Short Course on MIMO Systems

Diversity in Communications

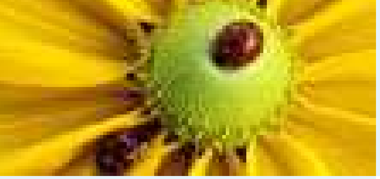
Raviraj S. Adve

University of Toronto

Dept. of Elec. and Comp. Eng.

Toronto, Ontario, Canada M5S 3G4

rsadve@comm.utoronto.ca



Multiple Input Multiple Output Systems

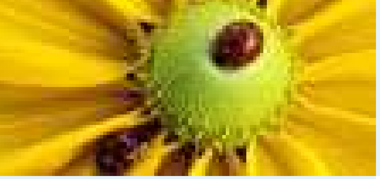
MIMO Systems: the use of an antenna array at the receiver (Multiple Output) and/or the transmitter (Multiple Input) in wireless communications

Outline of this Course:

- *Basic digital and wireless communications*
- *Diversity on Receive*
- *Diversity on Transmit*
- *Multiplexing and data rate*

Detailed notes available at

http://www.comm.utoronto.ca/˜ rsadve/teaching.html

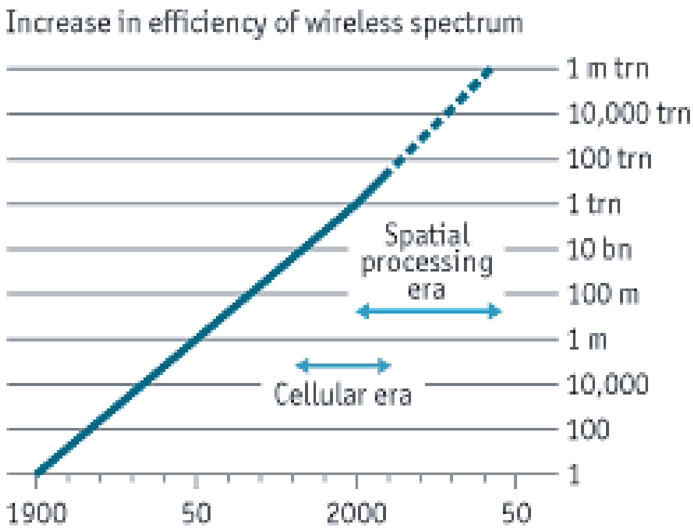


I like this research area because...

MIMO Systems bring together....

- Antenna array theory
- Probability theory
- Linear algebra
- Optimization
- Digital communications

....and it is useful!



(borrowed from *The Economist, April 28th - May 4th 2007*)

A Digital Communication System

Introduction and Overview

Basic Digital Communications
- BER
- Info Theory

Basic Wireless Communications

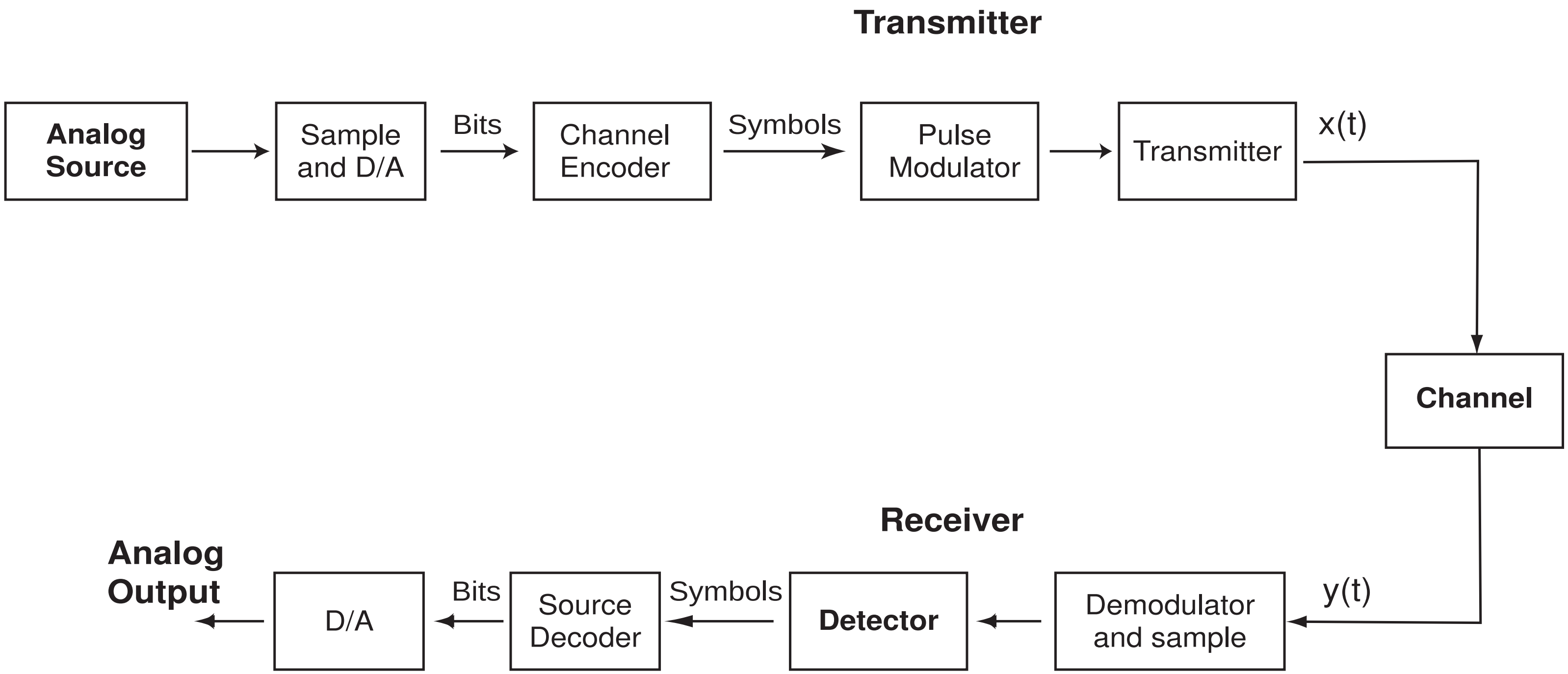
Receive Diversity

Transmit Diversity

MIMO Information Theory

Course Summary

END



$$x(t) = s(t) + n(t)$$

- The noise term, $n(t)$, is usually modelled as additive, white, Gaussian noise (AWGN)

Performance Measure : BER

- Performance of the system is generally measured via the bit error rate (BER)
  - BER is a function of signal-to-noise ratio (SNR)

$$\text{SNR} = \frac{E\{|s(t)|^2\}}{E\{|n(t)|^2\}}$$

where $E\{\cdot\}$ is the expectation operator.

- Binary Phase Shift Keying (BPSK): bit = 0 $\longrightarrow$ symbol = 1, bit = 1 $\longrightarrow$ symbol = -1
- For an AWGN channel and BPSK

$$\begin{aligned} \text{BER} &= Q\left(\sqrt{2\text{SNR}}\right) \\ Q(x) &= \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt \end{aligned}$$

  - $Q(x) < (1/2)e^{-x^2/2}$, i.e., *for an AWGN channel BER falls off exponentially with SNR.*

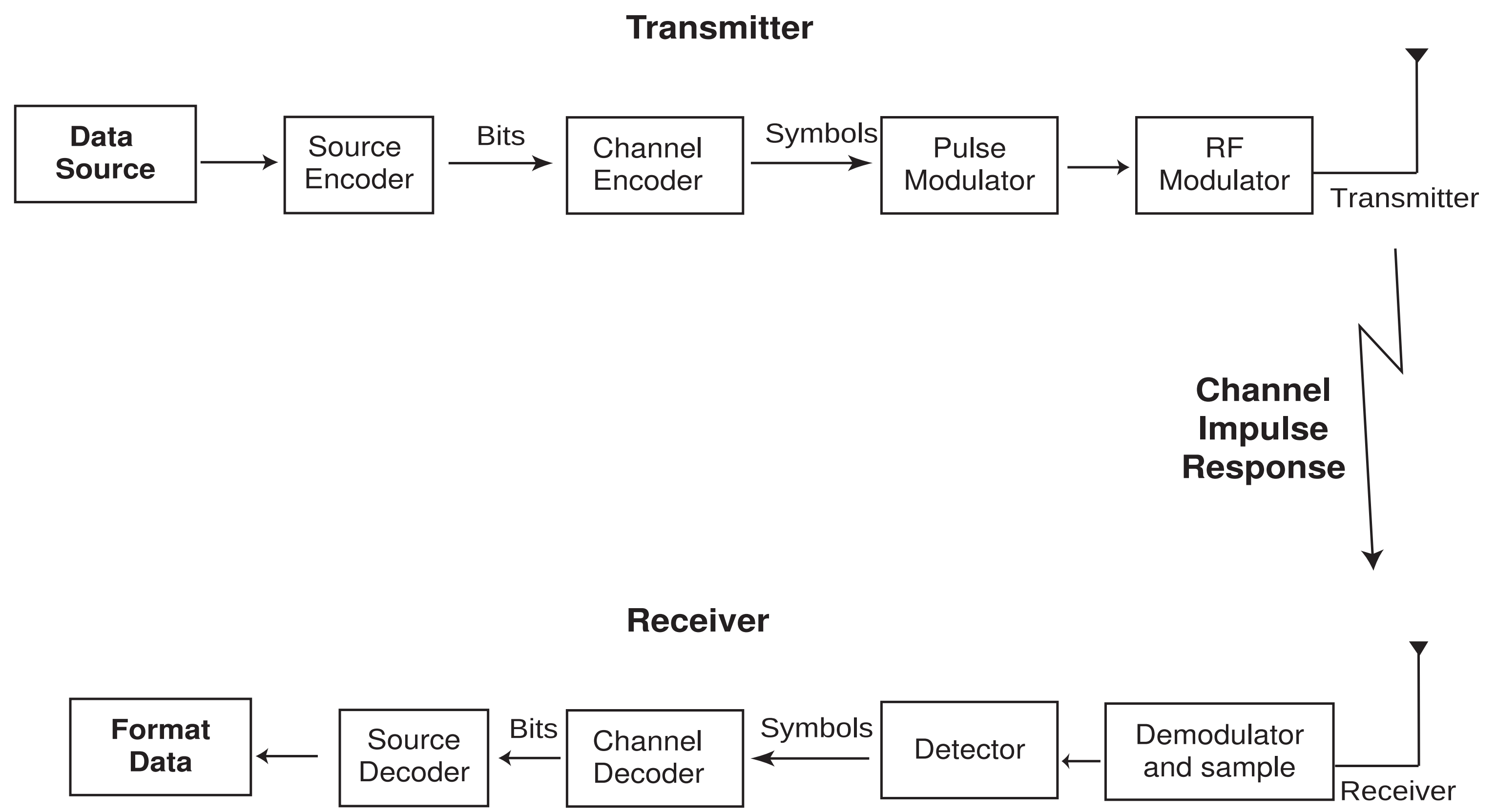
Basic Information Theory

- Channels are characterized by channel capacity C
- Shannon says: Given a channel with capacity C, one can find a coding scheme to transmit at a data rate $R < C$ without error. Furthermore, one cannot transmit without error at a data rate $R > C$.
  - C acts as the effective speed limit on the channel
- R is generally measured in bits per channel use.
- For an AWGN channel (with complex inputs and outputs)

$$C = \log_2 (1 + \text{SNR}) \qquad\qquad (\text{bits})$$

- Note that C is a non-linear function of SNR
  - At low SNR, $C \simeq$ SNR
  - At high SNR $C \simeq \log_2(\text{SNR})$

A Wireless Communication System



A wireless communication system is fundamentally limited by the random channel, i.e., fading.

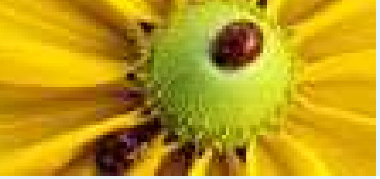
Fading Channels

Due to the unknown location of the mobile station and the unknown medium between the transmitter and receiver, the wireless channel is best characterized as *random*.

- Fading has three components:

  Overall fading = (Distance Attenuation) $\times$ (Large Scale Fading) $\times$ (Small Scale Fading)

- Distance attenuation: fall off in power with distance
  - In *line-of-sight* conditions, received power $\propto 1/d^2$ (where d is the distance between transmitter and receiver)
  - In *non line-of-sight* conditions, received power $\propto 1/d^n$, where n is the distance attenuation parameter
    - $1.5 < n < 4.5$.
    - In some scenarios $n < 2$ due to tunnelling
    - n large in urban areas, $\sim$ 4.5



Fading Channels (cont...)

Introduction and Overview

Basic Digital Communications

Basic Wireless Communications
- Fading Channels
- Small Scale Fading
- Rayleigh/Rician Fading
- Flat/Selective Fading
- Slow/Flat fading
- Summary
- SISO Data Model
- Channel Fluctuation
- Error Rate
- Outage Probability

Receive Diversity

Transmit Diversity

MIMO Information Theory

Course Summary

END

- Large Scale Fading: Occurs due to the attenuation each time a signal passes through an object

$$\text{Large Scale Fading} = 10^{-(x_1+x_2+\ldots)} = 10^{-x}$$

where x_i is the attenuation due to object # i.

- By the Cental Limit Theorem, $x = (x_1 + x_2 + \ldots)$ is a Gaussian random variable

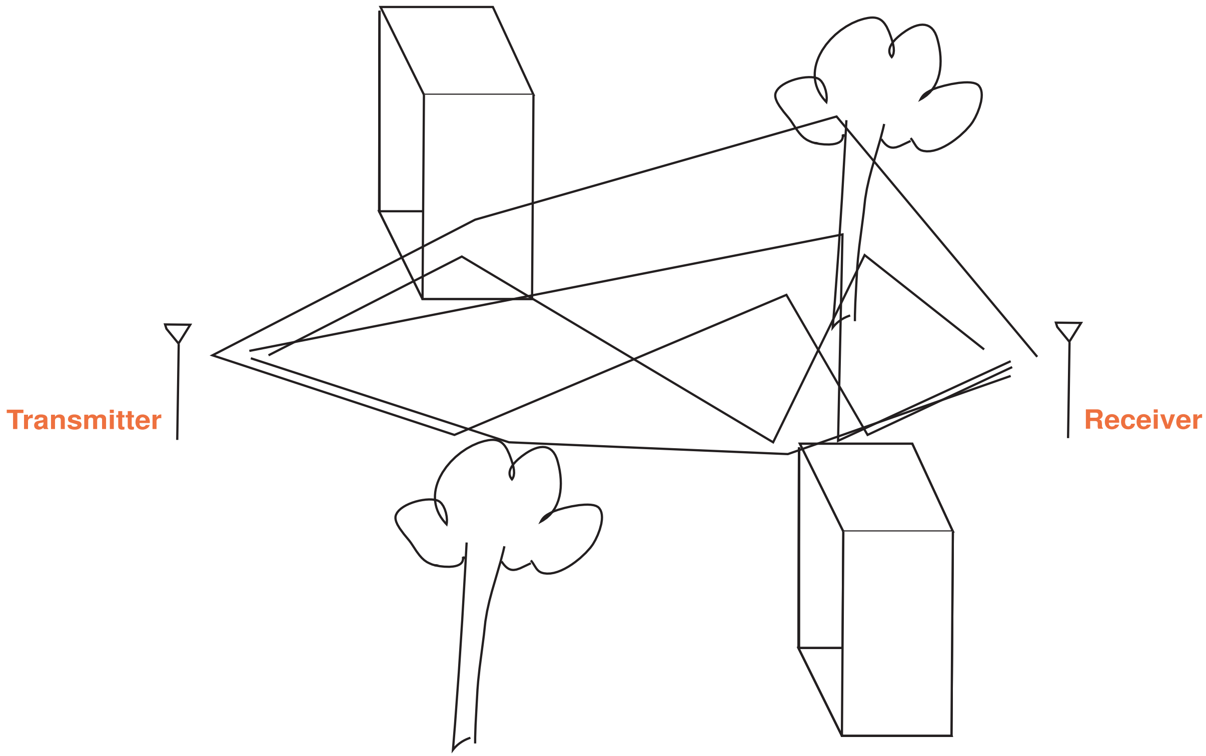
  $\Rightarrow$ large scale fading can be modelled as log-normal, i.e., the log of the fading term is distributed normal:

$$\text{Large Scale Fading} = 10^{-x};$$
$$x \sim \mathcal{N}(0, \sigma_{h1}^2)$$

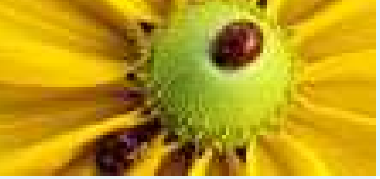
- Large scale fading varies slowly, remaining approximately constant over hundreds of wavelengths
  - There is not much one can do about large scale fading or distance attenuation, just *power control*

Small Scale Fading

- Occurs due to *multipath propagation*
  - Each signal arrives over many many paths



- Each path has slightly different length $\Rightarrow$ slightly different time of propagation
  $\Rightarrow$ with different phase that is effectively random



Small Scale Fading (cont...)

- Total received signal is the sum over all paths

$$\begin{aligned}\text{Received Signal} &= \text{Transmitted Signal} \times \text{Distance Attenuation}\\ &\quad \times \text{Large Scale Fading} \times (h_1 + h_2 \ldots)\\ h_i &= \alpha_i e^{j\phi_i}; \qquad \phi_i \text{ is random}\end{aligned}$$

- If all α_i are approximately the same (*Rayleigh* fading),

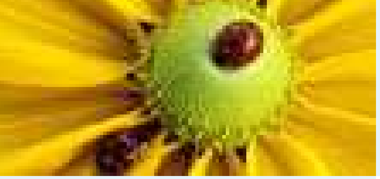
$$\text{Small Scale Fading} = h = (h_1 + h_2 + \ldots)$$

$$h \sim \mathcal{CN}(0, \sigma_h^2)$$

- If one of the components is dominant (*Rician* fading)

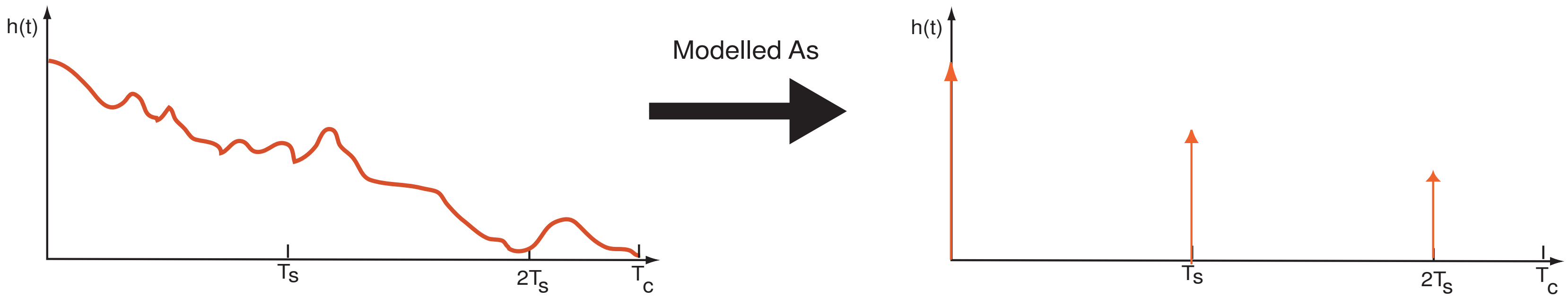
$$h \sim \mathcal{CN}(\mu, \sigma_h^2)$$

- $\mathcal{CN}(\mu, \sigma_h^2)$ represents the complex Gaussian distribution with mean μ and variance σ_h^2.
  - We will focus on Rayleigh fading, i.e., $h \sim \mathcal{CN}(0, \sigma_h^2)$
  - Note: Rayleigh fading is just one of many fading models



Small Scale Fading (cont...)

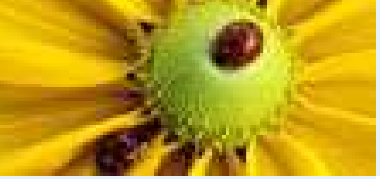
- Frequency flat versus frequency selective fading:
  - Symbol period: T_s; Channel time spread: T_c



- Channel modelled as a train of impulses
  - If $T_c < T_s$, $h(t) = \alpha\delta(t) \Rightarrow H(j\omega) =$ constant (Frequency flat fading)
  - If $T_c > T_s$, $h(t) = \sum_\ell \alpha_\ell \delta(t - \ell T_s) \Rightarrow H(j\omega) \neq$ constant (Frequency selective fading)
- Note: $\alpha \sim \mathcal{CN}(0, \sigma_h^2)$

Small Scale Fading (cont...)

- If the mobile is moving with radial velocity v, the channel changes as a function of time
  - Rate of change = Doppler frequency ($f_d = v/\lambda$)
- Symbol rate = $f_s = 1/T_s$
  - If $f_d \ll f_s$, channel is effectively constant over several symbols (slow fading)
  - If $f_d > f_s$, channel changes within a symbol period (fast fading)
- We shall focus on slow, flat, fading



In summary...

- Fading has three components
  - Distance attenuation $\propto 1/d^n$
  - Large scale fading; modelled as log-normal; constant over hundreds of λ
  - Small scale fading; modelled as Rayleigh, i.e., complex normal; fluctuates within fraction of λ
- Assume power control for distance attenuation and large scale fading.
  - This is all that can be done!

Data Model

- With flat, slow fading:

$$\boxed{x = hs + n}$$

- x: received signal, h: channel, s: transmitted complex symbol, n: noise
  - $h \sim \mathcal{CN}(0, \sigma_h^2)$
  - Instantaneous channel power: $|h|^2 \sim (1/\sigma_h^2)e^{-|h|^2/\sigma_h^2}$ (exponential)
    - channel is often in “bad shape”
  - Note: $\sigma_h^2 = E\{|h|^2\}$ = average power in channel
    - Set $\sigma_h^2 = 1$ for convenience (channel does not “introduce” power)

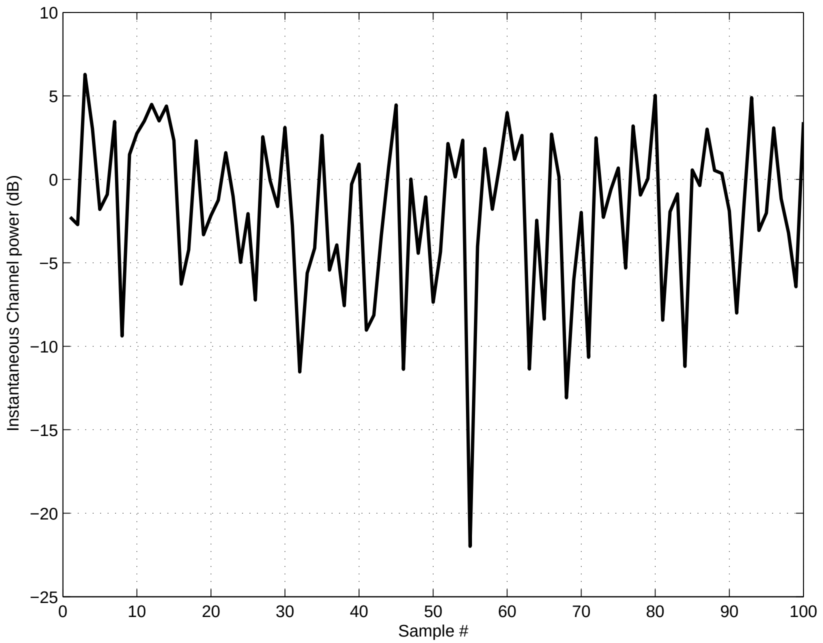
Introduction and Overview
Basic Digital Communications
Basic Wireless Communications
- Fading Channels
- Small Scale Fading
- Rayleigh/Rician Fading
- Flat/Selective Fading
- Slow/Flat fading
- Summary
- SISO Data Model
- Channel Fluctuation
- Error Rate
- Outage Probability
Receive Diversity
Transmit Diversity
MIMO Information Theory
Course Summary
END

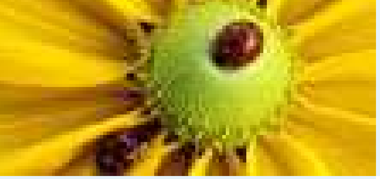
Impact of Fading

$$\text{Average SNR} = E\{|h|^2\}\frac{E\{|s|^2\}}{\sigma^2} = \sigma_h^2\Gamma = \Gamma$$

$$\gamma = \text{Instantaneous SNR} = |h|^2\frac{E\{|s|^2\}}{\sigma^2} = |h|^2\Gamma; \qquad \gamma \sim \frac{1}{\Gamma}e^{-\gamma/\Gamma}$$

- Note that the average SNR has not changed
- The fluctuation in power due to the fading seriously impacts on the performance of a wireless system





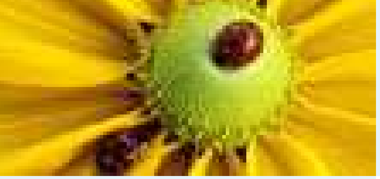
Bit Error Rate

- Without fading, BER = $Q(\sqrt{2\Gamma}) \simeq \exp(-\Gamma)$
  - Exponential drop off with SNR
- With fading, instantaneous SNR = γ, i.e., instantaneous BER = $Q(\sqrt{2\gamma})$

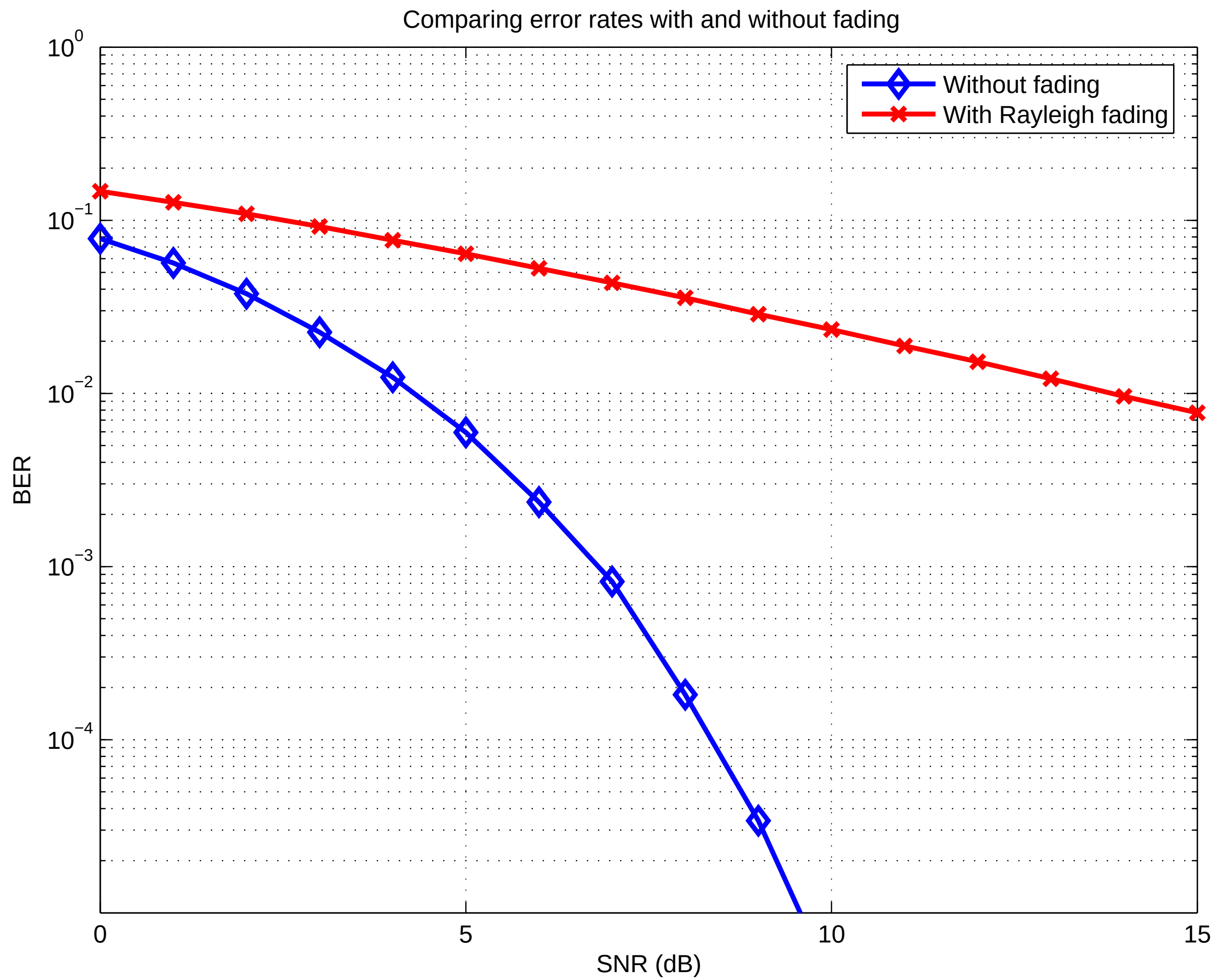
$$
\begin{aligned}
\Rightarrow \text{Average BER} &= E_\gamma\{Q(\sqrt{2\gamma}\} \\
&= \int_0^\infty Q(\sqrt{2\gamma})\,\frac{1}{\Gamma}e^{-\gamma/\Gamma}\,d\gamma \\
&= \frac{1}{2}\left(1-\sqrt{\frac{\Gamma}{1+\Gamma}}\right)
\end{aligned}
$$

  - At high SNR ($\Gamma \rightarrow \infty$),

$$\boxed{\text{BER} \propto \frac{1}{\Gamma}}$$



Bit Error Rate: Example



Note: In the case with fading, the BER v/s SNR plot (in log-log format appears as a straight line)

Outage Probability

Introduction and Overview
Basic Digital Communications
Basic Wireless Communications
- Fading Channels
- Small Scale Fading
- Rayleigh/Rician Fading
- Flat/Selective Fading
- Slow/Flat fading
- Summary
- SISO Data Model
- Channel Fluctuation
- Error Rate
- Outage Probability

Receive Diversity
Transmit Diversity
MIMO Information Theory
Course Summary
END

- If channel is known, capacity: $C = \log_2\left(1 + |h|^2\Gamma\right)$
- If channel is unknown, true capacity = 0!!
  - One cannot guarantee any data rate
  - Therefore, define outage probability for a rate R

$$\begin{aligned} P_{\text{out}} &= P(C < R) \\ P\left(\log_2\left(1+|h|^2\Gamma\right) < R\right) &= P\left(|h|^2 < \frac{2^R - 1}{\Gamma}\right) \\ &= 1 - \exp\left(-\frac{2^R-1}{\Gamma}\right) \end{aligned}$$

  - IMPORTANT: as average SNR gets large $(\Gamma \rightarrow \infty)$,

$$\exp\left(-\frac{2^R-1}{\Gamma}\right) \simeq 1 - \frac{2^R-1}{\Gamma}$$

$$\Rightarrow P_{\text{out}} \propto \frac{1}{\Gamma}$$

Outage Probability: Alternate Definition

- Note: Choosing a target rate R is equivalent to choosing a SNR threshold, γ_s
- Alternate definition of outage

$$\begin{aligned} P_{\text{out}} &= P\left[\gamma < \gamma_s\right] \\ &= \int_0^{\gamma_s} \frac{1}{\Gamma} e^{-\gamma/\Gamma} d\gamma \\ &= 1 - e^{-\gamma_s/\Gamma} \end{aligned}$$

- Again, as $(\Gamma \rightarrow \infty)$

$$\boxed{P_{\text{out}} \propto \frac{1}{\Gamma}}$$

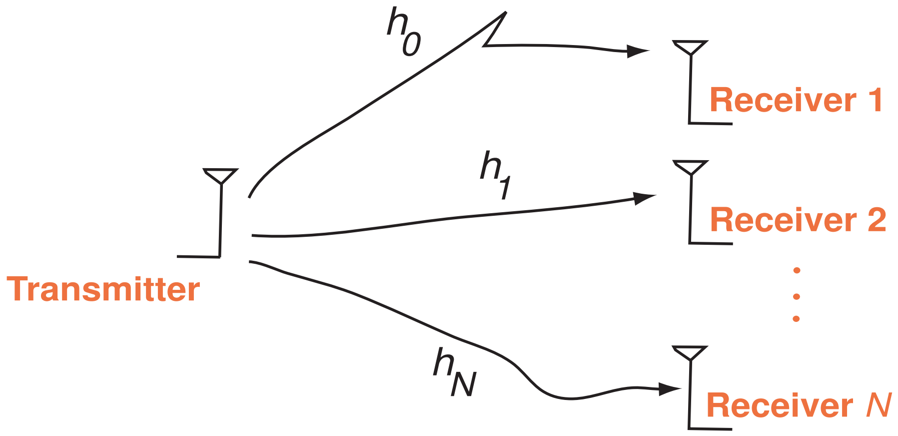
So, what are we going to do about this?

Use multiple antennas!

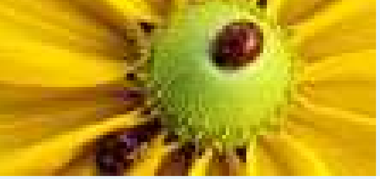
which provide *diversity*

Introduction to Receive Diversity

SIMO : Single antenna at the transmitter, multiple at the receiver



- If for a single receiver, $P_{\text{out}} = 0.1$, for two receivers $P_{\text{out}} = 0.01$
  - $\Rightarrow$ exponential gains in error rate with linear increase in number of antennas
  - fundamental assumption: the error events are *independent*, i.e, *the channels are independent*
- *Key:* provide the receiver with multiple *independent* copies of the message



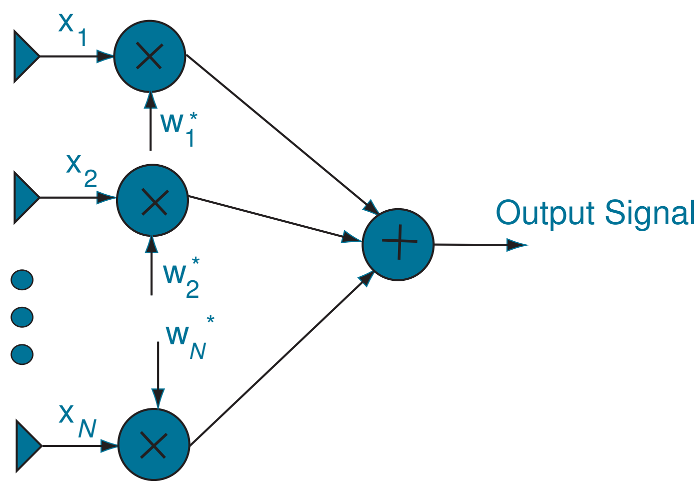
Diversity Basics

- Assume for now that the channels are independent and identically distributed (i.i.d.)
  - We will deal with the issue of correlation later
  - Channel to n^{th} receive element = h_n, i.e., h_n is assumed independent of h_m for $n \neq m$
- Signal to noise ratios are also i.i.d.: γ_n is independent of γ_m, $n \neq m$. Also,

$$\gamma_n \sim \frac{1}{\Gamma} e^{-\gamma_n/\Gamma}$$

  - Note: every channel has the same *average* SNR

Basics (cont...)



- This appears to be beamforming! (for a single user!).
- Write the received signal as a vector

$$\begin{aligned} \mathbf{x} &= \mathbf{h}s + \mathbf{n} \\ \mathbf{h} &= [h_1, h_2, \ldots h_N]^T \end{aligned}$$

  - The output signal is given by:

$$y = \sum_{n=1}^{N} w_n^* x_n = \mathbf{w}^H \mathbf{x} = \mathbf{w}^H \mathbf{h}s + \mathbf{w}^H \mathbf{n}$$

Receive Diversity Techniques

- The weight vector is $\mathbf{w} = [w_1, w_2, \ldots w_N]^T$
- Key: how are these weights chosen?

Receive Diversity Techniques:

- Selection Combining
- Maximal Ratio Combining (MRC)
- Equal Gain Combining (EGC)

Selection Combining

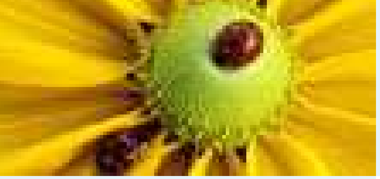
- Choose, for further processing, the receive element with the highest SNR

$$w_k = \begin{cases} 1 & \gamma_k = \max_n\{\gamma_n\} \\ 0 & \text{otherwise} \end{cases}$$

- The output SNR is therefore the maximum of the receive elements

$$\text{Output SNR} = \gamma_{\text{out}} = \max_n \{\gamma_n\}$$

- Note: channel phase information *not required* in the selection process
- This is the simplest diversity scheme
  - Seems to “waste” $(N-1)$ receivers



Analyzing Selection Diversity

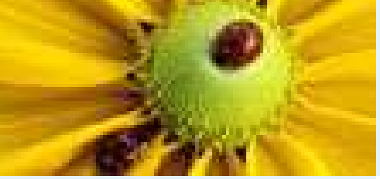
- Outage probability : the output SNR is below threshold γ_s if *all* receive elements have SNR below γ_s:

$$\begin{aligned} P_{\text{out}} &= P\left[\gamma_{\text{out}} < \gamma_s\right] \\ &= P\left[\gamma_1, \gamma_2, \ldots \gamma_N < \gamma_s\right] \\ &= \prod_{n=1}^{N} P[\gamma_n < \gamma_s] \\ \Rightarrow P_{\text{out}} &= \left[1 - e^{-\gamma_s/\Gamma}\right]^N \end{aligned}$$

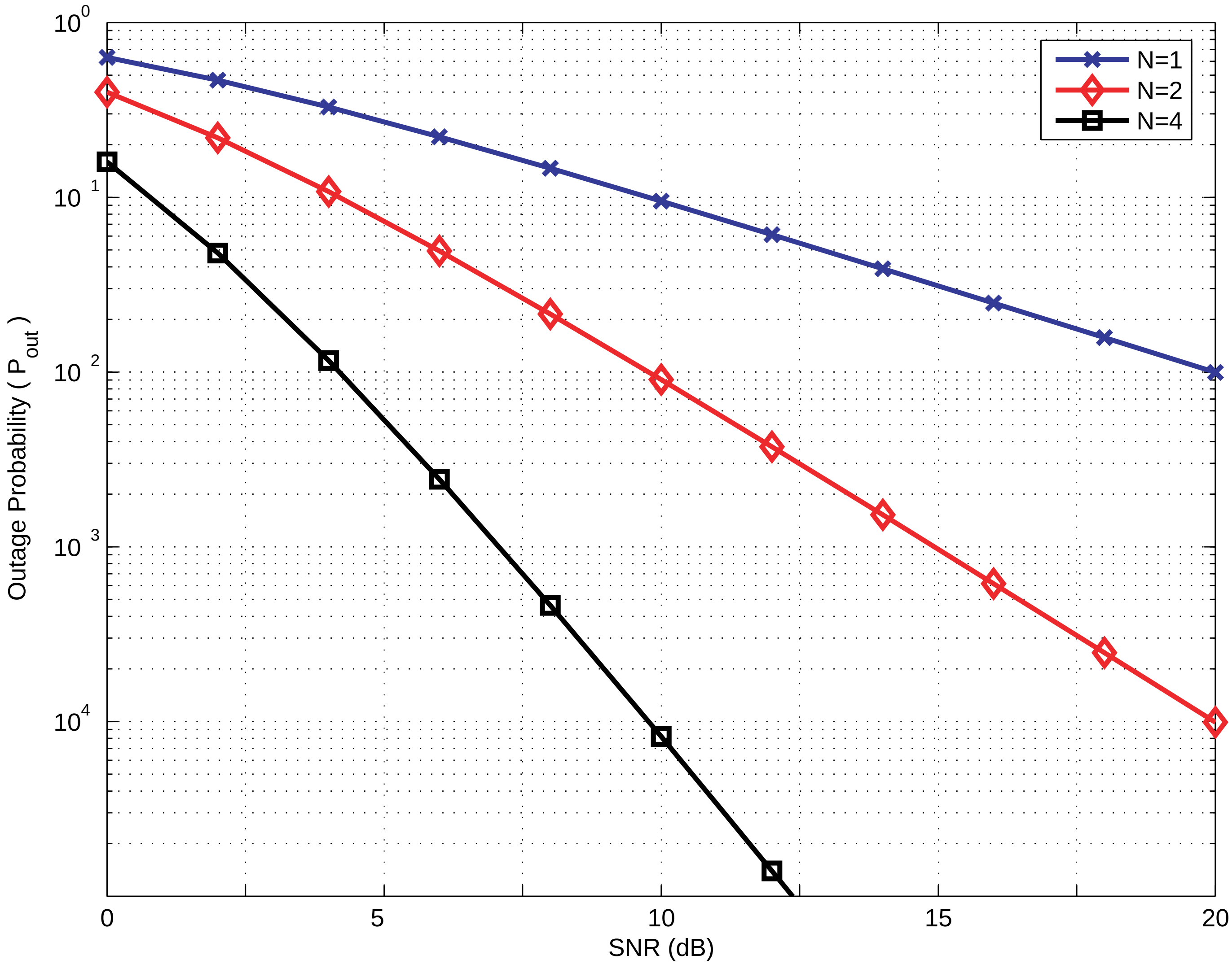
  - $P_{\text{out}} = \left[1 - e^{-\gamma_s/\Gamma}\right]$ for $N = 1$, i.e.,exponential gains in outage probability

- Note: at high SNR, as $\Gamma \to \infty$

$$P_{\text{out}} \propto \left(\frac{1}{\Gamma}\right)^N$$

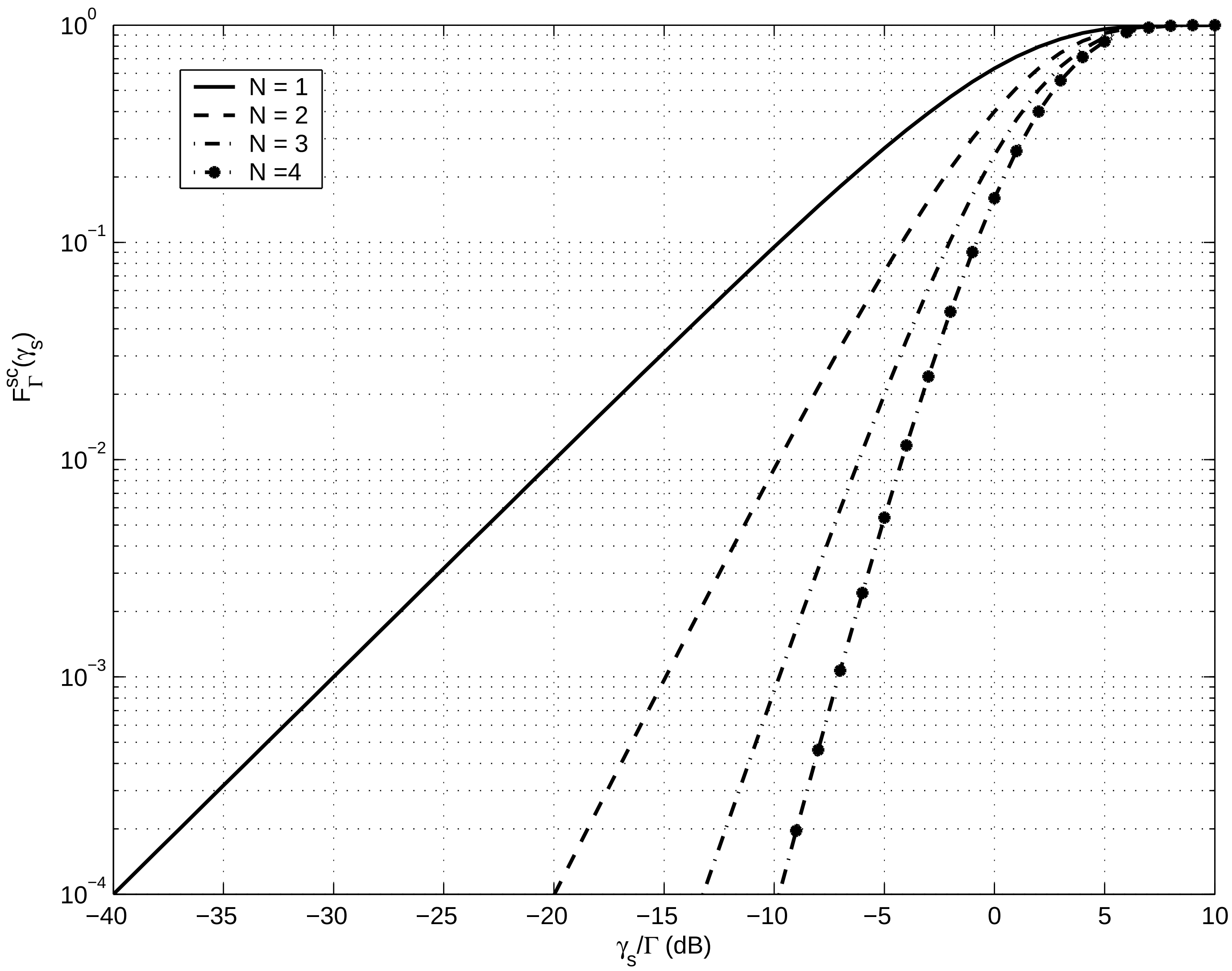


Performance: Outage Probability versus SNR

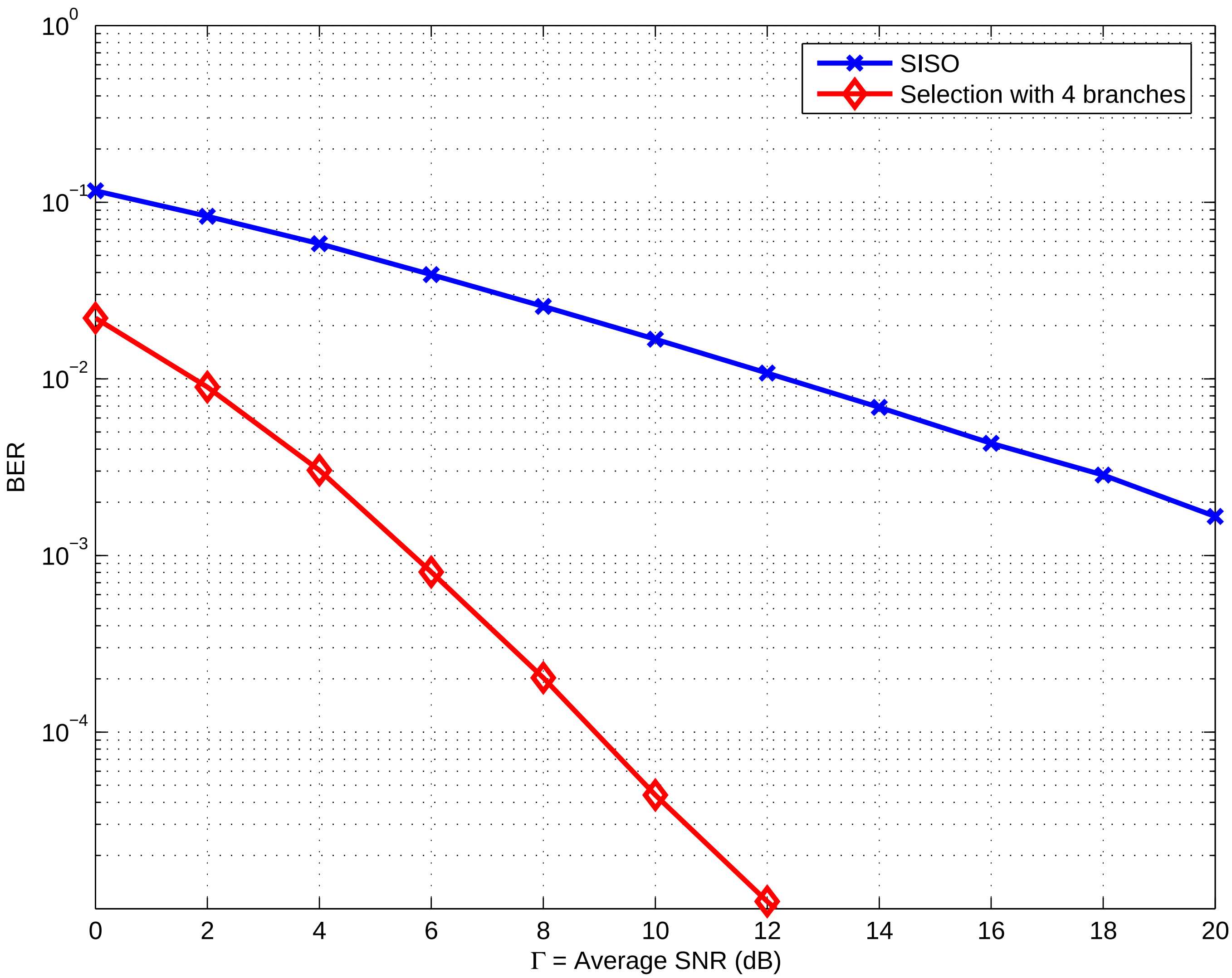


In this figure, $\gamma_s = 0$dB

Performance: Outage Probability versus γ_s/Γ



Performance: Bit Error Rate



Analysis of Selection Combining (cont...)

Q: Are the gains in selection due to gains in SNR?

A: In fact, no!!

SNR analysis:

- Note, $P_{\text{out}} = P\left(\gamma_{\text{out}} < \gamma_s\right)$ is also the cumulative density function (CDF) of output SNR
  - $\Rightarrow$ the probability density function, $f(\gamma_{\text{out}}) = dP_{\text{out}}/d\gamma_{\text{out}}$

$$f(\gamma_{\text{out}}) = \frac{N}{\Gamma} e^{-\gamma_{\text{out}}/\Gamma} \left[1 - e^{-\gamma_{\text{out}}/\Gamma}\right]^{N-1}$$

Also, $$E\{\gamma_{\text{out}}\} = \int_0^{\infty} \gamma_{\text{out}} f(\gamma_{\text{out}}) d\gamma_{\text{out}}$$

SNR Analysis (cont...)

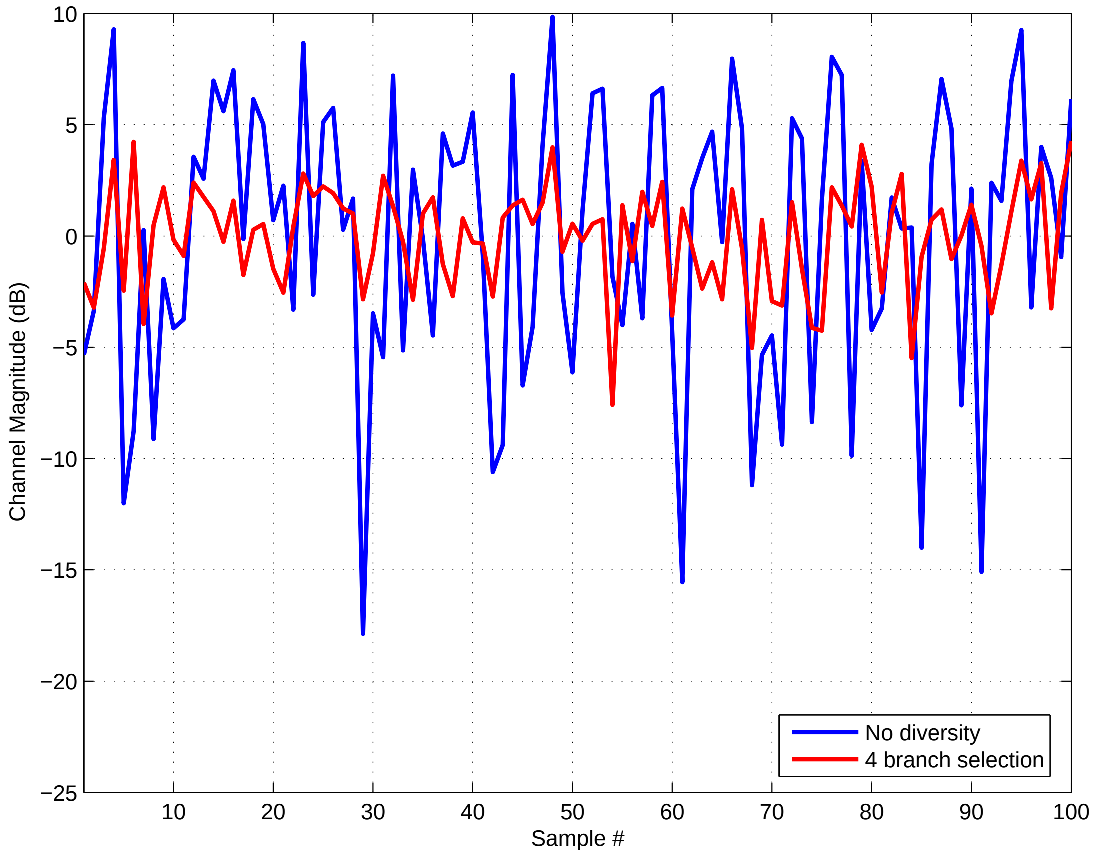
$$\begin{aligned} \Rightarrow E\left\{\gamma_{\text{out}}\right\} &= \Gamma \sum_{n=1}^{N} \frac{1}{n}, \\ &\simeq \Gamma\left(C + \ln N + \frac{1}{2N}\right), \end{aligned}$$

- The gain in SNR is *only ln(N)*!!!

SNR Analysis (cont...)

Q: So, where are the gains coming from?

A: Reduced variation in the channel



Maximal Ratio Combining

- Selection is simple, but wastes $(N-1)$ receive elements
- Maximal Ratio Combining maximizes output SNR γ_{out}

$$\mathbf{x} = \mathbf{h}s + \mathbf{n}$$

$$\mathbf{h} = [h_1,\, h_2, \ldots h_N]^T$$

Output Signal = $$y = \sum_{n=1}^{N} w_n^* x_n = \mathbf{w}^H\mathbf{h}s + \mathbf{w}^H\mathbf{n}$$

Output SNR = $$\gamma_{\text{out}} = \frac{\left|\mathbf{w}^H\mathbf{h}\right|^2 E\{|s|^2\}}{E\left\{\left|\mathbf{w}^H\mathbf{n}\right|^2\right\}} = \frac{\left|\mathbf{w}^H\mathbf{h}\right|^2 E\{|s|^2\}}{\sigma^2||\mathbf{w}||^2}$$

MRC: $$\mathbf{w}_{\text{MRC}} = \max_{\mathbf{w}}\left[\gamma_{\text{out}}\right] = \max_{\mathbf{w}}\left[\frac{\left|\mathbf{w}^H\mathbf{h}\right|^2}{\sigma^2||\mathbf{w}||^2}\right]$$

Introduction and Overview
Basic Digital Communications
Basic Wireless Communications
Receive Diversity
- Introduction
- Diversity Types
- Selection Combining
- Maximal Ratio Combining
- Diversity Order
- Equal Gain Combining
- Comparing Schemes
- Summary
- Correlation

Transmit Diversity
MIMO Information Theory
Course Summary
END

Maximal Ratio Combing (cont...)

- Using Cauchy-Schwarz inequality

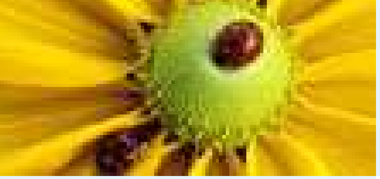
$$\mathbf{w} \propto \mathbf{h}$$

- Choose $\mathbf{w} = \mathbf{h}$,

$$\text{Output Signal} = y = \left(|h_1|^2 + |h_2|^2 + \ldots |h_N|^2\right) s + \text{noise}$$
$$= \left(\sum_{n=1}^{N} |h_n|^2\right) s + \text{noise}$$
$$\text{Output SNR} = \gamma_{\text{out}} = \sum_{n=1}^{N} \frac{E\{|s|^2\}|h_n|^2}{\sigma^2} = \sum_{n=1}^{N} \gamma_n$$

- i.e., the output SNR is the sum of the SNR over all receivers
- PDF of output SNR:

$$f(\gamma_{\text{out}}) = f(\gamma_1) \star f(\gamma_2) \star \cdots f(\gamma_N) = \frac{1}{(N-1)!} \frac{\gamma_{\text{out}}^{N-1}}{\Gamma^N} e^{-\gamma_{\text{out}}/\Gamma},$$



MRC Results

- Average SNR

$$\gamma_{\text{out}} = \sum_{n=1}^{N} \gamma_n \Rightarrow E\{\gamma_{\text{out}}\} = N\Gamma$$

- Outage probability

$$P_{\text{out}} \quad = \quad P\left[\gamma_{\text{out}} < \gamma_s\right] = 1 - e^{-\gamma_s/\Gamma} \sum_{n=0}^{N-1} \left(\frac{\gamma_s}{\Gamma}\right)^n \frac{1}{n!}$$

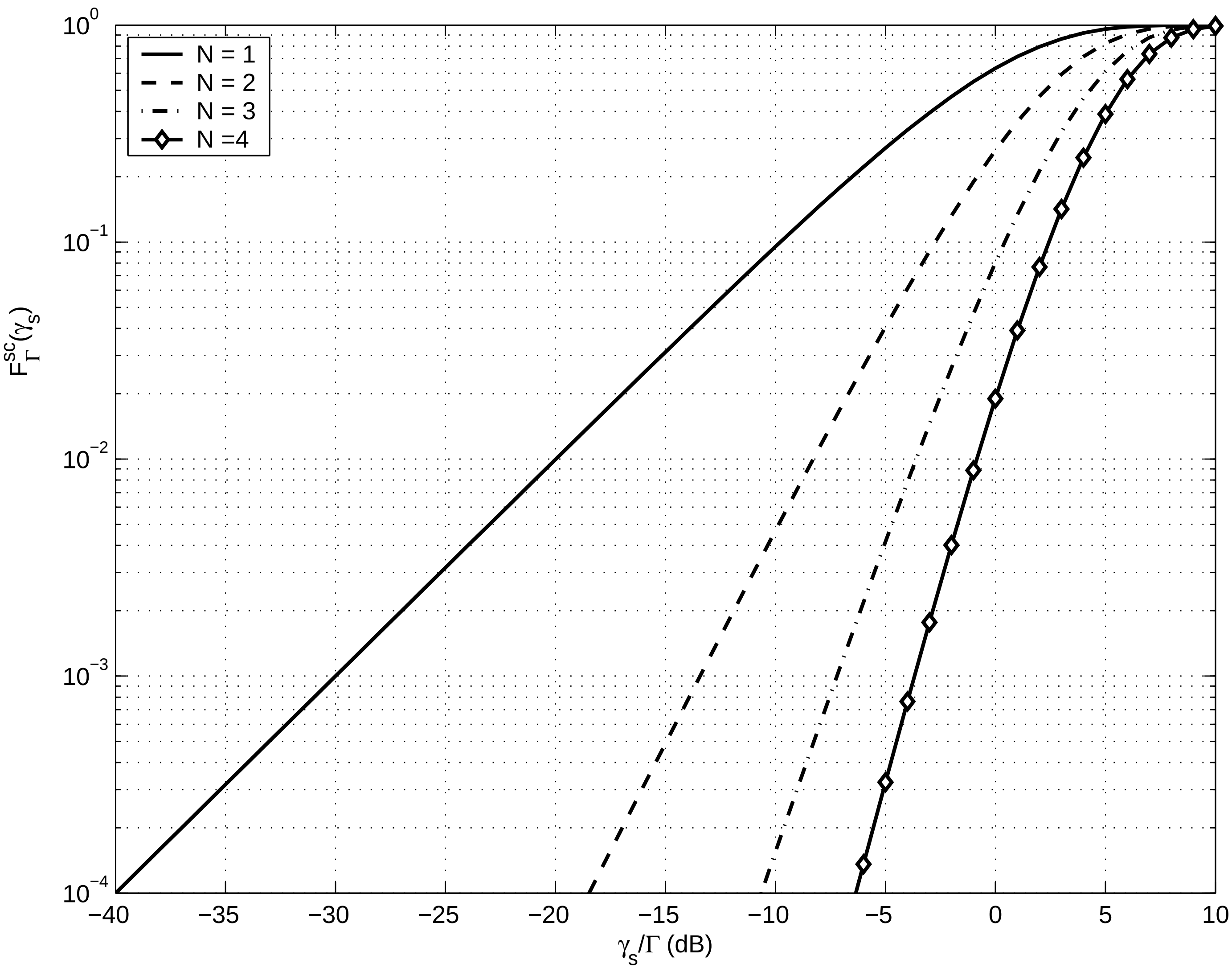
- At high SNR $(\Gamma \rightarrow \infty)$

$$P_{\text{out}} \propto \left(\frac{1}{\Gamma}\right)^N$$

- Similarly, bit error rate:

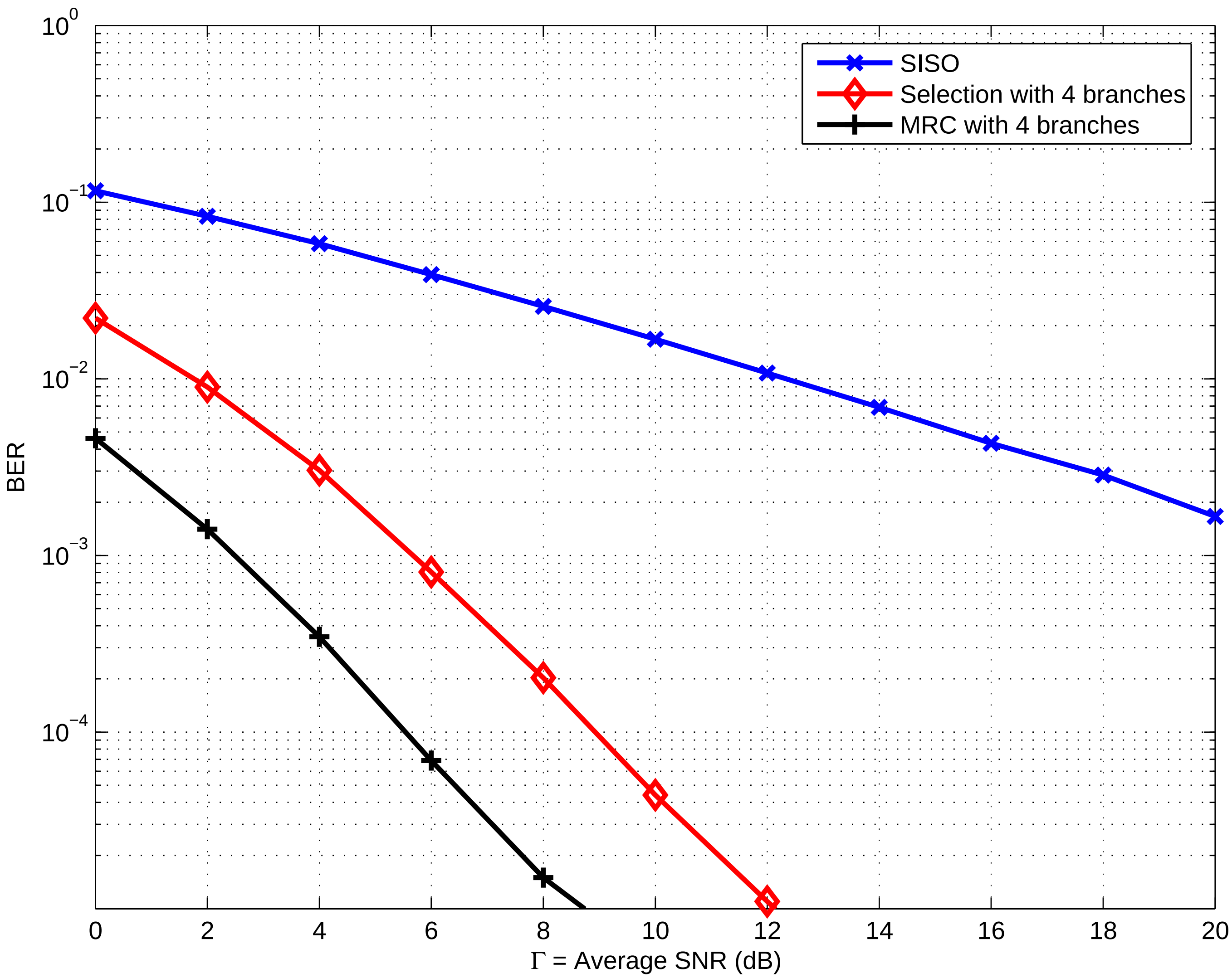
$$\text{BER} \quad = \quad \int_0^\infty \left[\text{BER}/\gamma_{\text{out}}\right] f(\gamma_{\text{out}}) d\gamma_{\text{out}} \quad \propto \left(\frac{1}{\Gamma}\right)^N \quad \Gamma \rightarrow \infty$$

Performance: Outage Probability



Outage probability versus γ_s/Γ

Performance: Bit Error Rate



Diversity Order

- A fundamental parameter of diversity-based systems
- Several times now we have seen that in the high-SNR regime

$$\text{BER or } P_{\text{out}} \propto \left(\frac{1}{\Gamma}\right)^N$$

$$\Rightarrow -\frac{\log(\text{BER})}{\log \Gamma} = N \quad \text{(at high SNR)}$$

- i.e., at high SNR *the slope of the curve in a log-log plot* is N
  - this is the informal definition of diversity order
  - *simulations show that SNR need not be very high for this to hold*

Diversity Order (cont...)

- Formal definition: *The diversity order*, *D, is defined as*

$$D = \lim_{\text{SNR}\to\infty} -\frac{\log \text{BER}}{\log \text{SNR}}$$

The diversity order measures the number of independent paths over which the data is received

- Can also use P_{out} in the definition
- Diversity order is (formally) a high-SNR concept
- Provides information of how useful incremental SNR is
- Sometimes diversity order is abused:
  - the high-SNR definition masks system inefficiencies
  - Note: *both selection and maximal ratio combining have the same diversity order*

The Gains are not due to SNR

Introduction and Overview

Basic Digital Communications

Basic Wireless Communications

Receive Diversity

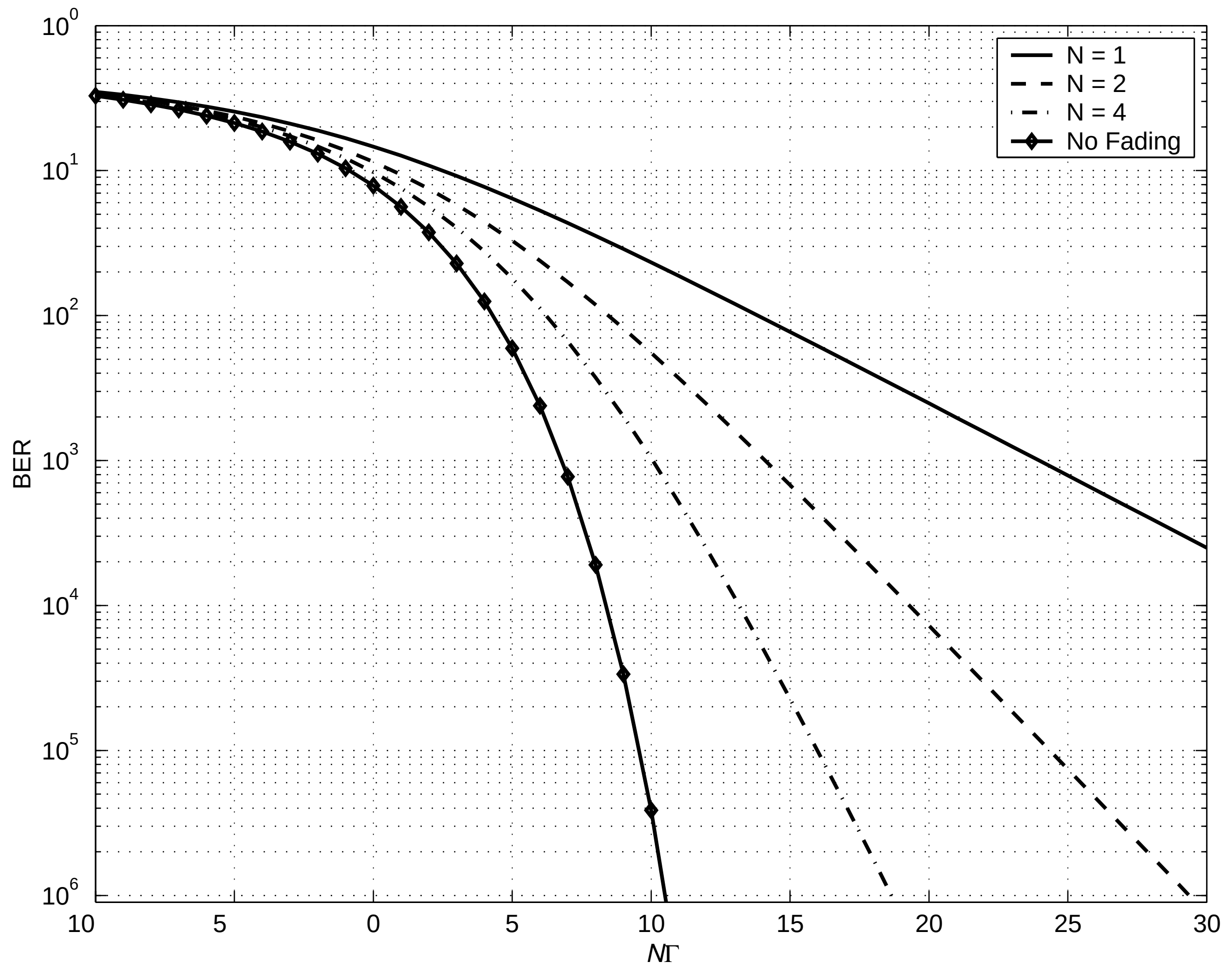
- Introduction
- Diversity Types
- Selection Combining
- Maximal Ratio Combining
- Diversity Order
- Equal Gain Combining
- Comparing Schemes
- Summary
- Correlation

Transmit Diversity

MIMO Information Theory

Course Summary

END



BER versus output SNR. Note that even though output SNR is the same, the BER is significantly different.

Equal Gain Combining

Introduction and Overview

Basic Digital Communications

Basic Wireless Communications

Receive Diversity
- Introduction
- Diversity Types
- Selection Combining
- Maximal Ratio Combining
- Diversity Order
- Equal Gain Combining
- Comparing Schemes
- Summary
- Correlation

Transmit Diversity

MIMO Information Theory

Course Summary

END

- MRC requires matching of both phase and magnitude
  - Magnitude can fluctuate by 10s of dB
  - Biggest gains are by the coherent addition
- Equal Gain Combining: only cancel the phase of the channel

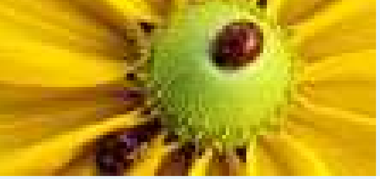
$$w_n = e^{j\angle h_n}$$

- And so...

$$\begin{aligned}\text{Output Signal} = y &= \mathbf{w}^H \mathbf{x} = \sum_{n=1}^{N} w_n^* x_n \\ &= s\left[\sum_{n=1}^{N} |h_n|\right] + \text{noise}\end{aligned}$$

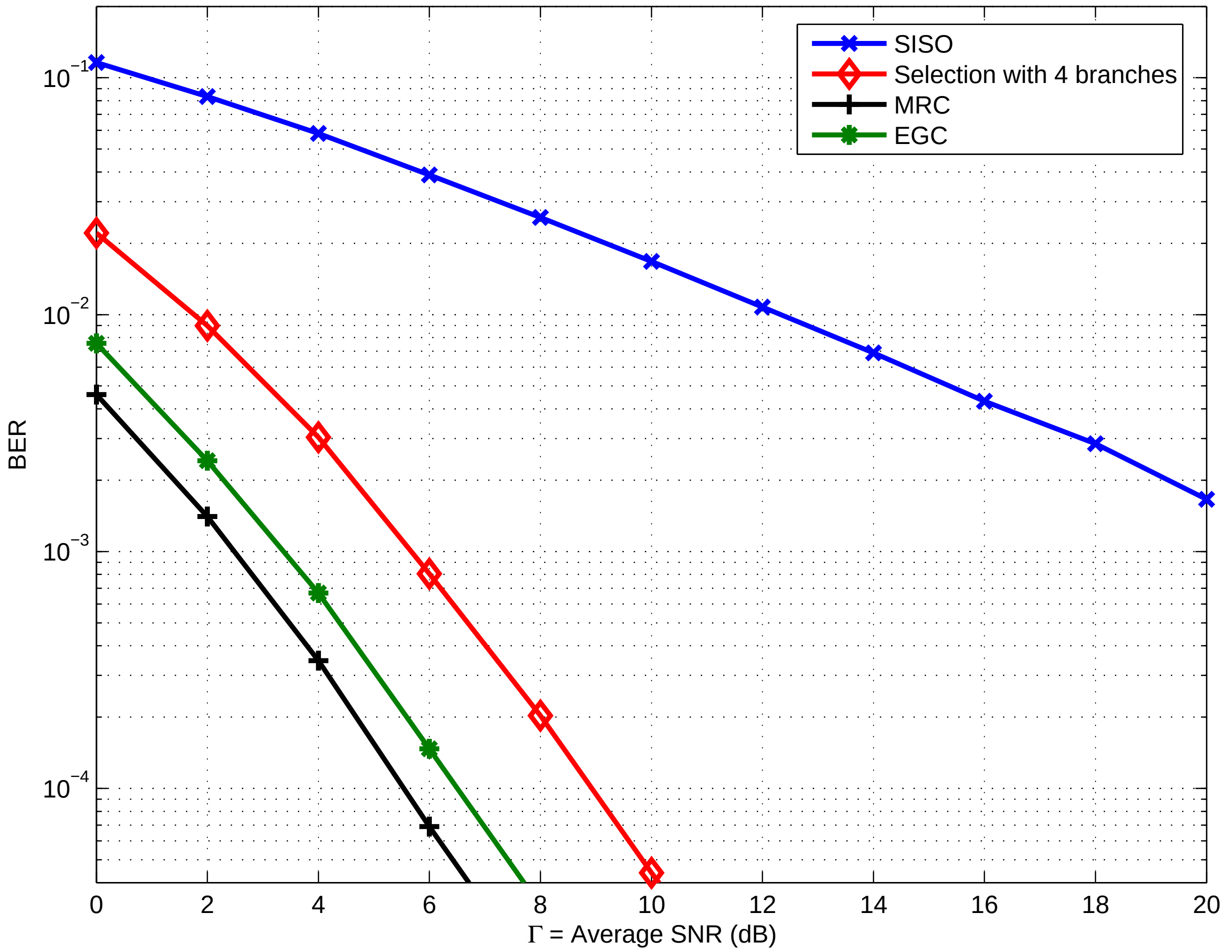
- ...resulting in a small loss in SNR...

$$\text{Average Output SNR} = \left[1 + (N-1)\frac{\pi}{4}\right]\Gamma$$



Comparing Diversity Schemes

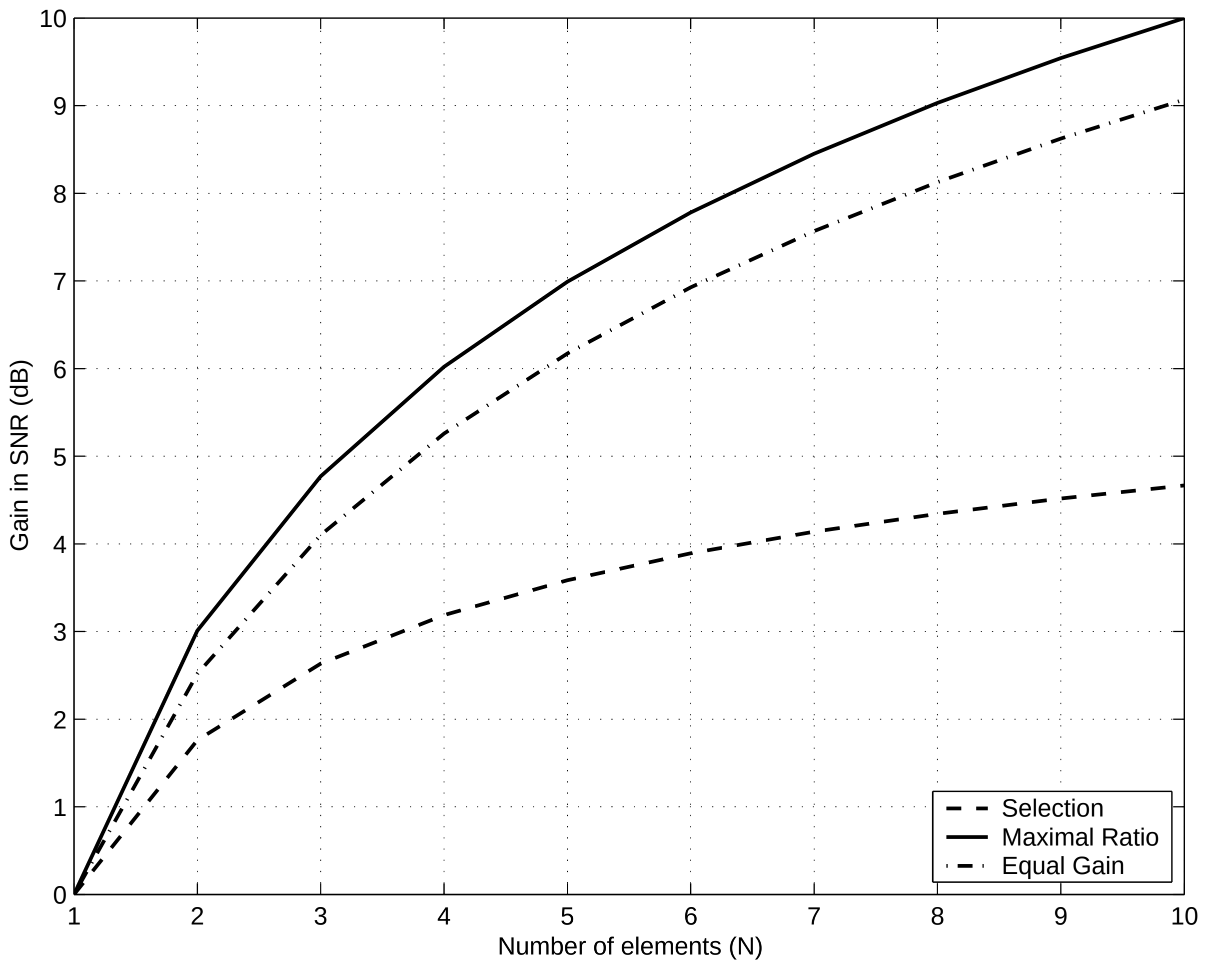
Bit error rate:



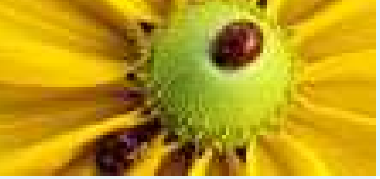
Note: MRC and EGC have similar performance

Comparing Diversity Schemes (cont...)

Gains in SNR:

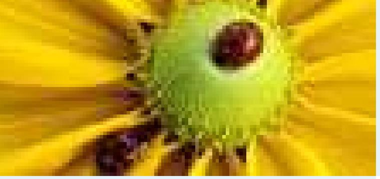


Note: MRC and EGC have similar performance



Comparing Diversity Schemes (cont...)

- All these diversity schemes have same diversity order
- "Work" by reducing *fluctuations* in overall channel
- Selection Combining
  - Simple to implement; only requires power measurement
  - Gain in SNR = $\ln(N)$
- Maximal Ratio Combining
  - Optimal in SNR sense
  - Gain in SNR = N
  - Requires knowledge of channel and matching over several 10s of dB
  - Easiest to analyze
- Equal Gain Combining
  - Small loss w.r.t. MRC
  - Very difficult to analyze, but may be a good trade off for implementation



In Summary...

Diversity is based on providing the receiver with multiple independent copies of the same signal

- The key is the independence between the copies of the same signal
  - The independence makes the gains in error rates exponential with linear gains in number of elements

So, the question is ....under what circumstances can we assume independence?

The Issue of Correlation

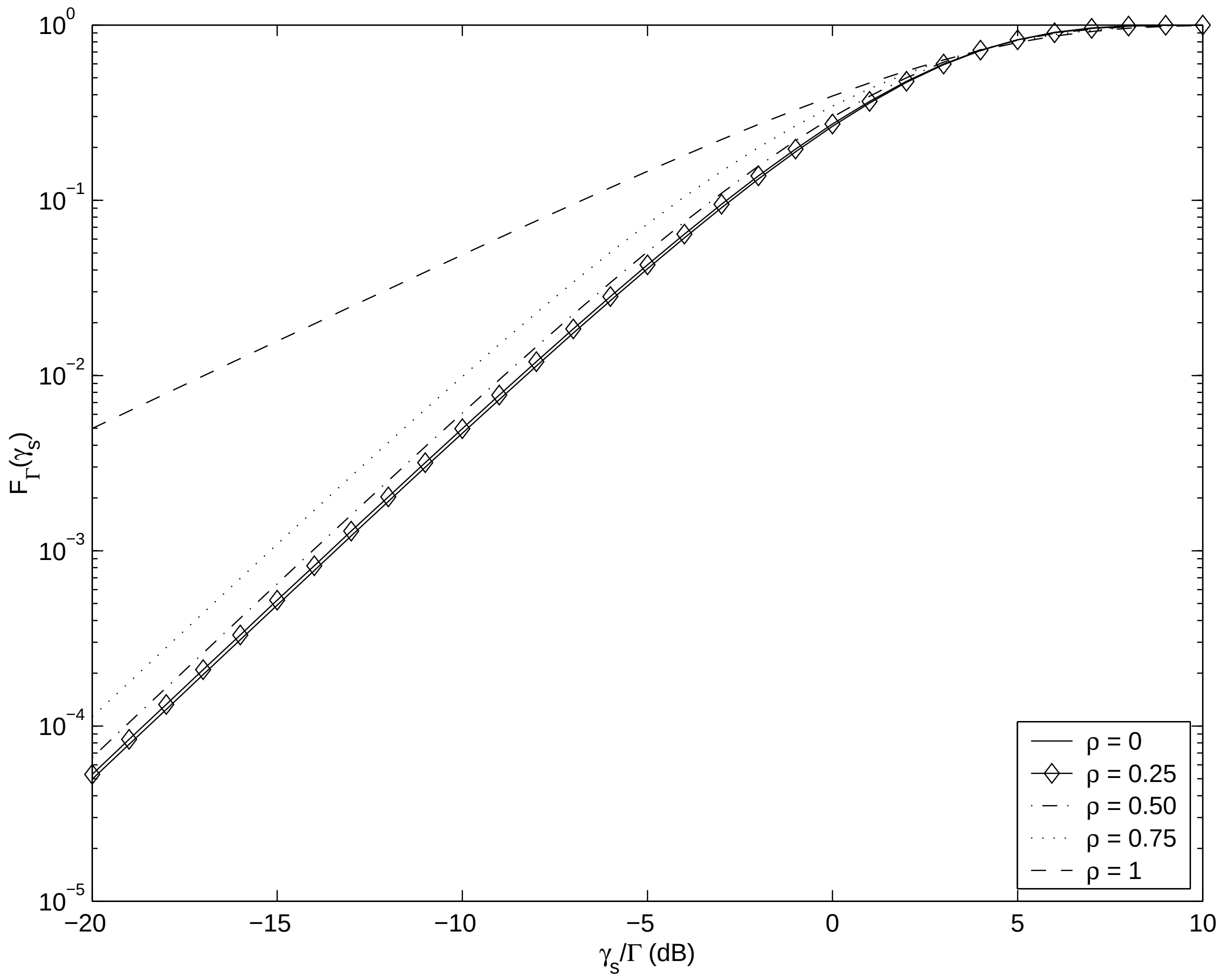
- Correlation between the received signals reduces the independence and hence the effective diversity order
- The extreme case: if all elements were perfectly correlated (e.g., line of sight conditions), diversity order = 1 (only SNR gains)

  For two receive antennas, with correlation of ρ:

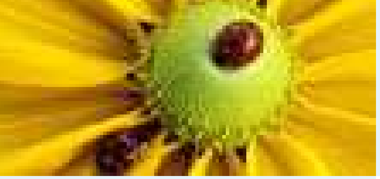
$$
\begin{aligned}
f(\gamma_{\text{out}}) &= \frac{1}{2|\rho|\Gamma}\left[e^{-\gamma_{\text{out}}/(1+|\rho|\Gamma)} - e^{-\gamma_{\text{out}}/(1-|\rho|\Gamma)}\right] \\
P_{\text{out}}(\gamma_s) &= 1 - \frac{1}{2|\rho|}\left[(1+|\rho|)e^{-\gamma_s/(1+|\rho|\Gamma)} \right. \\
&\quad - \left. (1-|\rho|)e^{-\gamma_s/(1-|\rho|\Gamma)}\right]
\end{aligned}
$$

- Correlation arises because
  - Electromagnetic Mutual Coupling
  - Finite distance between elements

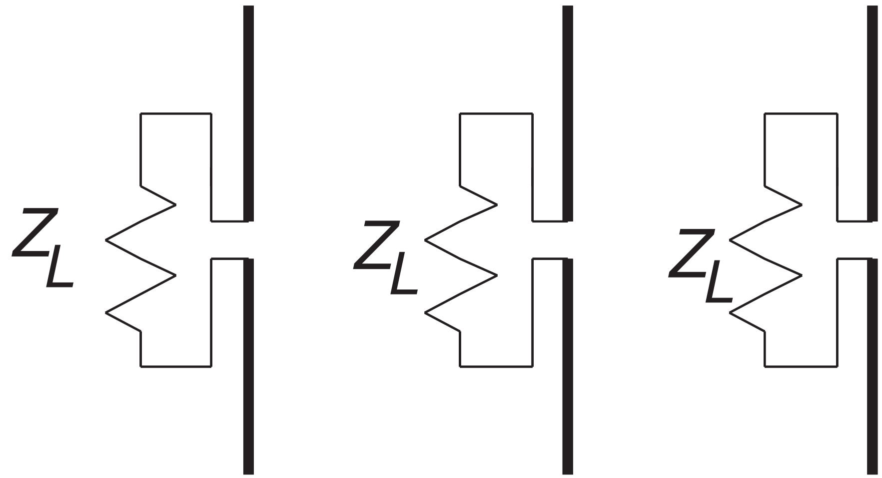
Impact of correlation : Outage Probability



Correlation below $\rho = 0.5$ is considered negligible

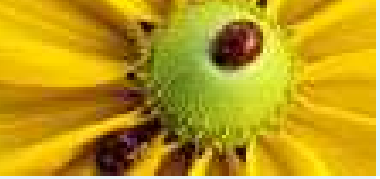


Mutual Coupling



$$\mathbf{V}_{\mathrm{oc}} = \left[\mathbf{Z} + \mathbf{Z}_L\right] \mathbf{Z}_L^{-1} \mathbf{V}$$

- $\mathbf{Z}$: A mutual impedance matrix
- $\mathbf{Z}_L$: Diagonal load matrix
- $\mathbf{V}_{\mathrm{oc}}$: Open circuits voltages that would arise without mutual coupling
- $\mathbf{V}$: True received voltages



Correlation due to Distance

- If spacing between elements is d and signal arrives from direction (θ, ϕ) only, correlation is given by

$$\rho(\theta, \phi) = e^{jkd\cos\phi\sin\theta},$$

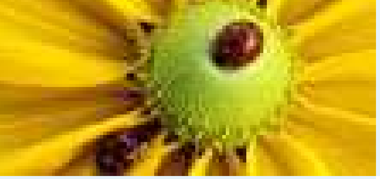
where $k = 2\pi/\lambda$

- Total correlation is therefore averaged over *angular power distribution*

$$\rho = E\{\rho(\theta, \phi)\} = \int_0^{\pi}\int_0^{2\pi} e^{jkdcos\phi\sin\theta} f_{\Theta,\Phi}(\theta, \phi)\, d\phi\, d\theta,$$

where $f_{\Theta,\Phi}(\theta, \phi)$ is the power distribution of the received signals over all angles

- So, the angular distribution is crucial
  - Depends on where the receiver is
    - at the mobile or the base station
    - the base station “looks down” on the mobile



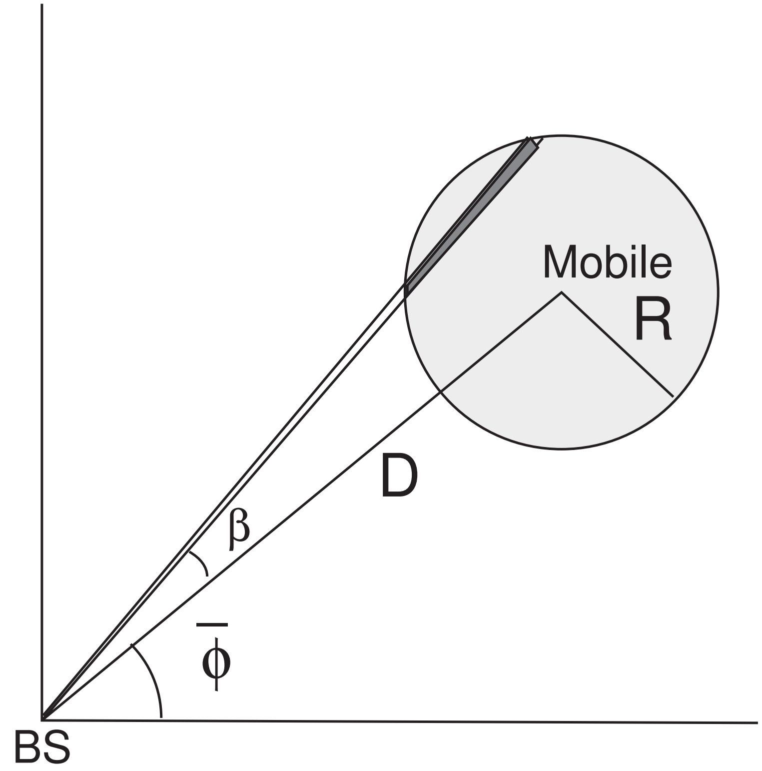
Correlation at the Mobile

- The mobile is (usually) surrounded by many scatterers
- In a dense multipath environment

$$
\begin{aligned}
f_{\Theta,\Phi}(\theta,\phi) &= \frac{1}{2\pi}\delta(\theta-\theta_0) \\
\rho &= \int_{\theta,\phi}\frac{1}{2\pi}\delta(\theta-\theta_0)e^{jkdcos\phi\sin\theta}d\theta d\phi \\
&= J_0(kd\sin\theta_0)
\end{aligned}
$$

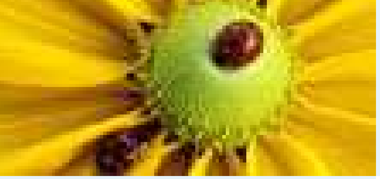
- $\theta_0 = \pi/2, \Rightarrow \rho < 0.5$ if $\underline{d > 0.24\lambda}$
  - Required distance increases as θ_0 decreases
  - Rule of thumb: $d \geq \lambda/2$
  - At 1GHz, $\lambda = 30$cm, i.e., received signals independent if $d >$ 15cm

Correlation at a Base Station



- Signal arrives from a small angular region surrounding the mobile

$$\rho = \int_{-\beta_{\max}}^{\beta_{\max}} e^{-jkd\cos(\bar{\phi}+\beta)\sin\theta_0} f_B(\beta)\,d\beta,$$



Correlation at a Base Station (cont...)

- Array along x-axis
- For a uniform disk of scatterers, $\bar{\phi} = \pi/2$

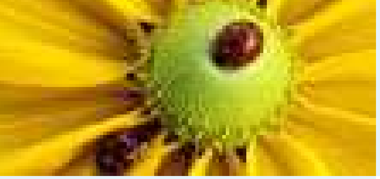
$$\rho = \frac{2J_1(kd\sin\beta_{\max}\sin\theta_0)}{kd\sin\beta_{\max}\sin\theta_0}.$$

- $R = 1.2$km, $D = 50$m, $\theta_0 = 80^o$ $\rho < 0.5$ for $d > 9\lambda$
  - At 1GHz, received signals independent if $d >$ 2.7m (approx. 9ft.)

The required distance is therefore determined by the array setting

Transmit Diversity

- So far, we had a SIMO situation: a single transmitter and multiple receivers
  - Achieving diversity was relatively easy: each receiver receives a copy of the transmitted signal
    - multiple receivers $\Rightarrow$ multiple copies
- What about the MISO situation?
  - Very useful in the expected asymmetrical communication scenarios with more traffic from base station to mobile
    - Base station is expensive, has more space, has multiple antennas
    - Mobile is cheap, has little space, has one antenna
    - Users are downloading information, e.g., a webpage



Transmit Diversity (cont...)

- MISO: N transmit antennas, one receive antenna
- Transmit diversity requires the time dimension. To see this, consider if we did not use the time dimension. Each transmit antenna transmits symbol s.

$$\text{Received Signal} = x = \sum_{n=1}^{N} h_n s + \text{noise}$$

$$= hs + n$$

$$\text{where} \quad h = \sum_{n=1}^{N} h_n$$

- Received signal is a scalar, there is no diversity here!!
  - And hence the concept of space-time coding

Space-Time Coding

- Simple example: Transmit the same symbol over two time slots (symbol periods)
  - On time slot 1, antenna $n = 1$ transmits symbol s

$$x_1 = h_1 s + n_1$$

  - On time slot 2, antenna $n = 2$ transmits the same symbol s

$$x_2 = h_2 s + n_2$$

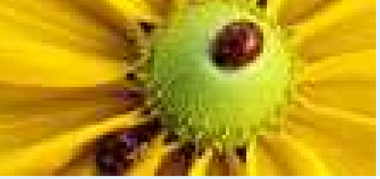
- At the receiver form a receive *vector* over the two time slots

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} s + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \mathbf{h}s + \mathbf{n}$$

- Maximum Ratio Combining:

$$y = \mathbf{h}^H \mathbf{x} = \left(|h_1|^2 + |h_2|^2\right) s + \text{noise}$$

and we would get order-2 diversity



The (Famous) Alamouti's Code

- The previous scheme "wastes" half the time
- A more efficient approach: consider two symbols s_1 and s_2
- In the first time slot,
  antenna $n = 1$ transmits s_1 and antenna $n = 2$ transmits s_2

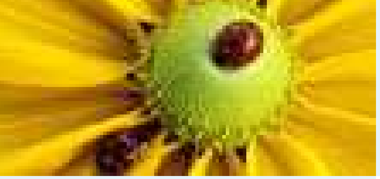
$$x_1 = h_1 s_1 + h_2 s_2 + n_1$$

- In the second time slot,
  antenna $n = 1$ transmits $-s_2^*$ while antenna $n = 2$ transmits s_1^*

$$x_2 = -h_1 s_2^* + h_2 s_1^* + n_2$$

- One subtle, but important point: each element transmits with half the available power
- Form the received data vector (note the conjugate on x_2)

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2^* \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix}$$



Alamouti (cont...)

Now work with $\mathbf{x}$

$$\begin{aligned} \mathbf{x} &= \mathbf{H}\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \mathbf{n} \\ \mathbf{y} &= \mathbf{H}^H\mathbf{x} = \mathbf{H}^H\mathbf{H}\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \text{noise} \\ &= \begin{bmatrix} |h_1|^2 + |h_2|^2 & 0 \\ 0 & |h_1|^2 + |h_2|^2 \end{bmatrix}\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \text{noise} \end{aligned}$$

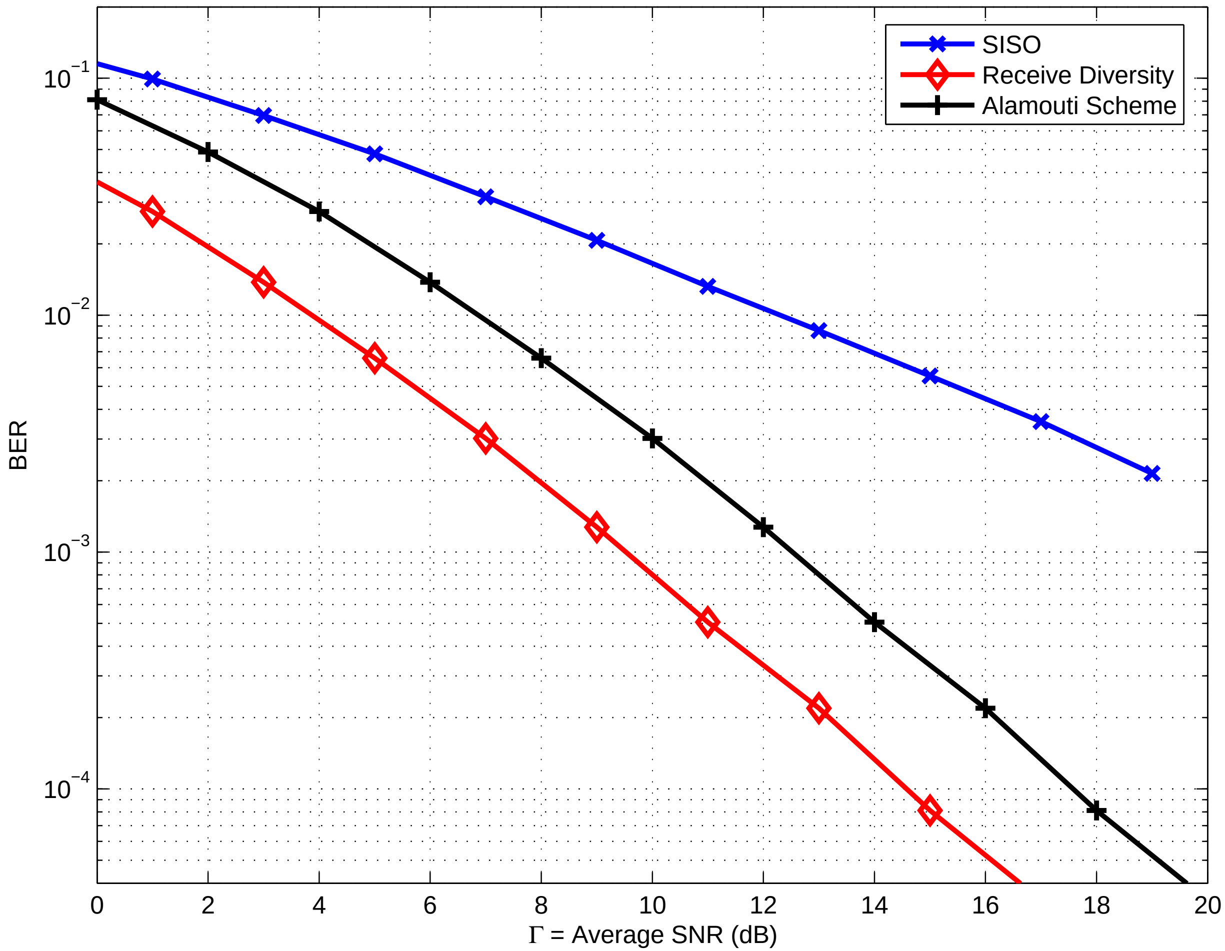
i.e.,

$$\begin{aligned} y_1 &= \left(|h_1|^2 + |h_2|^2\right) s_1 + \text{noise} \\ y_2 &= \left(|h_1|^2 + |h_2|^2\right) s_2 + \text{noise} \end{aligned}$$

$\Rightarrow$ we get order-2 diversity on both symbols!!

The key is that the effective channel matrix is orthogonal

Alamouti Code : Performance



Note: the diversity order of the Alamouti scheme and 2-branch receive diversity is the same. Alamouti's scheme suffers a 3dB loss because of the power splitting

Code Design Criteria: What Makes a Code Go

- Alamouti's scheme is a space-time code
  - A careful organization of data (or a function of the data) in space and time
- Alamouti's scheme is specific to two transmit antennas...
  - ...and, unfortunately, cannot be generalized to $N > 2$
- Consider a code with K symbols over N antennas and L time slots (code rate $R = K/L$)

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1L} \\ c_{21} & c_{22} & \cdots & c_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & \cdots & c_{NL} \end{bmatrix}$$

- here, c_{nl} is a function of the K symbols transmitted in time slot l using antenna n
  - e.g., in the Alamouti scheme, $c_{12} = -s_2^*$

Design Criteria (cont...)

- We want to minimize error rate, the probability that codeword $\mathbf{C}$ was transmitted and $\tilde{\mathbf{C}}$ was decoded

$$P(\mathbf{C} \to \tilde{\mathbf{C}}) \quad \leq \quad \exp\left[-d^2\left(\mathbf{C}, \tilde{\mathbf{C}}\right)\frac{E_s}{4\sigma^2}\right]$$

- E_s is the available energy, $\Gamma = E_s/\sigma^2$
- $d\left(\mathbf{C}, \tilde{\mathbf{C}}\right)$ is the effective "distance" between $\mathbf{C}$ and $\tilde{\mathbf{C}}$

$$d^2\left(\mathbf{C}, \tilde{\mathbf{C}}\right) \quad = \quad \mathbf{h}^H\mathbf{E}\mathbf{E}^H\mathbf{h}$$

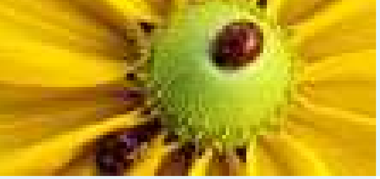
$$\mathbf{E} \quad = \quad \begin{bmatrix} c_{11}-\tilde{c}_{11} & c_{12}-\tilde{c}_{12} & \cdots & c_{1L}-\tilde{c}_{1L} \\ c_{21}-\tilde{c}_{21} & c_{22}-\tilde{c}_{22} & \cdots & c_{2L}-\tilde{c}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1}-\tilde{c}_{N1} & c_{N2}-\tilde{c}_{N2} & \cdots & c_{NL}-\tilde{c}_{NL} \end{bmatrix}$$

- $\mathbf{h} = [h_1,\, h_2,\, \ldots h_N]^T$ is the channel vector

Design Criteria (cont...)

$$\mathbf{E} = \begin{bmatrix} c_{11} - \tilde{c}_{11} & c_{12} - \tilde{c}_{12} & \cdots & c_{1L} - \tilde{c}_{1L} \\ c_{21} - \tilde{c}_{21} & c_{22} - \tilde{c}_{22} & \cdots & c_{2L} - \tilde{c}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1} - \tilde{c}_{N1} & c_{N2} - \tilde{c}_{N2} & \cdots & c_{NL} - \tilde{c}_{NL} \end{bmatrix}$$

- Tarokh et al. developed two design criteria based on this error matrix:
  - The Rank Criterion: The maximum diversity order is achieved if the rank of the error matrix is maximized (N)
    - If M receiving antennas, total diversity order available is NM
  - The Determinant Criterion: The error rate is minimized if the determinant of $\mathbf{E}\mathbf{E}^H$ is maximized *over all code pairs* $\mathbf{C}, \tilde{\mathbf{C}}$



Examples of Space-Time Codes

Introduction and Overview

Basic Digital Communications

Basic Wireless Communications

Receive Diversity

Transmit Diversity
- Transmit Diversity
- Space-Time Coding
- Alamouti's Scheme
- Design Criteria
- Examples of Codes
- Multiplexing
- Summary

MIMO Information Theory

Course Summary

END

- Space-Time Trellis Codes: Design a trellis code over space and time
- Orthogonal Block Codes: Codes in which each symbol can be independently decoded
  - Independent decoding is very convenient
  - Alamouti's code is orthogonal for $N = 2$
  - *Rate-1 orthogonal codes cannot exist for* $N > 2$
  - Rate-1/2 orthogonal codes are always available
    - rate-3/4 codes are available for $N = 3, 4$, e.g.,

$$\mathcal{G}_3 = \begin{pmatrix} s_1 & s_2 & s_3 \\ -s_2 & s_1 & -s_4 \\ -s_3 & s_4 & s_1 \\ -s_4 & -s_3 & s_2 \end{pmatrix},$$

Examples of Codes (cont...)

- Linear Dispersion Codes: Minimize error based on mutual information directly, not the design criteria
- Codes based on algebra: Algebraic codes, e.g., TAST etc.
- *In all cases, space-time codes provide diversity by giving the receiver independent copies of the same message*

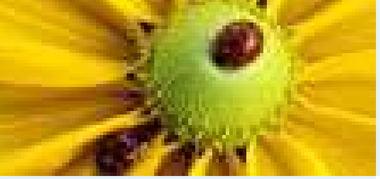
So far we have focused on diversity order (reliability) only.

What about data rate?

Space-Time Coding and Multiplexing

- Transmit more than one data stream (multiplexing)
  - Requires multiple receive antennas as well
- Instead of transmitting only a single data stream, transmit Q data streams in parallel.
- M receive, N transmit antennas. Divide the transmit antennas into Q groups, $N = N_1 + N_2 + \ldots N_Q$
  - Data stream q uses N_q antennas

$$
\begin{aligned}
\mathbf{x} &= \mathbf{H}\mathbf{c} + \mathbf{n}, \\
&= \left[\begin{array}{cccc|ccc}
h_{11} & h_{12} & \cdots & h_{1N_1} & h_{1(N_1+1)} & \cdots & h_{1N} \\
h_{21} & h_{22} & \cdots & h_{2N_1} & h_{2(N_1+1)} & \cdots & h_{2N} \\
\vdots & \vdots & \ldots & \vdots & \vdots & \ddots & \vdots \\
h_{M1} & h_{M2} & \cdots & h_{MN_1} & h_{M(N_1+1)} & \cdots & h_{MN}
\end{array}\right]
\begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_Q \end{bmatrix} \\
&\quad + \mathbf{n}
\end{aligned}
$$



STC and Multiplexing (cont...)

- The q^{th} data stream uses a space-time code over N_q antennas
- How do you isolate individual streams?

$$\mathbf{x} = \mathbf{H}_1 \mathbf{c}_1 + \bar{\mathbf{H}}_1 \begin{bmatrix} \mathbf{c}_2 \\ \mathbf{c}_3 \\ \vdots \\ \mathbf{c}_Q \end{bmatrix} + \mathbf{n}$$

- Now, let $\mathbf{H}_1^{\perp}$ be the null space of $\bar{\mathbf{H}}_1$ ($\mathbf{H}_1^{\perp H} \bar{\mathbf{H}}_1 = 0$)
  - this is possible if $M > N - N_1$. $\mathbf{H}_1^{\perp}$ is size $M \times (M - N + N_1)$

$$\mathbf{y}_1 = \mathbf{H}_1^{\perp H} \mathbf{x} = \left[\mathbf{H}_1^{\perp H} \mathbf{H}_1 \right] \mathbf{c}_1 + \text{noise}$$

which is a space-time coded system with N_1 transmitters and $(M - N + N_1)$ receivers

  - Diversity order = $N_1 \times (M - N + N_1)$

Multiplexing (cont...)

- Clearly we can apply the same idea for $q = 2, \ldots, Q$
  - Data stream q would achieve diversity order of $N_q \times (M - N + N_q)$.
- Can we do better? Yes!! Use interference cancellation...
  - ...since $\mathbf{c}_1$ has been decoded, subtract it!
  - Data stream 2 "sees" less interference (N_1 interfering transmissions are eliminated). Stream 2 can get diversity order of $N_2 \times (M - N + N_2 + N_1)$
  - Similarly, the q^{th} data stream can achieve diversity order of $N_q \times \left(M - N + \sum_{p=1}^{q} N_p\right)$
- BLAST: Bell Labs Layered Space-Time
  - $N_q = 1$
  - Requires $M \geq N$ (at least as many receivers as transmitters)
  - Achieved (in lab) spectral efficiency of 10s of b/s/Hz!

In Summary...

- Transmit diversity requires the time dimension
  - Space-time coding is the careful arrangement of data (or function of data) in space and time to achieve
    - Greatest diversity order (rank criterion)
    - Minimum error rate (determinant criterion)
- Several space-time code families are available
  - We focused on the simplest family of orthogonal space-time block codes
- Can also use the spatial degrees of freedom to multiplex
  - BLAST is one (famous) example

MIMO Information Theory

- We wish to investigate the fundamental limits of data transfer rate in MIMO wireless systems
- Remember:
  - A channel is fundamentally characterized by (and the data rate limited by) its capacity C
  - In the SISO case,

$$C = \log_2(1 + SNR)$$

- Our MIMO system: N transmit and M receive antennas

$$\boxed{\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}}$$

- $\mathbf{y}$: the length-M received signal vector
- $\mathbf{H}$: the $M \times N$ channel
- $\mathbf{x}$: The length-N transmit data vector
  - Let $\mathbf{S}_x = E\{\mathbf{x}\mathbf{x}^H\}$ be the covariance matrix of $\mathbf{x}$

Channel Unknown at Transmitter

- It is not too hard to show

$$C = \log_2 \det \left[\mathbf{I} + \frac{1}{\sigma^2} \mathbf{H} \mathbf{S}_x \mathbf{H}^H \right]$$

- If channel $\mathbf{H}$ is known at the transmitter, $\mathbf{S}_x$ can be chosen to best match the channel
- However, let's start with the case that the channel is not known. The best choice is

$$\begin{aligned} \mathbf{S}_x &= \frac{E_s}{N} \mathbf{I}_{N \times N} \\ \Rightarrow C &= \log_2 \det \left[\mathbf{I} + \frac{1}{\sigma^2} \frac{E_s}{N} \mathbf{H} \mathbf{H}^H \right] \end{aligned}$$

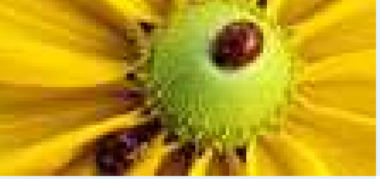
- Since $\mathbf{H}$ is not known, again, one cannot guarantee a data rate and the true capacity is zero!
  - Again, talk of an outage probability and/or expected capacity

Unknown Channel (cont...)

$$C = \log_2 \det \left[\mathbf{I} + \frac{1}{\sigma^2} \frac{E_s}{N} \mathbf{H}\mathbf{H}^H \right]$$

- Since we are assuming Rayleigh fading, the entries of $\mathbf{H}$ are complex Gaussian
  - Experts in STAP will recognize $\mathbf{H}\mathbf{H}^H$ as following the Wishart distribution
- Let the eigenvalues of $\mathbf{H}\mathbf{H}^H$ be $\lambda_m^2, m = 1, 2, \ldots, M$

Eigenvalues of $\left[\mathbf{I} + E_s/(N\sigma^2)\mathbf{H}\mathbf{H}^H\right] = \left(1 + \frac{E_s}{(N\sigma^2)}\lambda_m^2\right)$



Unknown Channel (cont...)

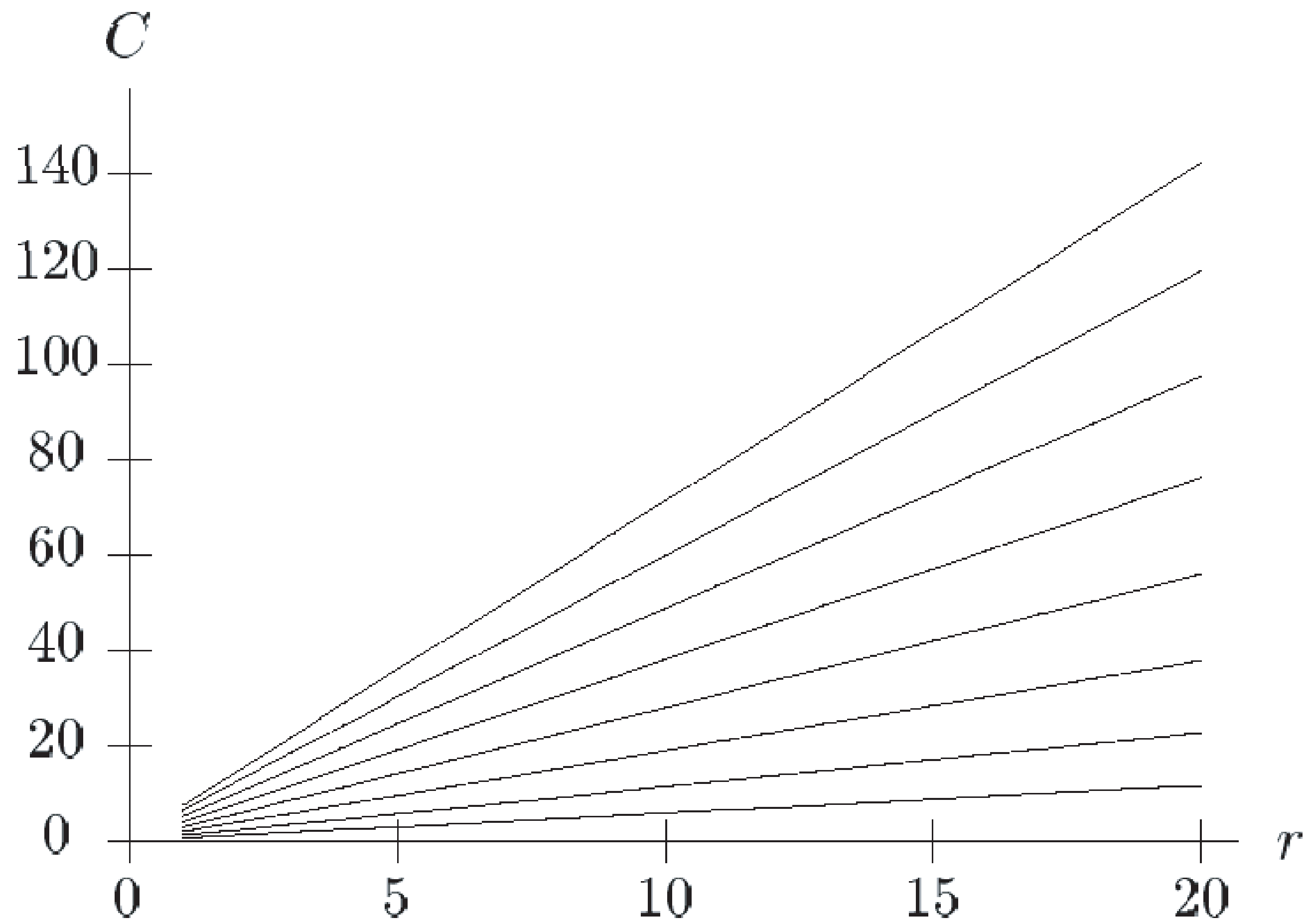
$$C = \log_2 \prod_{m=1}^{M} \left(1 + \frac{E_s}{N\sigma^2}\lambda_m^2\right) = \sum_{m=1}^{M} \log_2 \left(1 + \frac{E_s}{N\sigma^2}\lambda_m^2\right)$$

- The distribution of λ_m is known ($\mathbf{HH}^H$ is Wishart)
  - Without ordering, these eigenvalues are independent and identically distributed
  - There are $r = \min(M, N)$ eigenvalues
- Therefore, on average

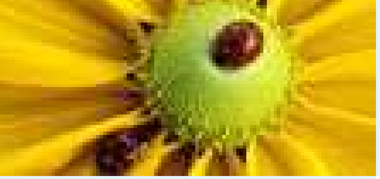
$$\begin{aligned} E\{C\} &= E_{\{\lambda_m\}} \left\{ \sum_{m=1}^{M} \log_2 \left(1 + \frac{E_s}{N\sigma^2}\lambda_m^2\right) \right\} \\ &= \min(N, M) E_\lambda \left\{ \log_2 \left(1 + \frac{E_s}{N\sigma^2}\lambda^2\right) \right\} \end{aligned}$$

- We get *linear* gains in capacity, not just power gains
  - As if we have $\min(N, M)$ parallel channels!

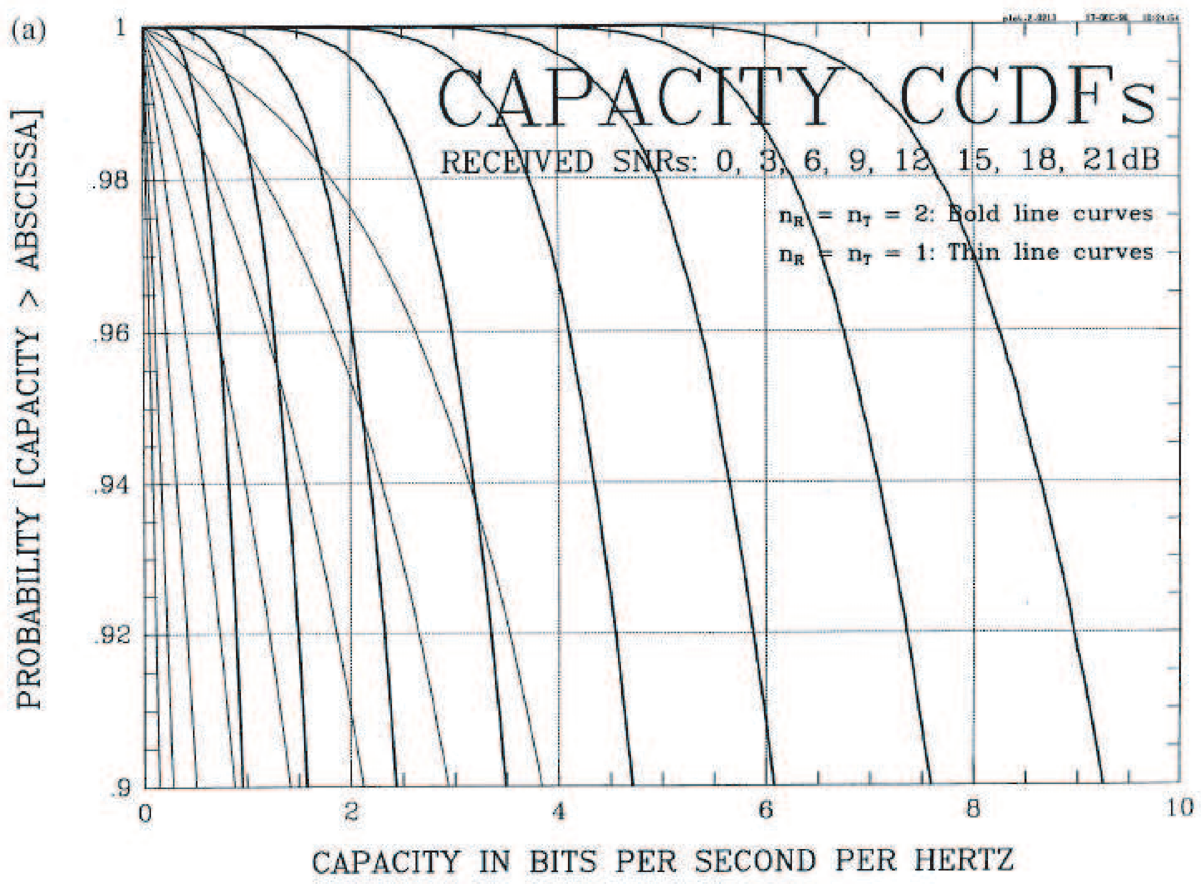
Unknown Channel: Ergodic Capacity



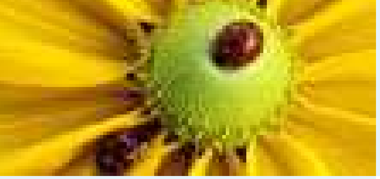
Note: Ergodic capacity for fixed SNR (from Telatar (1999)). Here r is the number of elements in the transmitter and receiver ($r = M = N$)



Unknown Channel: 1-(Outage Probability)



Note: "Success" rates (probability that capacity is above target) from Foschini and Gans (1998). Comparing "success" rates for SISO and a $N = M = 2$ system.



Channel Known at the Transmitter

- What if the channel is known at the transmitter?
  - Usually obtained via feedback (frequency division duplex - FDD) or via reciprocity (TDD)
  - The transmitter can tune the covariance matrix $\mathbf{S}_x$ to match the transmitter
  - This may be via power allocation

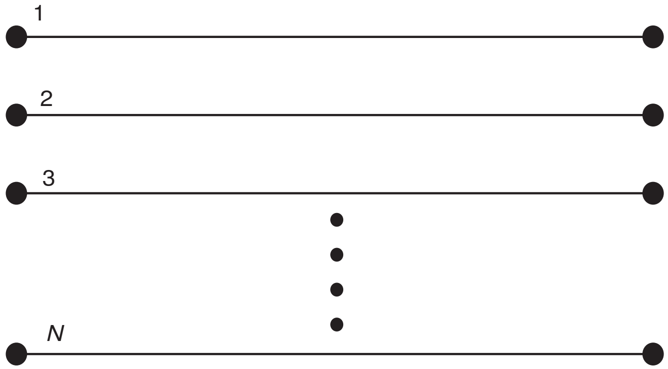
$$\begin{aligned} \mathbf{S}_x^{\text{opt}} &= \max_{\mathbf{S}_x} C \\ &= \max_{\mathbf{S}_x} \log_2 \det \left[\mathbf{I} + \frac{1}{\sigma^2} \mathbf{H}\mathbf{S}_x\mathbf{H}^H \right] \end{aligned}$$

Constraints: $\mathbf{S}_x$ must be positive-definite

$\mathbf{S}_x$ must satisfy a power constraint

- So, how do you optimize over a matrix?
  - Let's start with a simpler (and very instructive) system: parallel channels

Parallel Channels



- Each channel is independent of the other
- On the n^{th} channel

$$y_n = h_n x_n + \text{noise}$$

- The transmitter has one important constraint - a total available energy constraint of E_s
  - Since the transmitter knows the channel values, $h_n, n = 1, 2, \ldots, N$ it can allocate power to maximize the overall capacity

Parallel Channels (cont...)

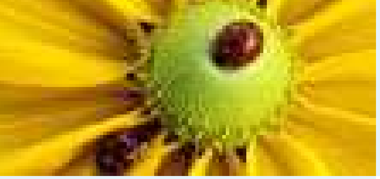
- The transmitter allocates power E_n to channel n

$$C = \sum_{n=1}^{N} \log_2 \left(1 + |h_n|^2 \frac{E_n}{\sigma^2} \right)$$

- Intuitively, the transmitter should allocate all its power to the strongest channel, right?
  - Strangely enough, “wrong”!
- This is because...

$$C = \log\left(1 + \text{SNR}\right)$$

- At high SNR, $C \sim \log\left(\text{SNR}\right)$
- At low SNR, $C \sim \text{SNR}$
  - There are diminishing marginal returns in allocating power



Parallel Channels (cont...)

The problem formulation:

$$\{E_n^{\text{opt}}\} = \max_{\{E_n\}} \sum_{n=1}^{N} \log_2 \left(1 + \frac{E_n |h_n|^2}{\sigma_n^2}\right)$$

$$\sum_{n=1}^{N} E_n \leq E_s$$

$$E_n \geq 0$$

The solution:

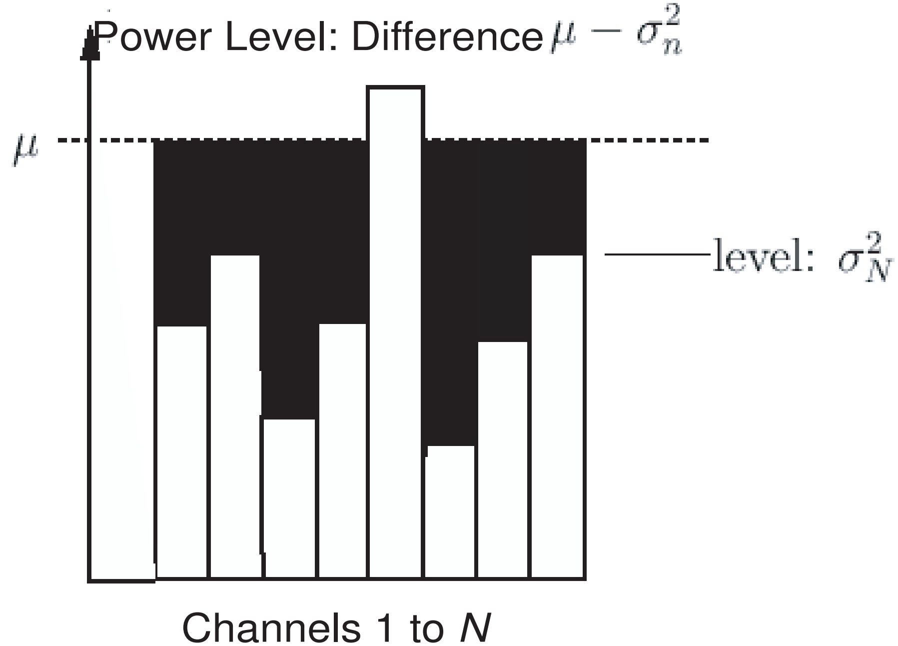
$$\left(\frac{\sigma^2}{|h_n|^2} + E_n\right) = \mu, \qquad n = 1, \ldots, N$$

$$E_n = \left(\mu - \frac{\sigma^2}{|h_n|^2}\right)^+,$$

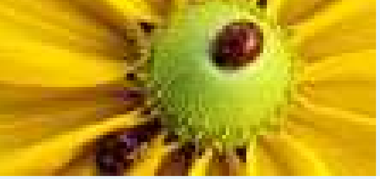
where $(x)^+ = 0$ if $x < 0$ and $(x)^+ = x$ if $x \geq 0$

Waterfilling

- Note that μ is a constant
- Channel sees an "effective" noise variance of $\sigma_n^2 = \sigma^2/|h_n|^2$



- Note that the better channels do get more power
- Some channels are so "bad" that they do not get any power



MIMO Systems with Known Channel

- So far, we have focused on parallel channels.

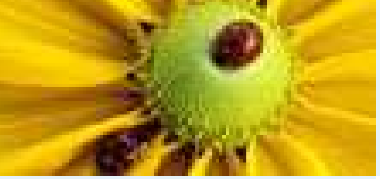
$$\text{So far,} \quad \mathbf{y} = \begin{bmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_N \end{bmatrix} \mathbf{x} + \mathbf{n}$$

- What does this tell us about a “regular” MIMO system?
- We have

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$$

- N transmitters, M receivers, $\mathbf{H}$ is the $M \times N$ channel

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N} \\ h_{21} & h_{22} & \cdots & h_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1} & h_{M2} & \cdots & h_{MN} \end{bmatrix}$$



MIMO Systems (cont...)

- One can use the singular value decomposition of $\mathbf{H}$

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$$

- $\mathbf{U}$ is the matrix of eigenvectors of $\mathbf{H}\mathbf{H}^H$
- $\mathbf{V}$ is the matrix of eigenvectors of $\mathbf{H}^H\mathbf{H}$
- $\boldsymbol{\Sigma}$ is a “diagonal” $M \times N$ matrix of singular values

$$\boldsymbol{\Sigma} = \left[\begin{array}{cccc|cc} \sigma_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_R & 0 & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{array}\right],$$

- R is the rank of $\mathbf{H}$
- The zeros pad the matrix to match the $M \times N$ dimensions

MIMO Systems (cont...)

- How does this help? We have

$$
\begin{aligned}
\mathbf{U}\mathbf{U}^H &= \mathbf{U}^H\mathbf{U} = \mathbf{I}_{M\times M} \\
\mathbf{V}\mathbf{V}^H &= \mathbf{V}^H\mathbf{V} = \mathbf{I}_{N\times N}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{n} \\
&= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H\mathbf{x} + \mathbf{n} \\
\Rightarrow \mathbf{U}^H\mathbf{y} &= \boldsymbol{\Sigma}\mathbf{V}^H\mathbf{x} + \mathbf{U}^H\mathbf{n}
\end{aligned}
$$

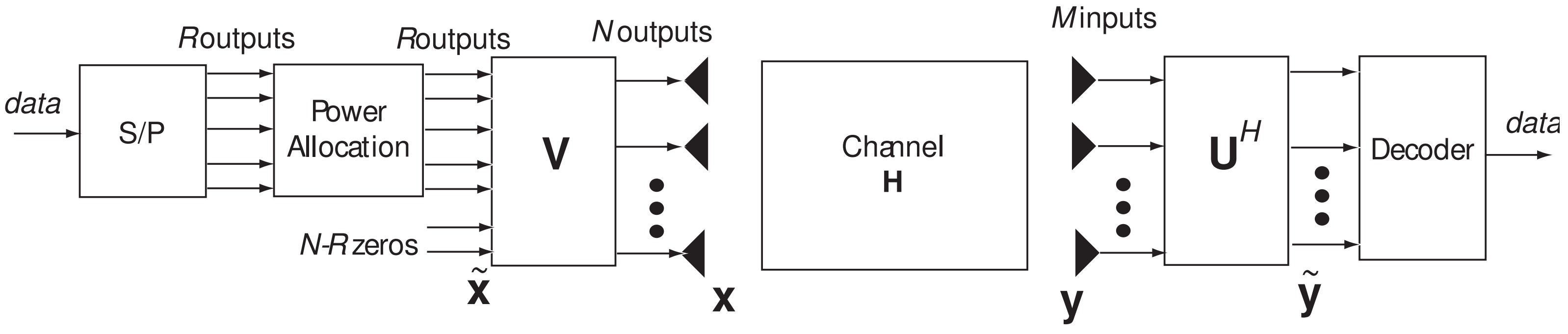
- $\tilde{\mathbf{y}} = \mathbf{U}^H\mathbf{y}$, $\tilde{\mathbf{x}} = \mathbf{V}^H\mathbf{x}$, $\tilde{\mathbf{n}} = \mathbf{U}^H\mathbf{n}$

$$
\boxed{\tilde{\mathbf{y}} = \Sigma\tilde{\mathbf{x}} + \tilde{\mathbf{n}}}
$$

- We have a set of R parallel channels!

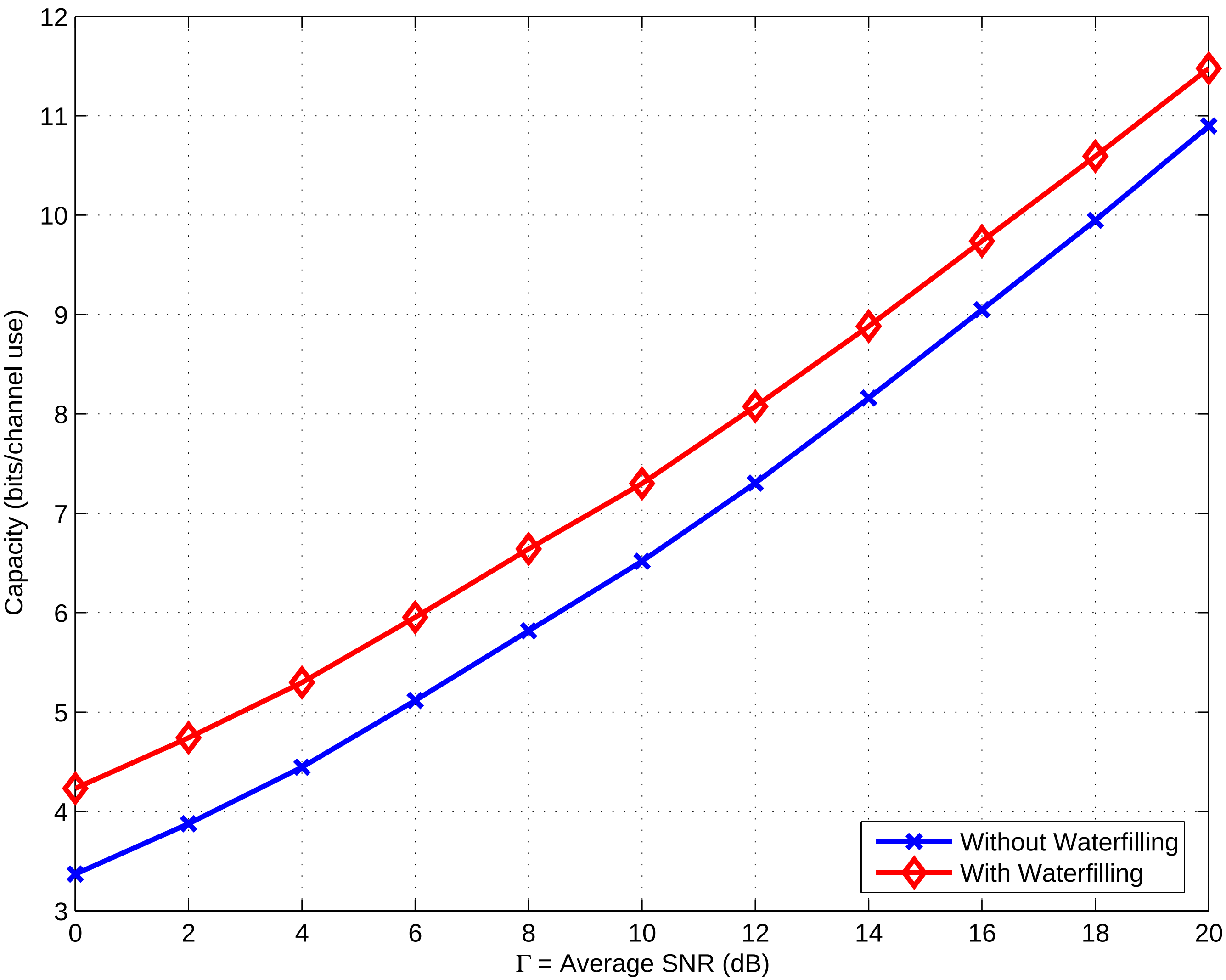
MIMO Systems (cont...)

Therefore,



- The transmitter precodes the transmitted signal using matrix $\mathbf{V}$. This matches the transmission to the "eigen-modes" of the channel
- The transmitter also waterfills over the singular values σ_n as the equivalent channel values as parallel channels (called h_n earlier)
- The receiver decodes using the matrix $\mathbf{U}$
- IMPORTANT: This diagonalization process (transmission on eigen-modes is a fundamental concept in wireless communications

Improvement in Capacity



Comparing capacities for $N = M = 4$

Some Illustrative Examples

- Example 1: SIMO System, 1 transmitter, M receivers

$$\begin{aligned}
\mathbf{H} = \mathbf{h} &= \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix} \\
\Rightarrow \mathbf{U} &= \begin{bmatrix} \frac{\mathbf{h}}{||\mathbf{h}||} & \mathbf{h}^{\perp} \end{bmatrix} \\
\mathbf{V} &= [1] \\
\mathbf{\Sigma} &= [||\mathbf{h}||, 0, \ldots, 0]^T
\end{aligned}$$

- Note that we have only a single “parallel” channel

$$C = \log_2\left(1 + \frac{E_s}{\sigma^2}||\mathbf{h}||^2\right)$$

- Effectively, all channel powers added together

Some Illustrative Examples (cont...)

Introduction and Overview
Basic Digital Communications
Basic Wireless Communications
Receive Diversity
Transmit Diversity
MIMO Information Theory
- Basics
- Unknown Channel
- Known Channel
- Waterfilling
- MIMO Systems
- Examples
- Summary
- DMT

Course Summary
END

- Example 2: MISO System, N transmitters, 1 receiver

$$\begin{aligned} \mathbf{H} = \mathbf{h} &= [h_1, h_2, \ldots, h_N] \\ \Rightarrow \mathbf{U} &= [1] \\ \mathbf{V} &= \left[\frac{\mathbf{h}}{||\mathbf{h}||} \quad \mathbf{h}^{\perp} \right] \\ \mathbf{U} &= [1] \\ \mathbf{\Sigma} &= [||\mathbf{h}||, 0, \ldots, 0] \end{aligned}$$

- Again we have only a single “parallel” channel

$$C = \log_2 \left(1 + \frac{E_s}{\sigma^2} ||\mathbf{h}||^2 \right)$$

- This is the same as the SIMO case!
- Note: without channel knowledge,

$$C = \log_2 \left(1 + \frac{E_s}{N\sigma^2} ||\mathbf{h}||^2 \right)$$

Some Illustrative Examples (cont...)

- Example 3: MIMO System, line of sight conditions

$$\mathbf{s}(\phi_r) = \left[1, z_r, z_r^2, \ldots, z_r^{M-1}\right]^T, \qquad z_r = e^{jkd_r \cos\phi_r}$$

$$\mathbf{s}(\phi_t) = \left[1, z_t, z_t^2, \ldots, z_t^{N-1}\right]^T, \qquad z_t = e^{jkd_t \cos\phi_t}$$

$$\begin{aligned} \mathbf{H} &= \mathbf{s}(\phi_r) \otimes \mathbf{s}^T(\phi_t) \\ &= \begin{bmatrix} 1 & z_t & z_t^2 & \cdots & z_t^{N-1} \\ z_r & z_r z_t & z_r z_t^2 & \cdots & z_r z_t^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_r^{M-1} & z_r^{M-1} z_t & z_r^{M-1} z_t^2 & \cdots & z_r^{M-1} z_t^{N-1} \end{bmatrix} \end{aligned}$$

- This is a rank-1 matrix!
  - This one singular value = $\sigma_1 = \sqrt{NM}$

$$C = \log_2\left(1 + \frac{E_s}{N\sigma^2} NM\right) = \log_2\left(1 + \frac{E_s}{\sigma^2} M\right)$$

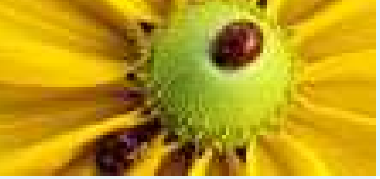
Some Illustrative Examples (cont...)

- Example 4: MIMO System, $M = N$, rich scattering conditions, full rank channel
  - We have N parallel channels
- For convenience, assume all singular values are equal
  - Let this singular value = σ_1
  - Since all parallel channels are equally powerful, power allocation is uniform

$$C = \sum_{n=1}^{N} \log_2 \left(1 + \frac{E_s}{N\sigma^2} \sigma_1^2 \right) = N \log_2 \left(1 + \frac{E_s}{N\sigma^2} \sigma_1^2 \right)$$

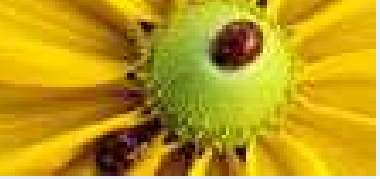
- Note the huge difference from “line of sight” scenario
  - The number of transmit or receive elements is outside the log term; we get linear gains in capacity

If $N \neq M$ $\qquad C = \min(N, M) \log_2 \left(1 + \frac{E_s}{N\sigma^2} \sigma_1^2 \right)$



Summary of Information Theoretic Analysis

- MIMO systems allow for huge increases in capacity
  - If the fading is independent then one can achieve linear gains in capacity over the SISO case
  - Notice that this inherently requires the concept of a diversity of paths
- The concept of diagonalization or transmission on eigen-channels and the associated concept of waterfilling are fundamental
- Waterfilling allocates more power to better channels
  - Note that this is, initially, counter-intuitive. Generally, if we have a poor channel, we add power, not reduce power



Diversity-Multiplexing Tradeoff

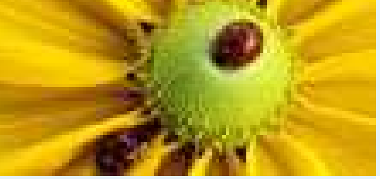
Consider a system with N transmit and M receive antennas in a rich scattering, Rayleigh fading environment

- Diversity: We have seen that through space-time coding and receive diversity we can achieve a diversity order of NM.
- Multiplexing: In the information theoretic analysis we saw that we could get a pre-log factor of $\min(M, N)$. Also, at high SNR

$$C \to \min(M, N) \log_2(\text{SNR})$$

Q: Can we get both diversity and multiplexing (rate) gains?

A: Yes! But, there is a trade-off between the two!



DMT (cont...)

- As before, define the diversity order as:

$$D = \lim_{\text{SNR}\to\infty} \left[\frac{\log P_{\text{out}}}{\log \text{SNR}} \right]$$

  - D tells us how fast the error rate falls with increases in log(SNR)

- Define a multiplexing gain r as

$$r = \lim_{\text{SNR}\to\infty} \left[\frac{R}{\log \text{SNR}} \right]$$

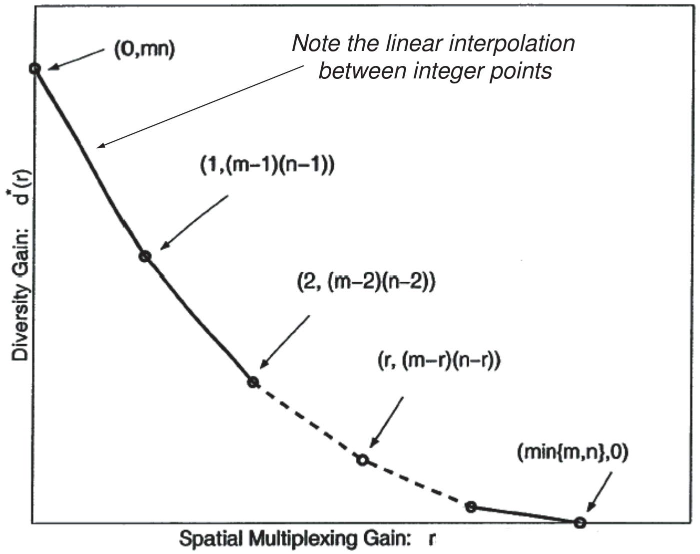
  - R is the rate of transmission
  - r is the rate at which the transmission rate increases with log(SNR)

DMT (cont...)

- Diversity Multiplexing Tradeoff: The optimal tradeoff curve, $d^\star(r)$ is given by the piecewise-linear function connecting the points $(r, d^\star(r))$, $r = 0, 1, \ldots, \min(M, N)$, where

$$d^\star(r) = (M - r)(N - r)$$

- Note: $d_{\max} = MN$ and $r_{\max} = \min(M, N)$.
- At integer points, r degrees of freedom are used for multiplexing, the rest are available for diversity



Course Summary

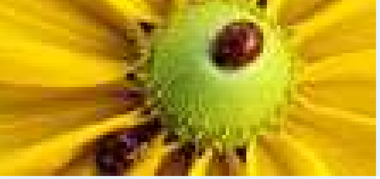
- We have explored uses of the spatial dimension in wireless communications
- Performance of wireless systems is fundamentally limited by fading
  - Fading makes the received signal to be a random copy of the transmitted signal
  - MIMO systems are based on independent fading to/from multiple elements
- We covered three major concepts:
  - Receive Diversity
  - Transmit Diversity
  - MIMO Information Theory
- Receive Diversity:
  - Selection, maximal ratio and equal gain combining
    - Trade off between complexity and performance

Course Summary (cont...)

- Receive Diversity (cont...)
  - Notion of diversity order
    - Measures the number of independent paths the signal is received
  - Impact of correlation
    - Array elements must be some minimum distance apart
    - The key is to create independence
- Transmit Diversity:
  - Requires the time dimension: space-time coding
  - An important example of block codes: Alamouti's Code
    - Unfortunately, only rate 1/2 codes are guaranteed for complex data constellations
  - Good codes designed using the rank and determinant criteria

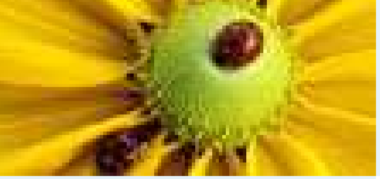
Course Summary (cont...)

- Multiplexing:
  - Transmitting multiple data streams simultaneously
  - Can combine space-time coding with multiplexing
    - A famous example: the BLAST scheme
    - We should emphasize that this is only one possible multiplexing scheme
- MIMO Information Theory:
  - Capacity via determinant of channel
  - Ergodic capacity and outage probability if channel is unknown
  - Transmission on eigen-channels and using waterfilling if channel is known to transmitter
    - Creates $r = \text{rank}\left(\mathbf{H}\right)$ parallel channels
  - There is a fundamental trade-off between diversity and multiplexing



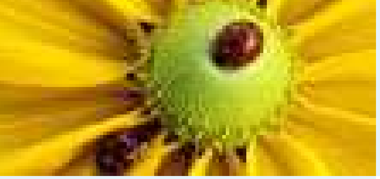
We could have done so much more...

- Channel Estimation: Throughout we assumed the receiver knows the channel
  - The receiver has to estimate and track a time varying channel
- Frequency selective channels: We focused on flat channels, frequency selectivity is becoming important
  - Everyone assumes 4G will be based on OFDM - MIMO-OFDM is a "hot" topic
- Error control coding to achieve the capacity of MIMO systems
- Feedback to inform the transmitter of the channel
  - Low data rate schemes, error bounds, impact of error



What else? (cont...)

- Multiuser Communications: Transmitting to/receiving from multiple users simultaneously
  - Is another form of multiplexing; requires interference cancellation, provides flexibility
  - Is getting more important, both theoretically and in implementation
- Cooperative Communications:
  - Probably the "hottest" research area now
  - Nodes with a single antenna share resources to act like a MIMO system
    - via relaying, forwarding, cooperative diversity
  - Especially applicable to the new "modern" kinds of networks. Also, distributed signal processing.
    - Sensor Networks: Large scale networks of small, cheap nodes
    - Mesh Networks: Networks of access points



END

That all folks!

Introduction and Overview
Basic Digital Communications
Basic Wireless Communications
Receive Diversity
Transmit Diversity
MIMO Information Theory
Course Summary
END

Again, detailed notes available at

http://www.comm.utoronto.ca/˜ rsadve/teaching.html