W}\mathbf{H}\mathbf{K}_x\mathbf{H}^*\right) \quad \text{bits/sec.}$$

This can be rewritten as:

$$\max_{\text{Trace}(\frac{\mathbf{K}_x}{P})\leq 1} W \log\det\left(\mathbf{I} + \frac{P}{N_o W}\mathbf{H}\frac{\mathbf{K}_x}{P}\mathbf{H}^*\right) \quad \text{bits/sec.}$$

When, both P and W are both doubled, the corresponding capacity can be written as:

$$2 \max_{\text{Trace}(\frac{\tilde{\mathbf{K}}_x}{2P})\leq 1} W \log\det\left(\mathbf{I} + \frac{P}{N_o W}\mathbf{H}\frac{\tilde{\mathbf{K}}_x}{2P}\mathbf{H}^*\right) \quad \text{bits/sec.}$$

Since the two optimization problems are essentially the same, the optimal solution for the second case is given by $\tilde{\mathbf{K}}_x^* = 2\mathbf{K}_x^*$. Therefore the capacity is exactly doubled.

EXERCISE 8.11.

EXERCISE 8.12.

EXERCISE 8.13.

EXERCISE 8.14.

EXERCISE 8.15.

EXERCISE 8.16. The general capacity expression is given by:

$$C = E\left[\log\det\left(\mathbf{I}_{n_r} + \frac{\mathsf{SNR}}{n_t}\mathbf{H}\mathbf{H}^*\right)\right].$$

The apparent paradox is because of the behavior of $\log\det\left(\mathbf{I}_{n_r} + \frac{\mathsf{SNR}}{n_t}\mathbf{H}\mathbf{H}^*\right)$. At low values of SNR it behaves like $\log\left(\text{Trace}\left(\frac{\mathsf{SNR}}{n_t}\mathbf{H}\mathbf{H}^*\right)\right)$, whereas at medium SNR it behaves like $\log\det\left(\frac{\mathsf{SNR}}{n_t}\mathbf{H}\mathbf{H}^*\right)$. Thus, we have the following consistent behavior:

	C_{1n}	C_{nn}
low SNR	$n\mathsf{SNR}$	nSNR
medium SNR	$\log\mathsf{SNR} + \log n$	$n\log\mathsf{SNR}$

EXERCISE 8.17.

EXERCISE 8.18.

Exercise 8.19.

Exercise 8.20. For matched filters, the interference seen by the kth stream is given by:

$$\sum_{i \neq k} \frac{\mathbf{h}_k^* \mathbf{h}_i}{||\mathbf{h}_k||} x_i + \mathbf{h}_k^* \mathbf{w}.$$

Thus, the variance of the interference seen by the kth stream is approximately given by:

$$\sum_{i \neq k} \frac{\mathbf{h}_k^* \mathbf{h}_i}{||\mathbf{h}_k||} P_i + N_0.$$

Therefore, the rate for the kth stream is given by (assuming the worst-case assumption that the interference is Gaussian):

$$\begin{aligned} R_k &= E\left[\log\left(1 + \frac{P_K ||\mathbf{h}_k||^2}{\sum_{i \neq k} \frac{\mathbf{h}_k^* \mathbf{h}_i}{||\mathbf{h}_k||} P_i + N_0}\right)\right], \\ &= E\left[\log\left(1 + \frac{\frac{P_K}{N_0} ||\mathbf{h}_k||^2}{\sum_{i \neq k} \frac{\mathbf{h}_k^* \mathbf{h}_i}{||\mathbf{h}_k||} \frac{P_i}{N_0} + 1}\right)\right], \\ &\approx E\left[\log\left(1 + \frac{P_K}{N_0} ||\mathbf{h}_k||^2\right)\right], \\ &\approx E\left[\frac{P_K}{N_0} ||\mathbf{h}_k||^2\right], \end{aligned}$$

where the last two steps follow from the low SNR assumption. Thus, the total sum-rate is given by:

$$\begin{aligned} \sum_k R_k &= \frac{\mathsf{SNR}}{n_t} \sum_k E\left[||\mathbf{h}_k||^2\right], \\ &= n_r \mathsf{SNR}, \end{aligned}$$

which at low SNR is the capacity of a $1 \times n_r$ channel (see Soln 8.16).

Exercise 8.21.

Exercise 8.22. We use the following matrix identity which follows from the matrix inversion lemma:

$$\log|\mathbf{A} + \mathbf{x}\mathbf{x}^*| - \log|\mathbf{A}| = \log(1 + \mathbf{x}^* \mathbf{A}^{-1} \mathbf{x}). \tag{8.1}$$

Taking $\mathbf{x} = \sqrt{\frac{P_1}{N_0}}\mathbf{h}_1$ implies:

$$\log\det\left(\mathbf{I} + \sum_{i=1}^{n_t} \frac{P_i}{N_0}\mathbf{h}_i\mathbf{h}_i^*\right) - \log(1+\mathrm{SINR}_1) \;=\; \log\det\left(\mathbf{I} + \sum_{i=2}^{n_t} \frac{P_i}{N_0}\mathbf{h}_i\mathbf{h}_i^*\right).$$

Similarly, taking $\mathbf{x} = \sqrt{\frac{P_k}{N_0}}\mathbf{h}_k$ we get:

$$\log\det\left(\mathbf{I} + \sum_{i=k}^{n_t} \frac{P_i}{N_0}\mathbf{h}_i\mathbf{h}_i^*\right) - \log(1+\mathrm{SINR}_k) \;=\; \log\det\left(\mathbf{I} + \sum_{i=k+1}^{n_t} \frac{P_i}{N_0}\mathbf{h}_i\mathbf{h}_i^*\right).$$

Adding all such equations, we get:

$$\log\det\left(\mathbf{I} + \sum_{i=1}^{n_t} \frac{P_i}{N_0}\mathbf{h}_i\mathbf{h}_i^*\right) \;=\; \sum_{i=1}^{n_t}\log(1+\mathrm{SINR}_i).$$

Thus, taking $\mathbf{K}_x$ to be a diagonal matrix with entries P_i we get:

$$\log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{K}_x\mathbf{H}_i^*\right) \;=\; \sum_{i=1}^{n_t}\log(1+\mathrm{SINR}_i).$$

Exercise 8.23. We have the following sequence of steps:

$$\begin{aligned}
p_{\text{out}}(R) &\overset{(a)}{\geq} \mathbb{P}\left\{\log\det\left(\mathbf{I}_{n_r} + \mathsf{SNR}\mathbf{H}\mathbf{H}^*\right) < R\right\},\\
&\overset{(b)}{\geq} \mathbb{P}\left\{\mathsf{SNR}\mathrm{Tr}[\mathbf{H}\mathbf{H}^*] < R\right\},\\
&\overset{(c)}{\geq} \mathbb{P}\left\{\mathsf{SNR}|h_{11}|^2 < \frac{R}{n_r n_t}\right\}^{n_r n_t},\\
&\overset{(d)}{=} \left(1 - e^{-\frac{R}{n_r n_t \mathsf{SNR}}}\right)^{n_r n_t},\\
&\overset{(e)}{\approx} \frac{R^{n_t n_R}}{(n_r n_t \mathsf{SNR})^{n_t n_r}}.
\end{aligned}$$

Each of these steps can be justified as follows:

- (a): follows from letting each antenna power be SNR rather than SNR/n_t.
- (b): follows from the equation: $\mathsf{SNR}\mathrm{Tr}[\mathbf{H}\mathbf{H}^*] < \det\left(\mathbf{I}_{n_r} + \mathsf{SNR}\mathbf{H}\mathbf{H}^*\right)$ and hence a simple set theoretic containment relationship.
- (c): again follows from a simple set theoretic containment relationship:

$$\left\{\mathsf{SNR}|h_{ij}|^2 < \frac{R}{n_r n_t}\ \forall\, i,j\right\} \subset \left\{\mathsf{SNR}\mathrm{Tr}[\mathbf{H}\mathbf{H}^*] < R\ \forall\, i,j\right\},$$

  and the facts that h_{ij}s are i.i.d.

- (d): follows from the fact $|h_{11}|^2$ is exponential.
- (e): follows from a simple Taylor series expansion.

EXERCISE 8.24. At high SNR, MMSE-SIC receiver is same as the decorrelator followed by SIC. At high SNR, for the first stream, a decorrelator projects n_r dimensional receive vector along a sub-space orthogonal to $n_t - 1$ other directions. Thus, the diversity seen by the first stream will be $n_r - (n_t - 1) = n_r - n_t + 1$. Decoding of this stream in fact will be bottleneck for all other streams. Thus, for each stream the diversity is given by $n_r - n_t + 1$. However, for the kth stream, if all the previous streams have been decoded correctly, then the diversity seen by the kth stream is given by $n_t - n_t + k$ (projection of n_r dimensional vector onto a sub-space orthogonal to a $n_t - k$ dimensional sub-space).

EXERCISE 8.25. 1. From MMSE estimation of streams, we have

$$\mathsf{SNR}|g_1|^2 = \mathbf{h}_1^* \left(\mathbf{I}/\mathsf{SNR} + \mathbf{h}_2\mathbf{h}_2^*\right)^{-1} \mathbf{h}_1.$$

Using the matrix inversion lemma we get:

$$\begin{aligned} |g_1|^2 &= \mathbf{h}_1^* \left(\mathbf{I} - \frac{\mathsf{SNR}\mathbf{h}_2\mathbf{h}_2^*}{1 + \mathsf{SNR}||\mathbf{h}_2||^2}\right) \mathbf{h}_1, \\ &= ||\mathbf{h}_1||^2 - \frac{\mathsf{SNR}||\mathbf{h}_1^*\mathbf{h}_2||^2}{1 + \mathsf{SNR}||\mathbf{h}_2||^2}. \end{aligned}$$

Now, consider:

$$\begin{aligned} ||\mathbf{h}_{1\perp 2}||^2 + \frac{||\mathbf{h}_{1||2}||^2}{1 + \mathsf{SNR}||\mathbf{h}_2||^2} &= ||\mathbf{h}_1||^2 - ||\mathbf{h}_{1||2}||^2 + \frac{||\mathbf{h}_{1||2}||^2}{1 + \mathsf{SNR}||\mathbf{h}_2||^2}, \\ &= ||\mathbf{h}_1||^2 - \frac{\mathsf{SNR}||\mathbf{h}_2||^2||\mathbf{h}_{1||2}||^2}{1 + \mathsf{SNR}||\mathbf{h}_2||^2}, \\ &= ||\mathbf{h}_1||^2 - \frac{\mathsf{SNR}||\mathbf{h}_1^*\mathbf{h}_2||^2}{1 + \mathsf{SNR}||\mathbf{h}_2||^2}, \end{aligned}$$

which matches with the expression above for $|g_1|^2$. Thus, we have

$$|g_1|^2 = ||\mathbf{h}_{1\perp 2}||^2 + \frac{||\mathbf{h}_{1||2}||^2}{1 + \mathsf{SNR}||\mathbf{h}_2||^2}$$

The fact that $|g_2|^2 = ||\mathbf{h}_2||^2$ follows directly since the first symbol doesn't see any interference.

2. At high SNR, the second term is small with high probability, thus the marginal distribution of $|g_1|^2$ is same as $||\mathbf{h}_{1\perp 2}||^2$. Now taking $\mathbf{h}_2$ as a basis vector, we see

that $\mathbf{h}_{1\perp 2}$ is always orthogonal to a basis vector. Thus it is a projection of $\mathbf{h}_1$ onto one dimension. Thus, statistically it should be similar to simple complex Gaussian. Note, that by circular symmetry of $\mathbf{h}_1$, the fact that $\mathbf{h}_2$ is along a random direction does not change the statistics. Since, $||\mathbf{h}_{1\perp 2}||^2$ is exponential, $|g_1|^2$ is marginally exponential at high SNR. Moreover, we have:

$$|g_1|^2 = ||\mathbf{h}_{1\perp 2}||^2 + \frac{||\mathbf{h}_{1||2}||^2}{1 + \mathsf{SNR}|g_2|^2},$$

where $\mathbf{h}_{1\perp 2}$ and $\mathbf{h}_{1||2}$ are independent of $|g_2|$. Thus, we see that $|g_1|$ and $|g_2|$ are negatively correlated.

3. The maximum diversity given by the parallel channel is same as the original MIMO channel since the D-BLAST structure preserves mutual information and hence the outage behavior. Thus, the total diversity is given by 4.

4. If $|g_1|^2$ and $|g_2|^2$ were independent with the same marginals, then the diversity offered by $|g_1|^2$ is 1 and that offered by $|g_2|^2$ is 2. Thus, the total diversity is given by 3.

Exercise 8.26. The coding scheme can be written as

$$\begin{bmatrix} 0 & \cdots & 0 & p_1^{(1)} & p_1^{(2)} & \cdots & p_1^{(T-n_t+1)} \\ \vdots & \iddots & \iddots & \iddots & \cdots & \iddots & \vdots \\ 0 & p_{n_t-1}^{(1)} & p_{n_t-1}^{(2)} & \iddots & & \iddots & 0 \\ p_{n_t}^{(1)} & p_{n_t}^{(2)} & \cdots & \cdots & & 0 & 0 \end{bmatrix},$$

where $P^{(k)} = \left[p_1^{(k)}, \ldots, p_{n_t}^{(k)}\right]$ are the independent data streams. The decoding can be done using successive interference cancelation: estimate stream $P^{(k)}$ one by one, then jointly decode it and then estimate $P^{(k+1)}$ after canceling out $P^{(k)}$.

Exercise 8.27. For an n_t transmit antenna channel a D-BLAST scheme with block-length T has a rate loss of:

$$\frac{n_t - 1}{T},$$

because a zero is sent on the first transmit antenna for the first $n_t - 1$ time slots. So instead of sending T streams, we send only $T - (n_t - 1)$ streams.

Chapter 9

Solutions to Exercises

Exercise 9.1.

Exercise 9.2.

Exercise 9.3.

Exercise 9.4.

Exercise 9.5.

Exercise 9.6.

Exercise 9.7.

Exercise 9.8.

Exercise 9.9.

Exercise 9.10.

Exercise 9.11.

Exercise 9.12.

Exercise 9.13.

Exercise 9.14.

Exercise 9.15.

Exercise 9.16.

Exercise 9.17.

Exercise 9.18.

Exercise 9.19.

Exercise 9.20.

Exercise 9.21.

Exercise 9.22.

Exercise 9.23.

Exercise 9.24.

Exercise 9.25.

Chapter 10

Solutions to Exercises

EXERCISE 10.1.

EXERCISE 10.2.

EXERCISE 10.3.

EXERCISE 10.4.

EXERCISE 10.5.

EXERCISE 10.6.

EXERCISE 10.7.

EXERCISE 10.8.

EXERCISE 10.9.

EXERCISE 10.10.

EXERCISE 10.11.

EXERCISE 10.12.

EXERCISE 10.13.

EXERCISE 10.14.

EXERCISE 10.15.

EXERCISE 10.16.

EXERCISE 10.17.

EXERCISE 10.18.

EXERCISE 10.19.

EXERCISE 10.20.

EXERCISE 10.21.

EXERCISE 10.22.

EXERCISE 10.23.

EXERCISE 10.24.

EXERCISE 10.25.

Appendix A

Solutions to Exercises

EXERCISE A.1. 1. $n = 1$. Let $A = w^2$, then using the formula for the density of a function of a random variable, we get:

$$
\begin{aligned}
f_1(a) &= \frac{f_w(\sqrt{a})}{2\sqrt{a}} + \frac{f_w(-\sqrt{a})}{2\sqrt{a}} \\
&= = \frac{1}{\sqrt{2\pi a}} \exp\left(-a/2\right)
\end{aligned}
$$

2. Let $\Phi_n(\omega)$ denote the characteristic function of $||\mathbf{w}||^2$. Then since convolution corresponds to multiplication of characteristic functions, we get

$$
\begin{aligned}
\Phi_n(\omega) &= \Phi_1(\omega)^n \\
&= \Phi_2(\omega)^{n/2} \\
&= \left(\frac{1}{1-2j\omega}\right)^{n/2},
\end{aligned}
$$

where the last step follows from the fact for $n = 2$, $||\mathbf{w}||^2$ is an exponential random variable. Then, from this we see that

$$
\begin{aligned}
\frac{d\Phi_n(\omega)}{d(j\omega)} &= n\left(\frac{1}{1-2j\omega}\right)^{n/2+1} \\
&= n\Phi_{n+2}(\omega).
\end{aligned}
$$

Since differentiation corresponds to multiplication in time domain, we get

$$
f_{n+2}(a) = \frac{a}{n} f_n(a).
$$

3. Using simple recursion:

$$
f_n(a) = \frac{1}{\sqrt{2\pi}} \frac{a^{n/2-1}}{1 \cdot 3 \cdots (n-2)} \exp(-a/2) \quad \text{for } n \text{ odd}
$$

$$f_n(a) = \frac{1}{2}\frac{a^{n/2-1}}{2 \cdot 4 \cdots (n-2)} \exp(-a/2) \quad \text{for } n \text{ even}$$

EXERCISE A.2. $(z_i)_{i=1}^M$ is a linear transformation of a Gaussian process, so it has a jointly Gaussian distribution, which is completely specified by the first and second moments.

$$E[z_i] = E\left[\int_{-\infty}^{\infty} w(t)s_i(t)dt\right] = \int_{-\infty}^{\infty} E[w(t)]s_i(t)dt = 0$$

$$\begin{aligned} E[z_i z_j] &= E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(t)w(\tau)s_i(t)s_j(\tau)dtd\tau\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{N_0}{2}\delta(t-\tau)s_i(t)s_j(\tau)dtd\tau \\ &= \int_{-\infty}^{\infty} \frac{N_0}{2} s_i(\tau)s_j(\tau)d\tau = \frac{N_0}{2}\delta_{i,j} \end{aligned}$$

Therefore $E[\mathbf{z}\mathbf{z}^T] = (N_0/2)\mathbf{I}_M$ and $\mathbf{z} \sim \mathbf{N}(\mathbf{0}, (N_0/2)\mathbf{I}_M)$.

EXERCISE A.3. For simplicity, let us assume that $\mathbf{x}$ is zero mean.

1. The covariance matrix, $\mathbf{K}$, of $\mathbf{x}$ is given by

$$\begin{aligned} \mathbf{K} &= E[\mathbf{x}\mathbf{x}^*] \\ &= E[(\mathcal{R}[\mathbf{x}] + j\mathcal{C}[\mathbf{x}])(\mathcal{R}[\mathbf{x}]^t - j\mathcal{C}[\mathbf{x}]^t)] \\ &= E[\mathcal{R}[\mathbf{x}]\mathcal{R}[\mathbf{x}]^t] + E[\mathcal{C}[\mathbf{x}]\mathcal{C}[\mathbf{x}]^t] - jE[\mathcal{R}[\mathbf{x}]\mathcal{C}[\mathbf{x}]^t] + jE[\mathcal{C}[\mathbf{x}]\mathcal{R}[\mathbf{x}]^t]. \end{aligned}$$

   Similarly, the pseudo-covariance matrix of $\mathbf{x}$ is given by

$$\mathbf{J} = E[\mathcal{R}[\mathbf{x}]\mathcal{R}[\mathbf{x}]^t] - E[\mathcal{C}[\mathbf{x}]\mathcal{C}[\mathbf{x}]^t] + jE[\mathcal{R}[\mathbf{x}]\mathcal{C}[\mathbf{x}]^t] + jE[\mathcal{C}[\mathbf{x}]\mathcal{R}[\mathbf{x}]^t].$$

   The covariance of matrix of $[\mathcal{R}[\mathbf{x}], \mathcal{C}[\mathbf{x}]]^t$ is given by

$$\begin{bmatrix} E[\mathcal{R}[\mathbf{x}]\mathcal{R}[\mathbf{x}]^t] & E[\mathcal{R}[\mathbf{x}]\mathcal{C}[\mathbf{x}]^t] \\ E[\mathcal{C}[\mathbf{x}]\mathcal{R}[\mathbf{x}]^t] & E[\mathcal{C}[\mathbf{x}]\mathcal{C}[\mathbf{x}]^t] \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \mathcal{R}(K+J) & \mathcal{C}(J-K) \\ \mathcal{C}(K+J) & \mathcal{R}(K-J) \end{bmatrix} \quad \text{(A.1)}$$

2. For a circularly symmetric $\mathbf{x}$, $\mathbf{J} = \mathbf{0}$ and the covariance of matrix of $[\mathcal{R}[\mathbf{x}], \mathcal{C}[\mathbf{x}]]^t$ is given by

$$\begin{bmatrix} E[\mathcal{R}[\mathbf{x}]\mathcal{R}[\mathbf{x}]^t] & E[\mathcal{R}[\mathbf{x}]\mathcal{C}[\mathbf{x}]^t] \\ E[\mathcal{C}[\mathbf{x}]\mathcal{R}[\mathbf{x}]^t] & E[\mathcal{C}[\mathbf{x}]\mathcal{C}[\mathbf{x}]^t] \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \mathcal{R}(K) & -\mathcal{C}(K) \\ \mathcal{C}(K) & \mathcal{R}(K) \end{bmatrix} \quad \text{(A.2)}$$

EXERCISE A.4. 1. Necessity of the two conditions is proved in appendix A. For proving sufficiency, let $\mathbf{y} = e^{j\theta}\mathbf{x}$, then

$$\begin{aligned} E[\mathbf{y}] &= e^{j\theta} E[\mathbf{x}] \\ &= \mathbf{0}. \end{aligned}$$

Then the pseudo-covariance of $\mathbf{y}$ is given by

$$\begin{aligned} E[\mathbf{y}\mathbf{y}^t] &= e^{2j\theta} E[\mathbf{x}\mathbf{x}^t] \\ &= 0. \end{aligned}$$

The covariance of $\mathbf{y}$ is given by

$$\begin{aligned} E[\mathbf{y}\mathbf{y}^x] &= e^{j\theta} E[\mathbf{x}\mathbf{x}^*] e^{-j\theta} \\ &= E[\mathbf{x}\mathbf{x}^*]. \end{aligned}$$

Thus, $\mathbf{y}$ and $\mathbf{x}$ have the same second order statistic and hence have identical distribution.

2. Since $\mathbf{x}$ is not given to be zero mean, the answer is no. But in addition if we assume that $\mathbf{x}$ is zero mean, then from (A.1) and (A.2) we see that $\mathbf{J}$ must be zero and hence $\mathbf{x}$ will be circularly symmetric.

EXERCISE A.5. Let $x = x_r + jx_i$. Then x_r and x_i are zero mean and are jointly Gaussian. Since the pseudocovariance for a circularly symmetric Gaussian is zero, we get

$$\begin{aligned} 0 &= E[x^2] \\ &= E[x_r^2] - E[x_i^2] + 2jE[x_r x_i]. \end{aligned}$$

Thus, $E[x_r^2] = E[x_i^2]$ and $E[x_r x_i] = 0$. For jointly Gaussian random variables, uncorrelated implies independent. Also, since they are zero mean and have the same second moment, x_r and x_i are i.i.d. random variables.

EXERCISE A.6. Let $\mathbf{x}$ be i.i.d. complex Gaussian with real and imaginary part distributed as $\mathcal{N}(0, \mathbf{K}_x)$. Then the covariance and pseudo-covariance of $\mathbf{x}$ is given by:

$$\begin{aligned} \mathbf{K} &= E[\mathbf{x}\mathbf{x}^*], \\ &= (\mathbf{K}_x(1,1) + \mathbf{K}_x(2,2))\mathbf{I}, \\ \mathbf{J} &= E[\mathbf{x}\mathbf{x}^t], \\ &= (\mathbf{K}_x(1,1) - \mathbf{K}_x(2,2) + 2j\mathbf{K}_x(1,2))\mathbf{I}. \end{aligned}$$

Now, let $\mathbf{y} = \mathbf{U}\mathbf{x}$, then the covariance of $\mathbf{y}$ is given by

$$E[\mathbf{y}\mathbf{y}^*] = \mathbf{U}E[\mathbf{x}\mathbf{x}^*]\mathbf{U}^*$$

$$= (\mathbf{K}_x(1,1) + \mathbf{K}_x(2,2))\mathbf{I}.$$

The pseudo-covariance of $\mathbf{y}$ is given by

$$\begin{aligned} E[\mathbf{y}\mathbf{y}^t] &= \mathbf{U}E[\mathbf{x}\mathbf{x}^t]\mathbf{U}^t \\ &= (\mathbf{K}_x(1,1) - \mathbf{K}_x(2,2) + 2j\mathbf{K}_x(1,2))\mathbf{U}\mathbf{U}^t. \end{aligned}$$

Since for general $\mathbf{U}$, $\mathbf{U}\mathbf{U}^t$ cannot be identity, we get $\mathbf{K}_x(1,1) = \mathbf{K}_x(2,2)$ and $\mathbf{K}_x(1,2) = 0$. That is, $\mathbf{x}$ should be circularly symmetric.

Exercise A.7. The ML decision rule is given through the likelihood ratio which is

$$\begin{aligned} \frac{\mathbb{P}(\mathbf{y}|x=1)}{\mathbb{P}(\mathbf{y}|x=-1)} &= \frac{\mathbb{P}_{\mathbf{z}}(\mathbf{y}-\mathbf{h})}{\mathbb{P}_{\mathbf{z}}(\mathbf{y}+\mathbf{h})}, \\ &= \prod_{i=1}^{n} \frac{\mathbb{P}_{\mathbf{z}_i}(\mathbf{y}_i-\mathbf{h}_i)}{\mathbb{P}_{\mathbf{z}_i}(\mathbf{y}_i-\mathbf{h}_i)}, \end{aligned}$$

where $\mathbf{z}_i, \mathbf{y}_i$ and $\mathbf{h}_i$ are two dimensional vectors with entries as the real and complex parts of the ith entry of $\mathbf{z}, \mathbf{y}$ and $\mathbf{h}$ respectively. Now,

$$\begin{aligned} \frac{\mathbb{P}_{\mathbf{z}_i}(\mathbf{y}_i-\mathbf{h}_i)}{\mathbb{P}_{\mathbf{z}_i}(\mathbf{y}_i-\mathbf{h}_i)} &= \frac{\exp\left(-(\mathbf{y}_i-\mathbf{h}_i)^t\mathbf{K}_x^{-1}(\mathbf{y}_i-\mathbf{h}_i)/2\right)}{\exp\left(-(\mathbf{y}_i+\mathbf{h}_i)^t\mathbf{K}_x^{-1}(\mathbf{y}_i+\mathbf{h}_i)/2\right)} \\ , &= \exp\left(-(\mathbf{h}_i^t\mathbf{K}_x^{-1}\mathbf{y}_i + \mathbf{y}_i^t\mathbf{K}_x^{-1}\mathbf{h}_i)/2\right), \\ &= \exp\left(-\mathbf{h}_i^t\mathbf{K}_x^{-1}\mathbf{y}_i\right). \end{aligned}$$

Thus, the likelihood ratio can be written as

$$\exp\left(-\sum_{i=1}^{n}\mathbf{h}_i^t\mathbf{K}_x^{-1}\mathbf{y}_i\right).$$

Note that $\sum_i \mathbf{h}_i^t\mathbf{y}_i = \mathbf{h}^*\mathbf{y}$. Thus for the likelihood ratio to be a function of only $\mathbf{h}^*\mathbf{y}$, we need that every $\mathbf{h}_i$ should be a right-eigenvector of $\mathbf{K}_x^{-1}$ with the same eigenvalue. Note that this condition is trivially satisfied if $\mathbf{K}_x$ is a scalar multiple of the identity matrix.

Exercise A.8. 1. $\mathbf{z} \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$. Let $H_0 = \{x = 1\}$ and $H_1 = \{x = -1\}$. Then,

$$\begin{aligned} p(\mathbf{y} \mid H_0) &= K\exp(-\|\mathbf{y}-\mathbf{h}\|^2/(2\sigma^2)) \\ p(\mathbf{y} \mid H_1) &= K\exp(-\|\mathbf{y}+\mathbf{h}\|^2/(2\sigma^2)) \\ LLR(\mathbf{y}) &= \log\left[\frac{p(\mathbf{y} \mid H_0)}{p(\mathbf{y} \mid H_1)}\right] = -\frac{1}{2\sigma^2}[\|\mathbf{y}\|^2 + \|\mathbf{h}\|^2 - 2\mathbf{y}^T\mathbf{h} - \|\mathbf{y}\|^2 - \|\mathbf{h}\|^2 - 2\mathbf{y}^T\mathbf{h}] \end{aligned}$$

$$= \frac{2}{\sigma^2}\mathbf{y}^T\mathbf{h} \underset{\hat{H}=1}{\overset{\hat{H}=0}{\gtrless}} 0$$

This last expression is the ML detection rule. Therefore $\mathbf{y}^T\mathbf{h}$ is a sufficient statistic.

$$\begin{aligned} P_e &= P(\mathbf{y}^T\mathbf{h} > 0 \mid x = -1) = P[(-\mathbf{h}+\mathbf{z})^T\mathbf{h} > 0] = P(\mathbf{h}^T\mathbf{z} > \|\mathbf{h}\|^2) \\ &= P\left(\frac{\mathbf{h}^T\mathbf{z}}{\|\mathbf{h}\|} > \|\mathbf{h}\|\right) = Q\left(\frac{\|\mathbf{h}\|}{\sigma}\right) \end{aligned}$$

where we have used the fact that $E\left[\frac{\mathbf{h}^T\mathbf{z}}{\|\mathbf{h}\|}\frac{\mathbf{z}^T\mathbf{h}}{\|\mathbf{h}\|}\right] = \sigma^2$.

We see that P_e is minimized by choosing any $\mathbf{h}$ such that $\|\mathbf{h}\| = 1$. In this case $P_e = Q(1/\sigma)$.

2. $\mathbf{z} \sim \mathbf{N}(\mathbf{0}, \mathbf{K}_z)$, with non-singular $\mathbf{K}_z$. Let $H_0 = \{x = 1\}$ and $H_1 = \{x = -1\}$. Then,

$$\begin{aligned} p(\mathbf{y} \mid H_0) &= K \exp\left[-(\mathbf{y}-\mathbf{h})^T\mathbf{K}_z^{-1}(\mathbf{y}-\mathbf{h})/2\right] \\ p(\mathbf{y} \mid H_1) &= K \exp\left[-(\mathbf{y}+\mathbf{h})^T\mathbf{K}_z^{-1}(\mathbf{y}+\mathbf{h})/2\right] \\ LLR(\mathbf{y}) &= \log\left[\frac{p(\mathbf{y} \mid H_0)}{p(\mathbf{y} \mid H_1)}\right] = \mathbf{y}^T\mathbf{K}_z^{-1}\mathbf{h} + \mathbf{h}^T\mathbf{K}_z^{-1}\mathbf{y} \\ &\qquad \hat{H} = 0 \\ &= 2\mathbf{y}^T\mathbf{K}_z^{-1}\mathbf{h} \underset{\hat{H}=1}{\overset{\hat{H}=0}{\gtrless}} 0 \\ &\qquad \hat{H} = 1 \end{aligned}$$

This last expression is the ML detection rule. Therefore $\mathbf{y}^T\mathbf{K}_z^{-1}\mathbf{h}$ is a sufficient statistic.

$$\begin{aligned} P_e &= P(\mathbf{y}^T\mathbf{K}_z^{-1}\mathbf{h} > 0 \mid x = -1) = P[(-\mathbf{h}+\mathbf{z})^T\mathbf{K}_z^{-1}\mathbf{h} > 0] = P(\mathbf{z}^T\mathbf{K}_z^{-1}\mathbf{h} > \mathbf{h}^T\mathbf{K}_z^{-1}\mathbf{h}) \\ &= P\left(\frac{\mathbf{z}^T\mathbf{K}_z^{-1}\mathbf{h}}{\sqrt{\mathbf{h}^T\mathbf{K}_z^{-1}\mathbf{h}}} > \sqrt{\mathbf{h}^T\mathbf{K}_z^{-1}\mathbf{h}}\right) = Q\left(\sqrt{\mathbf{h}^T\mathbf{K}_z^{-1}\mathbf{h}}\right) \end{aligned}$$

We see that P_e is minimized by choosing $\mathbf{h}$ to be the norm one eigenvector of $\mathbf{K}_z^{-1}$ associated with its maximum eigenvalue, which is the norm one eigenvector of $\mathbf{K}_z$ associated with its minimum eigenvalue λ_{min}. In this case $P_e = Q\left(\sqrt{\lambda_{min}^{-1}}\right)$.

This choice of $\mathbf{h}$ makes the signal of interest $\mathbf{h}x$ lie in the direction where the noise is smallest, maximizing the signal to noise ratio in the received signal $\mathbf{y}$.

3. If $\mathbf{K}_z$ is singular then the noise vector $\mathbf{z}$ lies in a proper subspace of $\mathcal{R}^n$ which we call $\mathcal{S}$. Then letting $\mathcal{S}^\perp$ be the orthogonal subspace associated to $\mathcal{S}$, we can choose $\mathbf{h} \in \mathcal{S}^\perp$, project onto this direction, and obtain a noise-free sufficient statistic of x. The corresponding probability of error is 0.

Appendix B

Solutions to Exercises

EXERCISE B.1.

EXERCISE B.2. 1. $f(x)$ is concave if for any $\lambda \in [0, 1]$:

$$\lambda f(x_1) + (1 - \lambda)f(x_2) \leq f[\lambda x_1 + (1 - \lambda)x_2] \tag{B.1}$$

for all $x_1, x_2 \in Dom(f)$.

2. Jensen's inequality: for a random variable X and a concave function $f(\cdot)$:

$$E[f(X)] \leq f[E(X)] \tag{B.2}$$

Proof by picure:

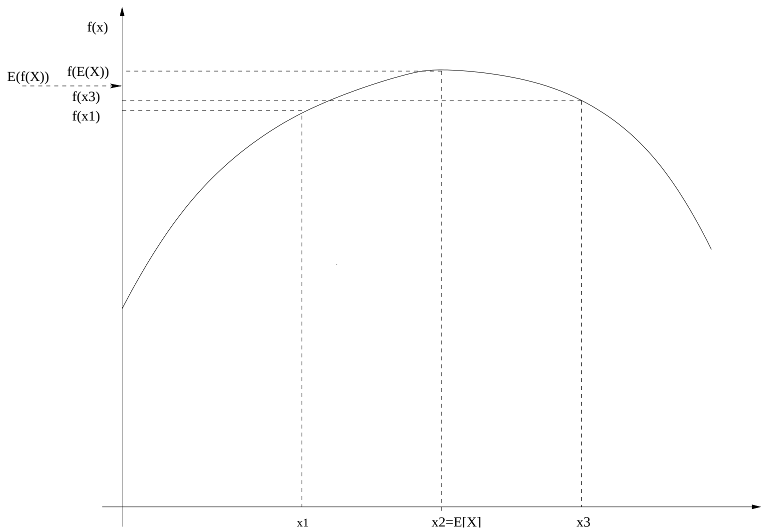


Figure B.1: Example of Jensen's inequality for a discrete random variable that takes only 3 values with equal probability and a concave function $f(\cdot)$.

3.

$$\begin{aligned} H(X) - H(X|Y) &= I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \\ &= -\sum_{x,y} p(x,y) \log \frac{p(x)p(y)}{p(x,y)} && \text{(B.3)} \\ &\geq -\log \left[\sum_{x,y} p(x,y) \frac{p(x)p(y)}{p(x,y)} \right] && \text{(B.4)} \\ &= -\log 1 = 0 \end{aligned}$$

where the inequality follows from Jensen's inequality and the convexity of $-\log(\cdot)$. We have equality iff $p(x)p(y) = p(x,y)$ for all x, y, i.e. X and Y are independent.

For the required example consider $X|y = 0 \sim \textit{Bernoulli}(1/2)$, $X|y = 1 \sim \textit{Bernoulli}(0)$ and $Y \sim \textit{Bernoulli}(1/2)$. It is easy to check that $X \sim \textit{Bernoulli}(1/4)$ and we can compute the different entropies. $H(X) = H(1/4) = 0.811$, $H(X|y = 0) = H(1/2) = 1$, $H(X|y = 1) = 0$, and $H(X|Y) = (1/2)H(X|y = 0) + (1/2)H(X|y = 1) = 1/2$, where we used $H(p)$ to denote the entropy of a $\textit{Bernoulli}(p)$ random variable. We see that $H(X|y = 0) > H(X)$ but $H(X|Y) < H(X)$ in agreement with the inequality that we just proved.

Exercise B.3.

Exercise B.4.

Exercise B.5.

Exercise B.6.

Exercise B.7.

Exercise B.8.

Exercise B.9.

Exercise B.10.

Exercise B.11.

Exercise B.12.