Fundamentals of Wireless Communication

David Tse
Dept of EECS
U.C. Berkeley

Course Objective

- Past decade has seen a surge of research activities in the field of wireless communication.
- Emerging from this research thrust are new points of view on how to communicate effectively over wireless channels.
- The goal of this course is to study in a unified way the fundamentals as well as the new research developments.
- The concepts are illustrated using examples from several modern wireless systems (GSM, IS-95, CDMA 2000 1x EV-DO, Flarion's Flash OFDM, ArrayComm systems.)

Course Outline

Part I: Basics

1. The Wireless Channel
2. Diversity
3. Capacity of Wireless Channels

Course Outline (2)

Part II: MIMO

4. Spatial Multiplexing and Channel Modelling

5. Capacity and Multiplexing Architectures

6. Diversity-Multiplexing Tradeoff

Course Outline (3)

Part III: Wireless Networks

7. Multiple Access and Interference Management: A comparison of 3 systems.

8. Opportunistic Communication and Multiuser Diversity

This short course only gives an overview of the ideas.

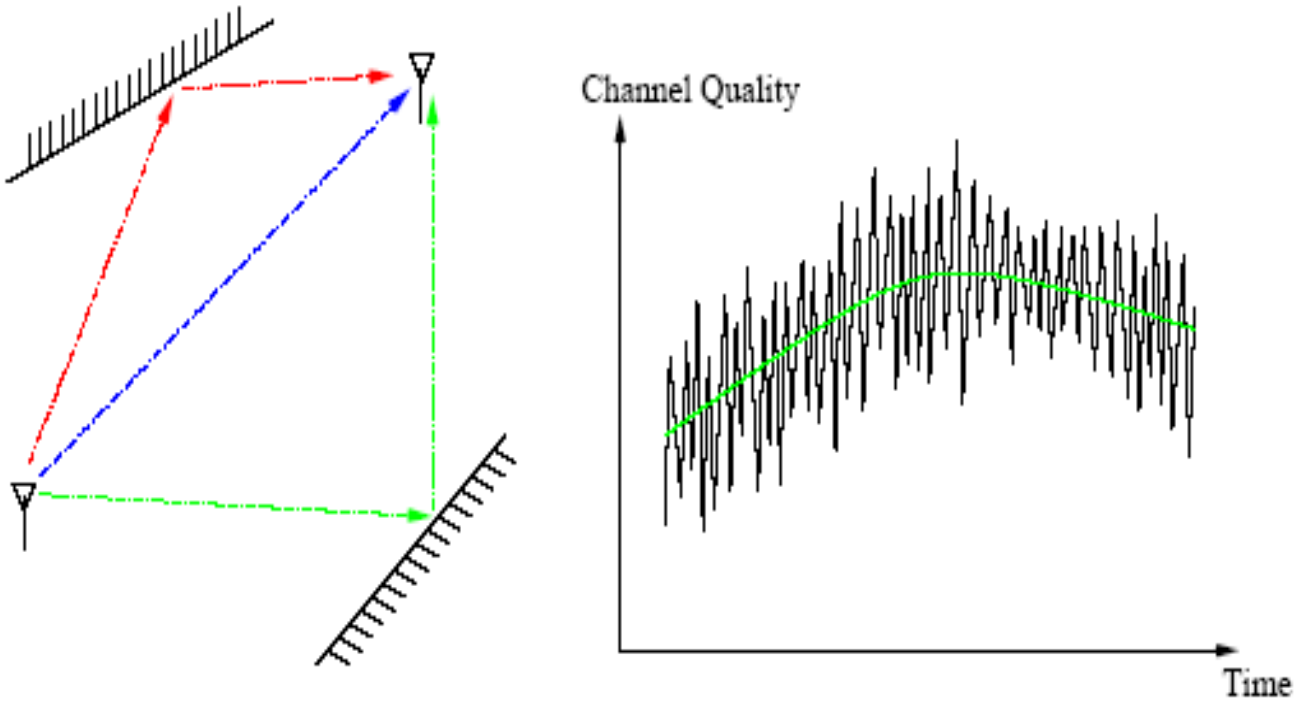
Full details can be found in

D. Tse and P. Viswanath, Fundamentals of Wireless Communication, Cambridge University Press, 2005.

http://www.eecs.berkeley.edu/~dtse

1. The Wireless Channel

Wireless Mulipath Channel



Channel varies at two spatial scales:

large scale fading

small scale fading

Large-scale fading

- In free space, received power attenuates like $1/r^2$.
- With reflections and obstructions, can attenuate even more rapidly with distance. Detailed modelling complicated.
- Time constants associated with variations are very long as the mobile moves, many seconds or minutes.
- More important for cell site planning, less for communication system design.

Small-scale multipath fading

- Wireless communication typically happens at very high carrier frequency. (eg. f_c = 900 MHz or 1.9 GHz for cellular)
- Multipath fading due to constructive and destructive interference of the transmitted waves.
- Channel varies when mobile moves a distance of the order of the carrier wavelength. This is about 0.3 m for 900 Mhz cellular.
- For vehicular speeds, this translates to channel variation of the order of 100 Hz.
- Primary driver behind wireless communication system design.

Game plan

- We wish to understand how physical parameters such as carrier frequency, mobile speed, bandwidth, delay spread impact how a wireless channel behaves from the communication system point of view.
- We start with deterministic physical model and progress towards statistical models, which are more useful for design and performance evaluation.

Physical Models

- Wireless channels can be modeled as linear time-varying systems:

$$y(t) = \sum_i a_i(t) x(t - \tau_i(t))$$

where $a_i(t)$ and $\tau_i(t)$ are the gain and delay of path i.

- The time-varying impulse response is:

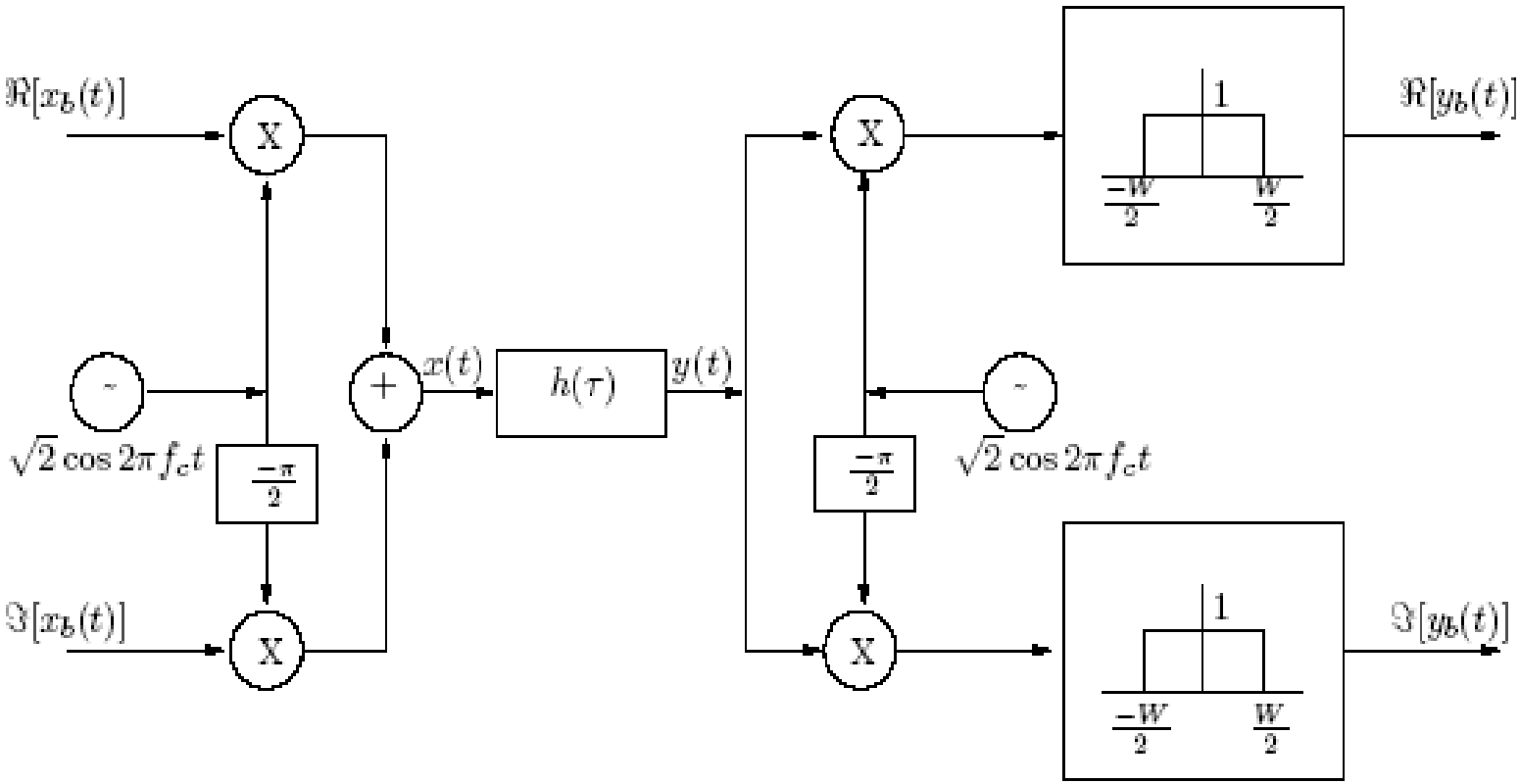
$$h(t, \tau) = \sum_i a_i(t) \delta(\tau - \tau_i(t))$$

- Consider first the special case when the channel is time-invariant:

$$h(\tau) = \sum_i a_i \delta(\tau - \tau_i)$$

Passband to Baseband Conversion

- Communication takes place at [f_c-W/2, f_c+ W/2].
- Processing takes place at baseband [-W/2,W/2].



Baseband Equivalent Channel

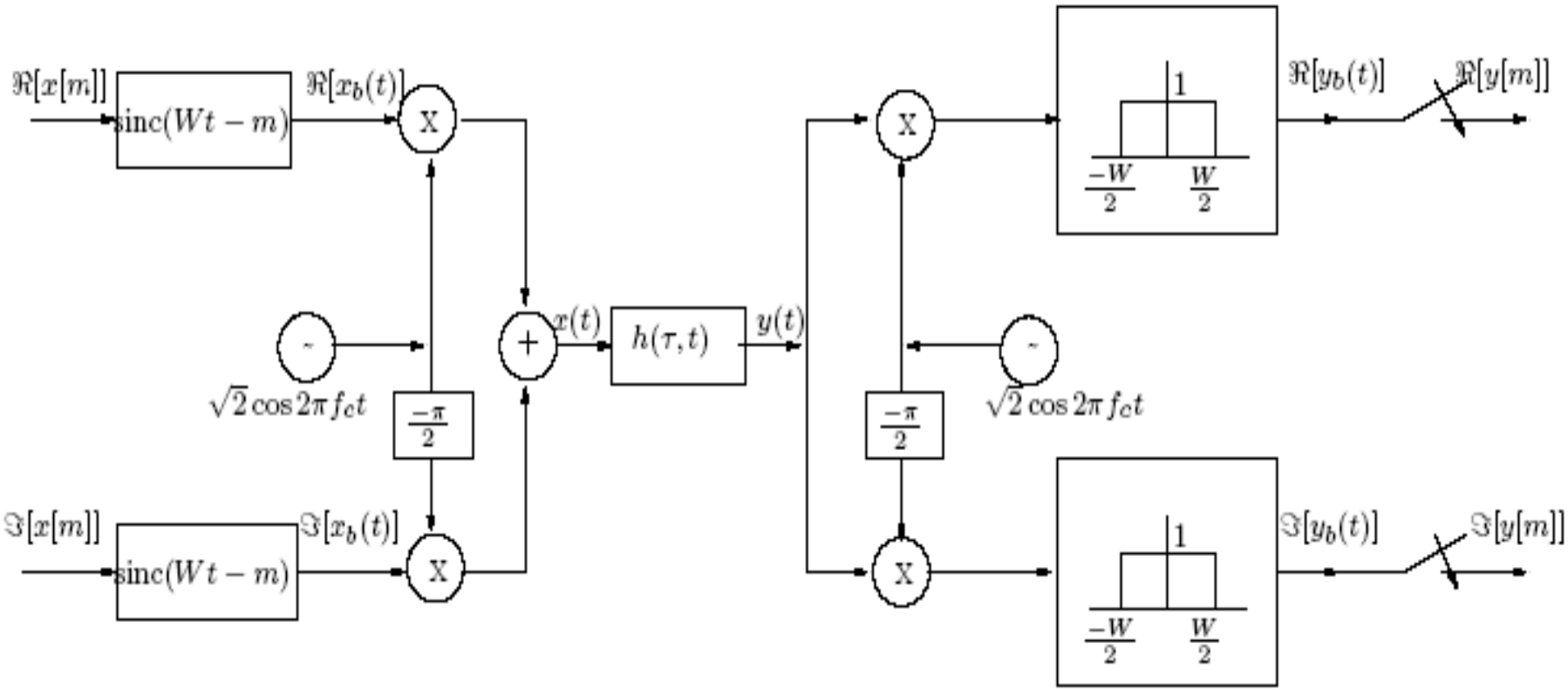
- The frequency response of the system is shifted from the passband to the baseband.

$$H_b(f) = H(f + f_c)$$

$$\begin{aligned} h_b(\tau) &= \sum_i a_i^b \delta(\tau - \tau_i) \\ a_i^b &= a_i e^{-j2\pi f_c \tau_i} \end{aligned}$$

- Each path is associated with a delay and a complex gain.

Modulation and Sampling



Multipath Resolution

Sampled baseband-equivalent channel model:

$$y[m] = \sum_{\ell} h_\ell \, x[m - \ell]$$

where h_l is the l th complex channel tap.

$$h_\ell \approx \sum_{i} a_i e^{j2\pi f_c \tau_i}$$

and the sum is over all paths that fall in the delay bin

$$\left[\frac{\ell}{W} - \frac{1}{2W}, \frac{\ell}{W} + \frac{1}{2W}\right]$$

System resolves the multipaths up to delays of 1/W .

Flat and Frequency-Selective Fading

- Fading occurs when there is destructive interference of the multipaths that contribute to a tap.

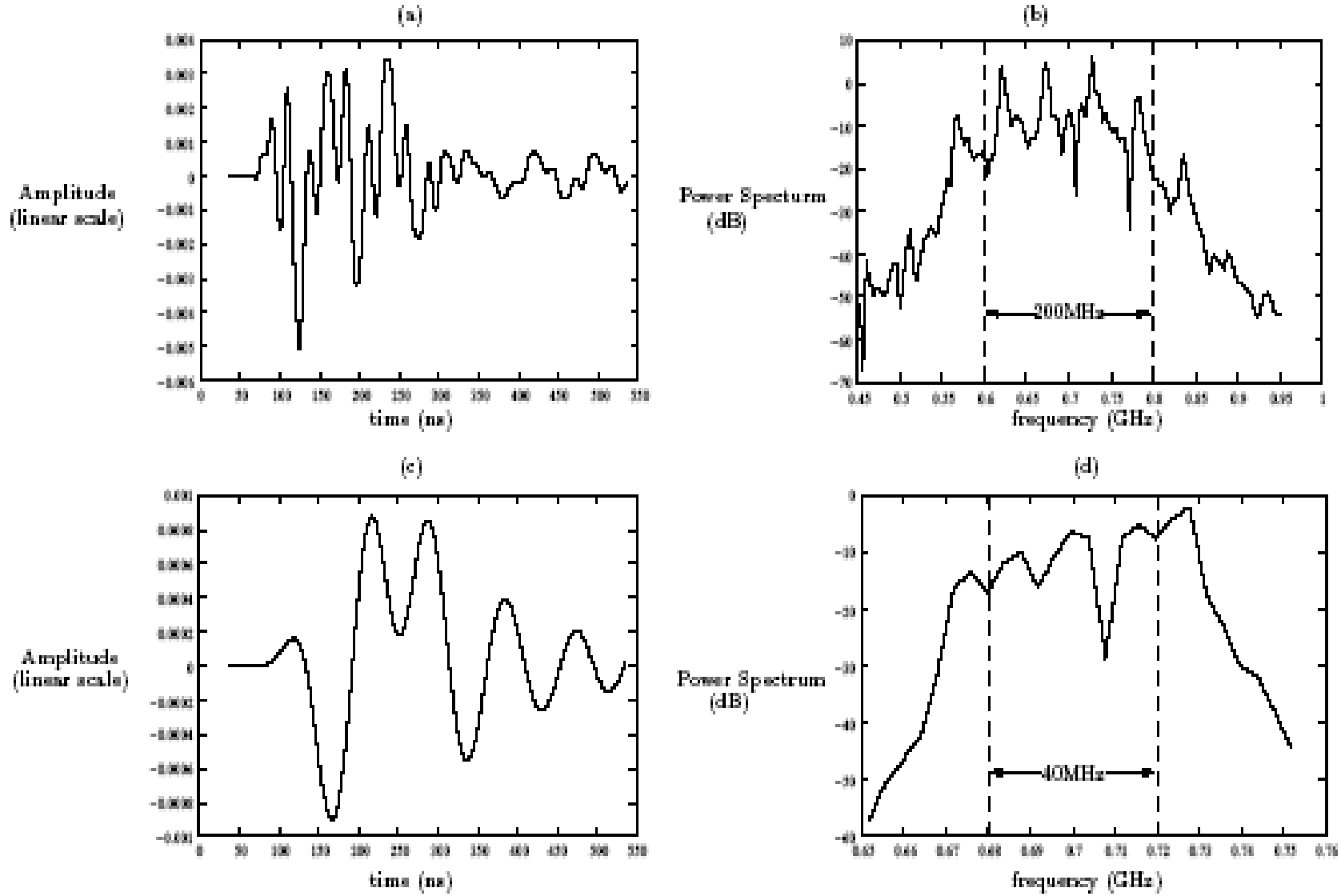
$$h_\ell \approx \sum_i a_i e^{j2\pi f_c \tau_i}$$

$$\text{Delay spread } T_d := \max_{i,j} |\tau_i(t) - \tau_j(t)|$$

$$\text{Coherence bandwidth } W_c := \frac{1}{T_d}$$

$$T_d \ll \frac{1}{W}, W_c \gg W \Rightarrow \text{single tap, flat fading}$$

$$T_d > \frac{1}{W}, W_c \ll W \Rightarrow \text{multiple taps, frequency-selective}$$


(a)
Amplitude
(linear scale)
time (ns)
(b)
Power Spectrum
(dB)
200MHz
frequency (GHz)
(c)
Amplitude
(linear scale)
time (ns)
(d)
Power Spectrum
(dB)
40MHz
frequency (GHz)

Time Variations

$$y[m] = \sum_{\ell} h_\ell[m] x[m-\ell]$$

$$h_\ell[m] \approx \sum_i a_i \left(\frac{m}{W}\right) e^{j2\pi f_c \tau_i(\frac{m}{W})}$$

$f_c\ \tau_i'(t)$ = Doppler shift of the i th path

$$\text{Doppler spread } D_s := \max_{i,j} |f_c\tau_i'(t) - f_c\tau_j'(t)|$$

$$\text{Coherence time } T_c := \frac{1}{D_s}$$

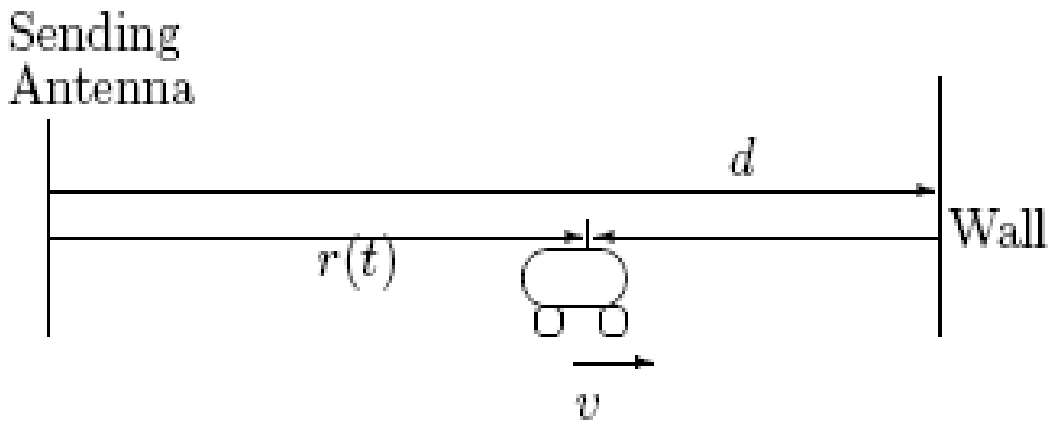
Two-path Example

v= 60 km/hr, f_c = 900 MHz:

direct path has Doppler shift of + 50 Hz

reflected path has shift of - 50 Hz

Doppler spread = 100 Hz



Key Channel Parameters and Time Scales	Symbol	Representative Values
carrier frequency	f_c	1 GHz
communication bandwidth	W	1 MHz
distance between transmitter and receiver	d	1 km
velocity of mobile	v	64 km/h
Doppler shift for a path	$D = \frac{f_c v}{c}$	50 Hz
Doppler spread of paths corresponding to a tap	D_s	100 Hz
time scale for change of path amplitude	$\frac{d}{v}$	1 minute
time scale for change of path phase	$\frac{1}{4D}$	5 ms
time scale for a path to move over a tap	$\frac{c}{vW}$	20 s
coherence time	$T_c = \frac{1}{4D_s}$	2.5 ms
delay spread	T_d	1 μ s
coherence bandwidth	$W_c = \frac{1}{2\pi T_d}$	500 kHz

Types of Channels

Types of Channel	Defining Characteristic
fast fading	$T_c \ll$ delay requirement
slow fading	$T_c \gg$ delay requirement
flat fading	$W \ll W_c$
frequency-selective fading	$W \gg W_c$
underspread	$T_d \ll T_c$

Statistical Models

- Design and performance analysis based on statistical ensemble of channels rather than specific physical channel.

$$h_\ell[m] \approx \sum_i a_i(m/W) e^{j2\pi f_c \tau_i (m/W)}$$

- Rayleigh flat fading model: many small scattered paths

$$h[m] \sim \mathcal{N}(0, \frac{1}{2}) + j\mathcal{N}(0, \frac{1}{2}) \sim \mathcal{CN}(0, 1)$$

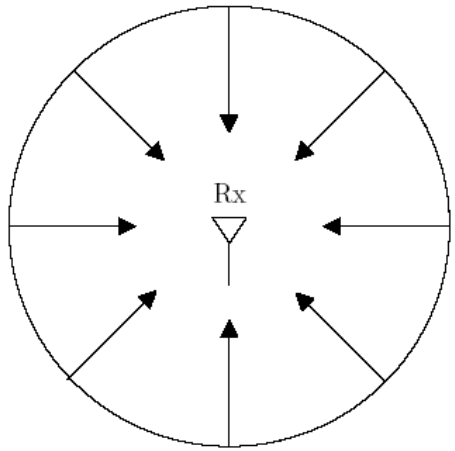
Complex circular symmetric Gaussian .

- Rician model: 1 line-of-sight plus scattered paths

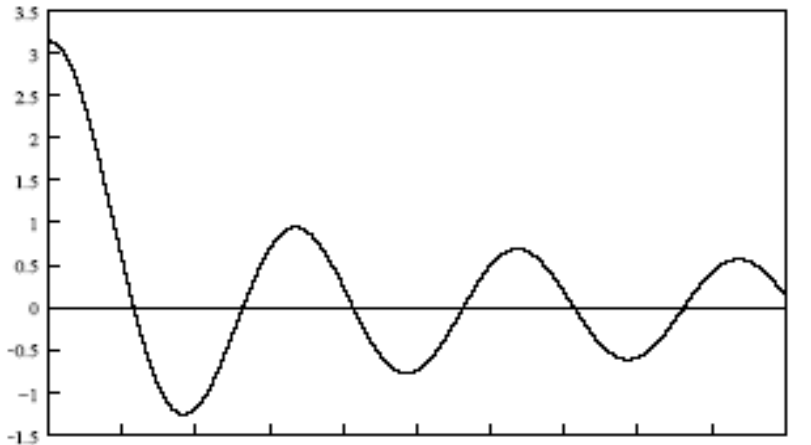
$$h[m] \sim \sqrt{\kappa} + \mathcal{CN}(0, 1)$$

Correlation over Time

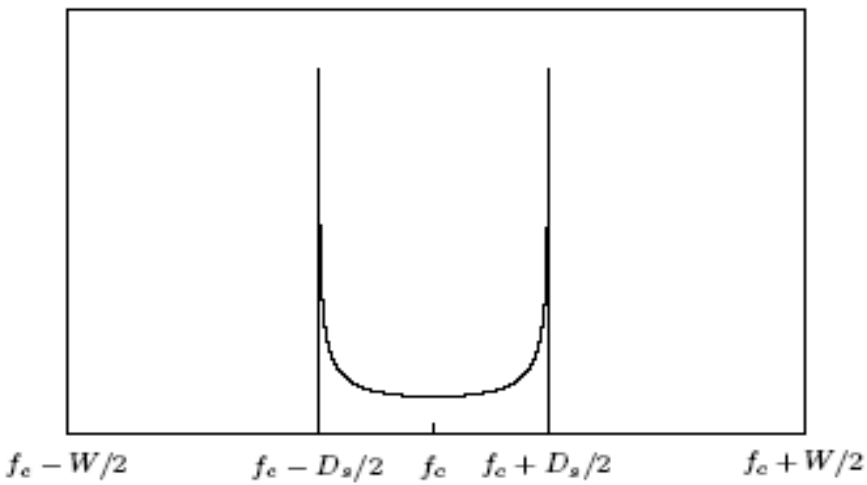
- Specified by autocorrelation function and power spectral density of fading process.
- Example: Clarke's (or Jake's) model.



$R_0[n]$



$S(f)$



Additive Gaussian Noise

- Complete baseband-equivalent channel model:

$$y[m] = \sum_{\ell} h_{\ell}[m]x[m-\ell] + w[m]$$

$$w[m] \sim \mathcal{CN}(0, N_0)$$

- Special case: flat fading:

$$y[m] = h[m]x[m] + w[m]$$

- Will use this throughout the course.

2. Diversity

Main story

- Communication over a flat fading channel has poor performance due to significant probability that channel is in deep fading.
- Reliability is increased by providing more signal paths that fade independently.
- Diversity can be provided across time, frequency and space.
- Name of the game is how to exploit the added diversity in an efficient manner.

Baseline: AWGN Channel

y = x+ w

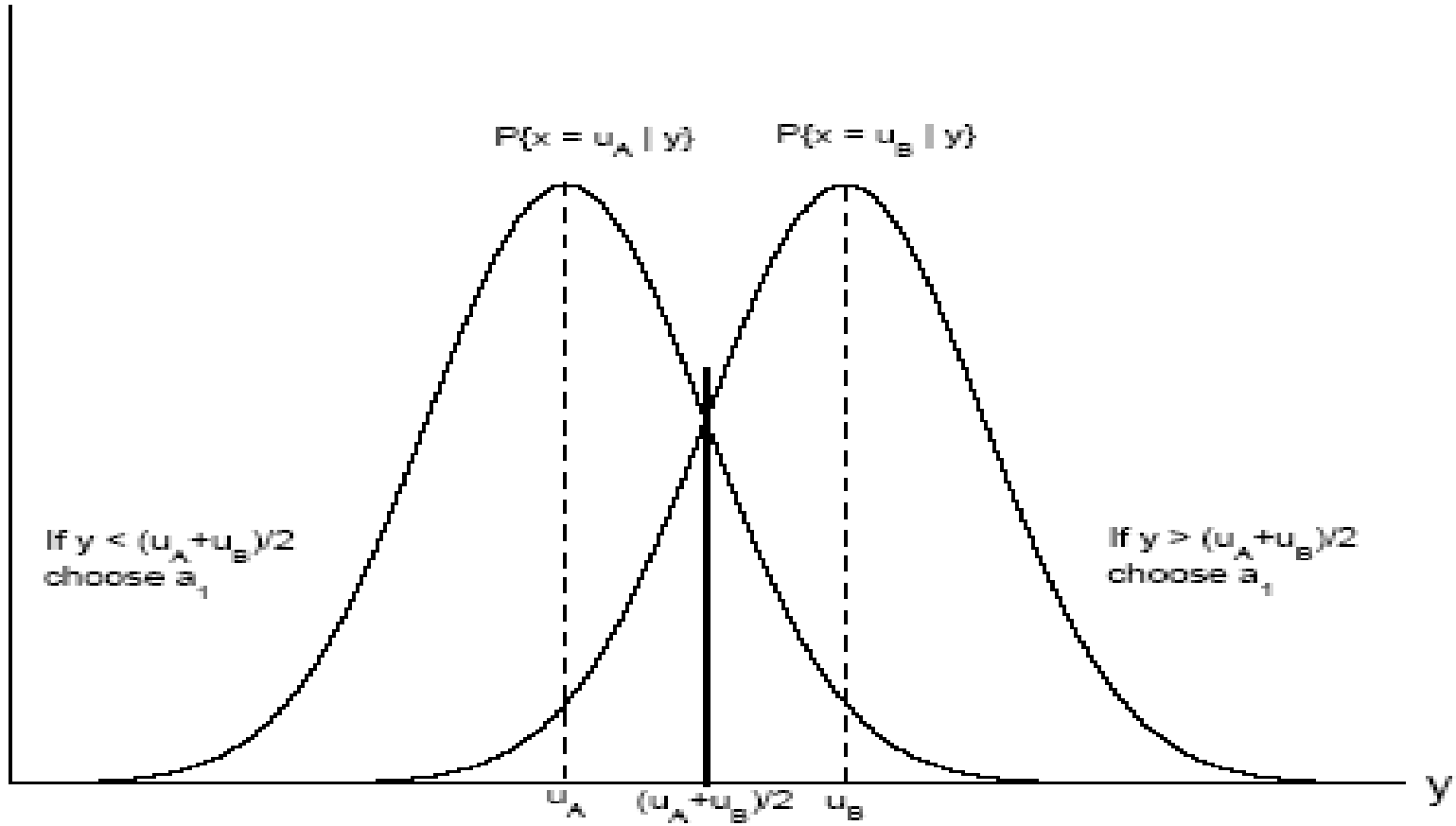
BPSK modulation x = $\pm$ a

$$p_e = Q\left(\frac{a}{\sqrt{N_0/2}}\right) = Q\left(\sqrt{2\mathsf{SNR}}\right)$$

$$\mathsf{SNR} := \frac{a^2}{N_0}$$

Error probability decays exponentially with SNR.

Gaussian Detection



Rayleigh Flat Fading Channel

$$y = hx + w$$

$$h \sim \mathcal{CN}(0, 1)$$

BPSK: $x = \pm a$.

Conditional on h,

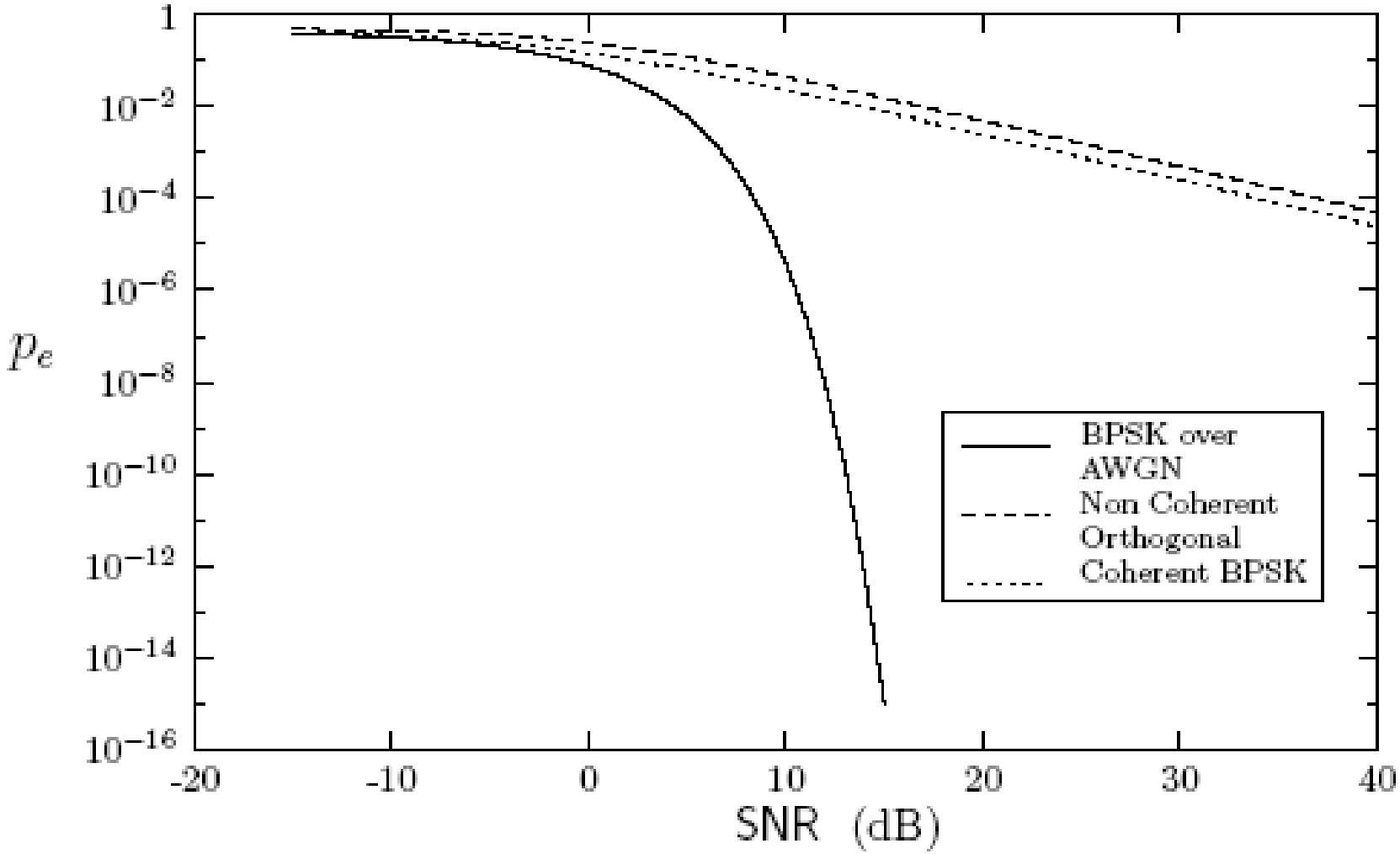
$$p_e = Q\left(\sqrt{2|h|^2 \mathsf{SNR}}\right)$$

Averaged over h,

$$p_e = \frac{1}{2}\left(1 - \sqrt{\frac{\mathsf{SNR}}{1 + \mathsf{SNR}}}\right) \approx \frac{1}{4\mathsf{SNR}}$$

at high SNR.

Rayleigh vs AWGN



Typical Error Event

Conditional on h,

$$p_e = Q\left(\sqrt{2|h|^2 \mathsf{SNR}}\right)$$

When $|h|^2 >> \frac{1}{\mathsf{SNR}}$, error probability is very small.

When $|h|^2 < \frac{1}{\mathsf{SNR}}$, error probability is very significant.

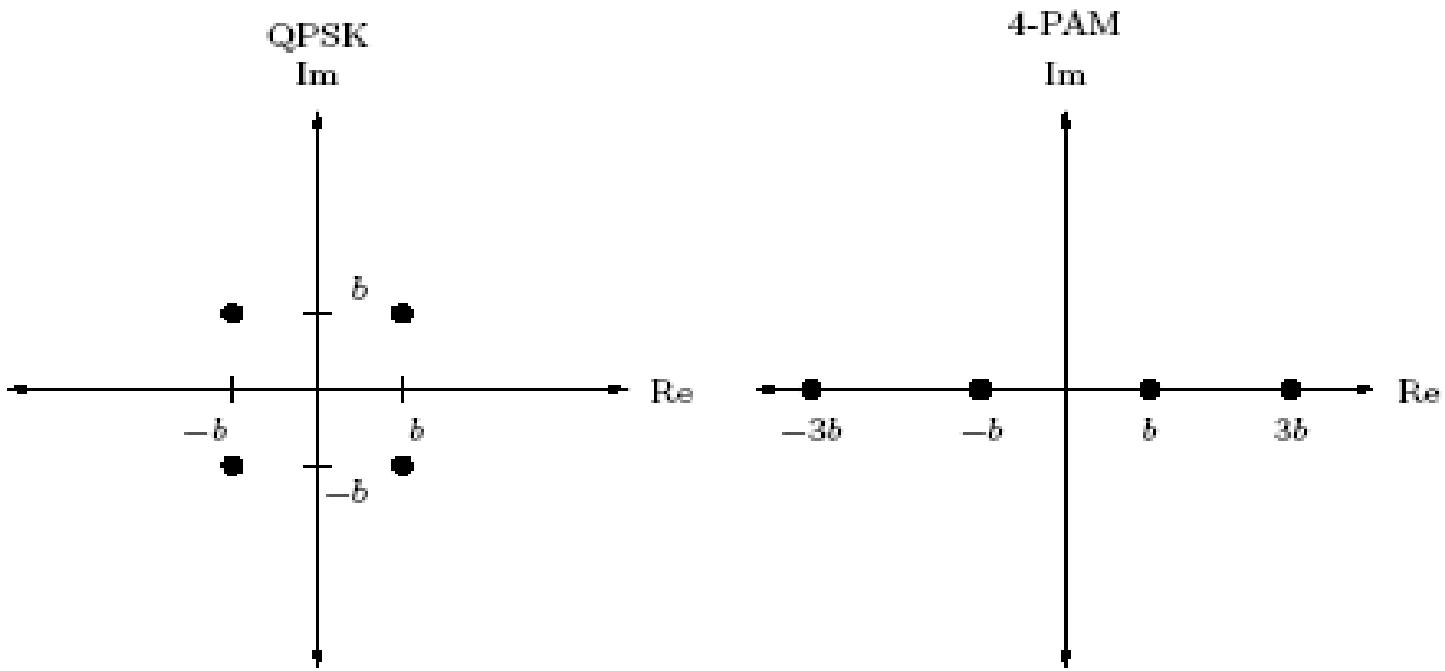
$$p_e \approx P\left(|h|^2 < \frac{1}{\mathsf{SNR}}\right) \approx \frac{1}{\mathsf{SNR}}$$

$$|h|^2 \sim \exp(1).$$

Typical error event is due to channel being in deep fade rather than noise being large.

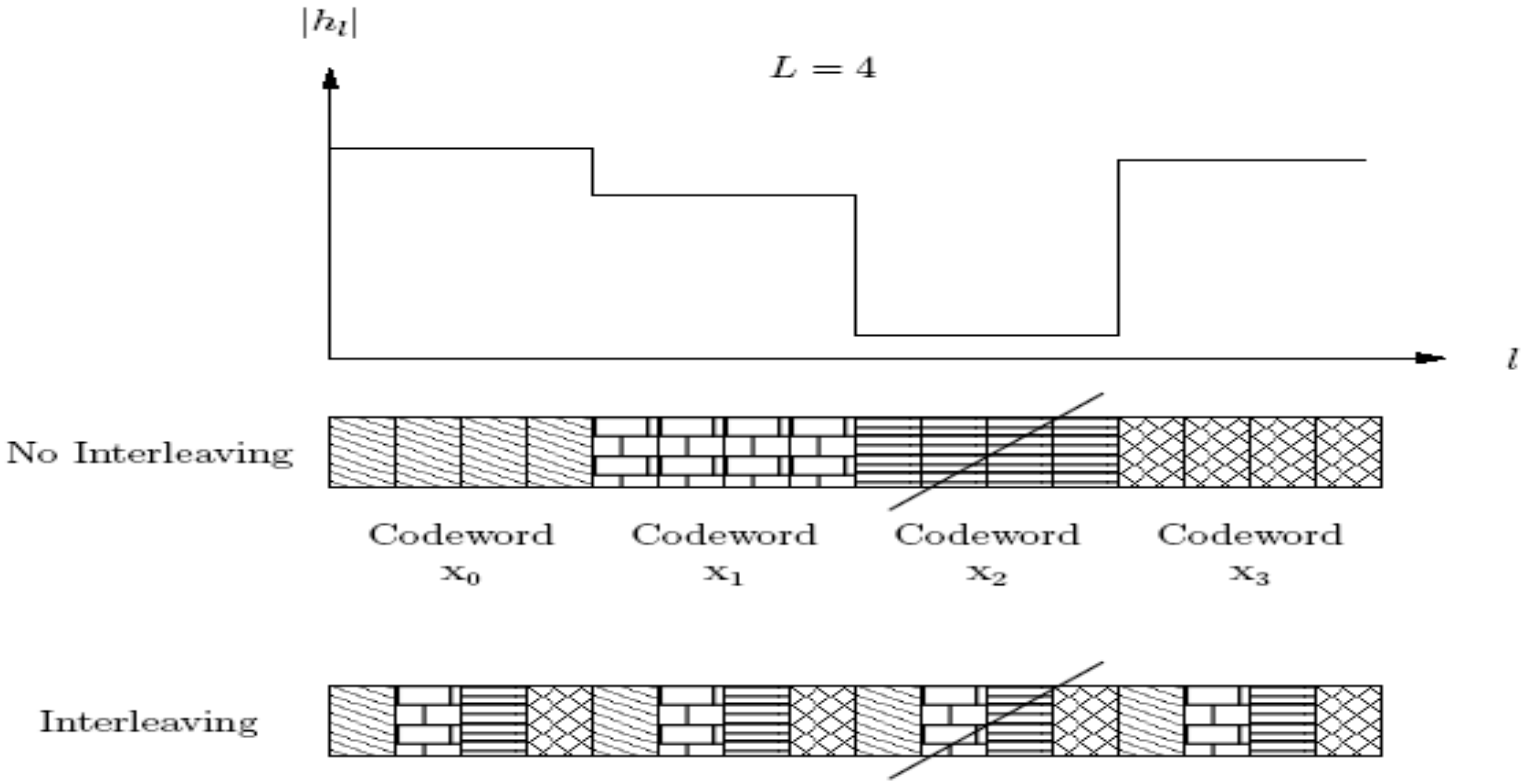
BPSK, QPSK and 4-PAM

- BPSK uses only the I-phase.The Q-phase is wasted.
- QPSK delivers 2 bits per complex symbol.
- To deliver the same 2 bits, 4-PAM requires 4 dB more transmit power.
- QPSK exploits the available degrees of freedom in the channel better.

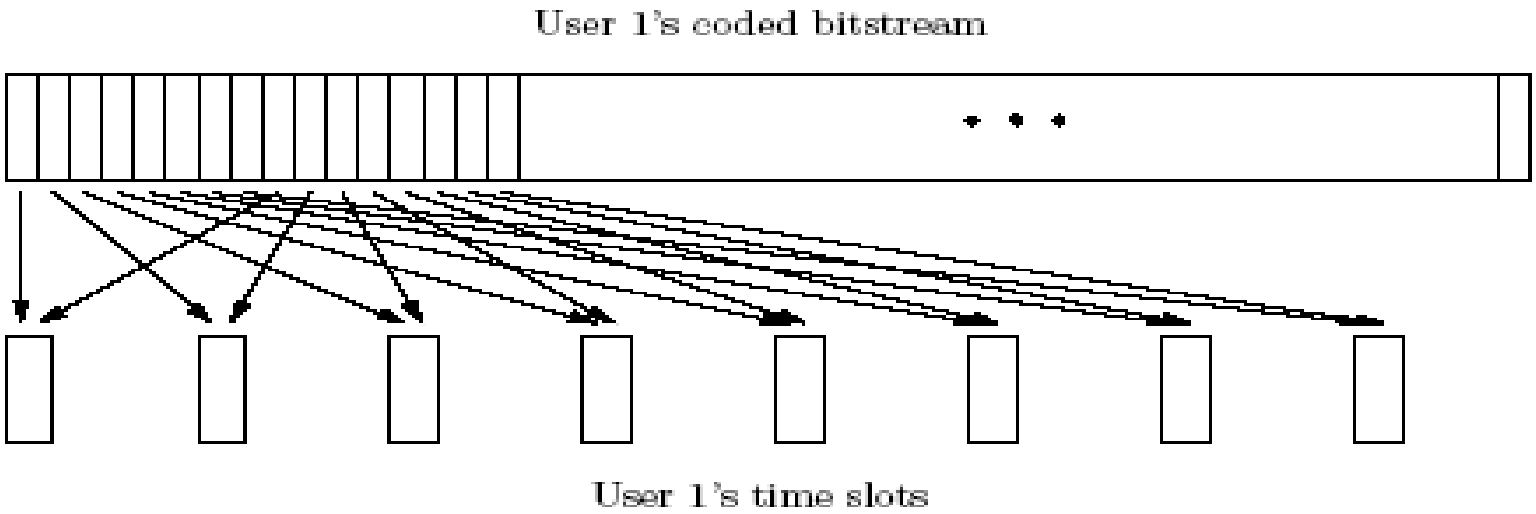


Time Diversity

- Time diversity can be obtained by interleaving and coding over symbols across different coherent time periods.



Example:GSM



- Amount of diversity limited by delay constraint and how fast channel varies.
- In GSM, delay constraint is 40ms (voice).
- To get full diversity of 8, needs v > 30 km/hr at f_c = 900Mhz.

Repetition Coding

After interleaving over L coherence time periods,

$$y_\ell = h_\ell x_\ell + w_\ell, \qquad \ell = 1, \ldots, L$$

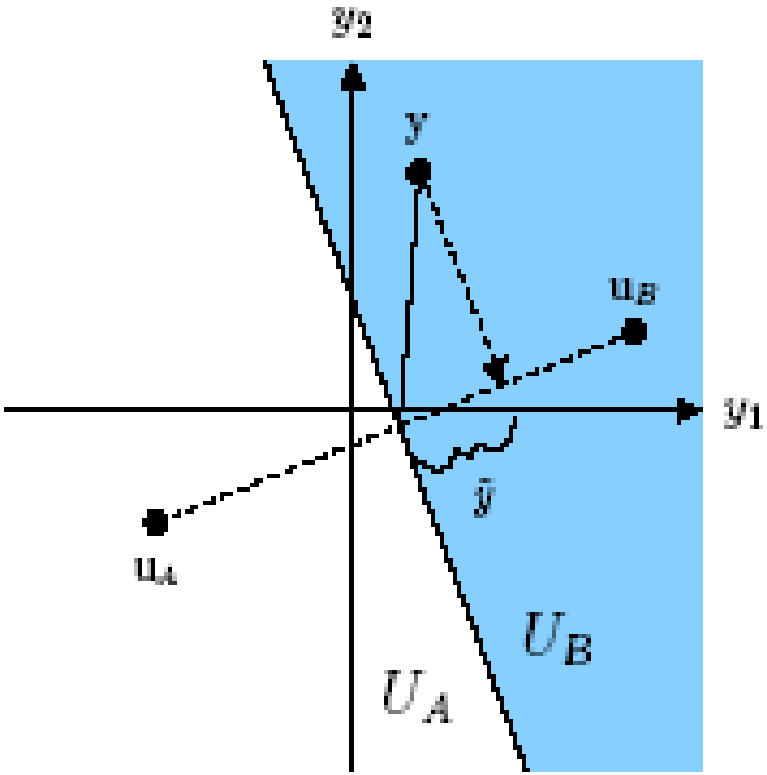
Repetition coding: $x_\ell = x$ for all ℓ.

$$\mathbf{y} = \mathbf{h}x + \mathbf{w}$$

where $\mathbf{y} = [y_1, \ldots, y_L]^t$, $\mathbf{h} = [h_1, \ldots, h_L]^t$ and $\mathbf{w} = [w_1, \ldots, w_L]^t$.

This is classic vector detection in white Gaussian noise.

Geometry



For BPSK $x = \pm a$,

$$\mathbf{u}_A = +a\mathbf{h}, \mathbf{u}_B = -a\mathbf{h}.$$

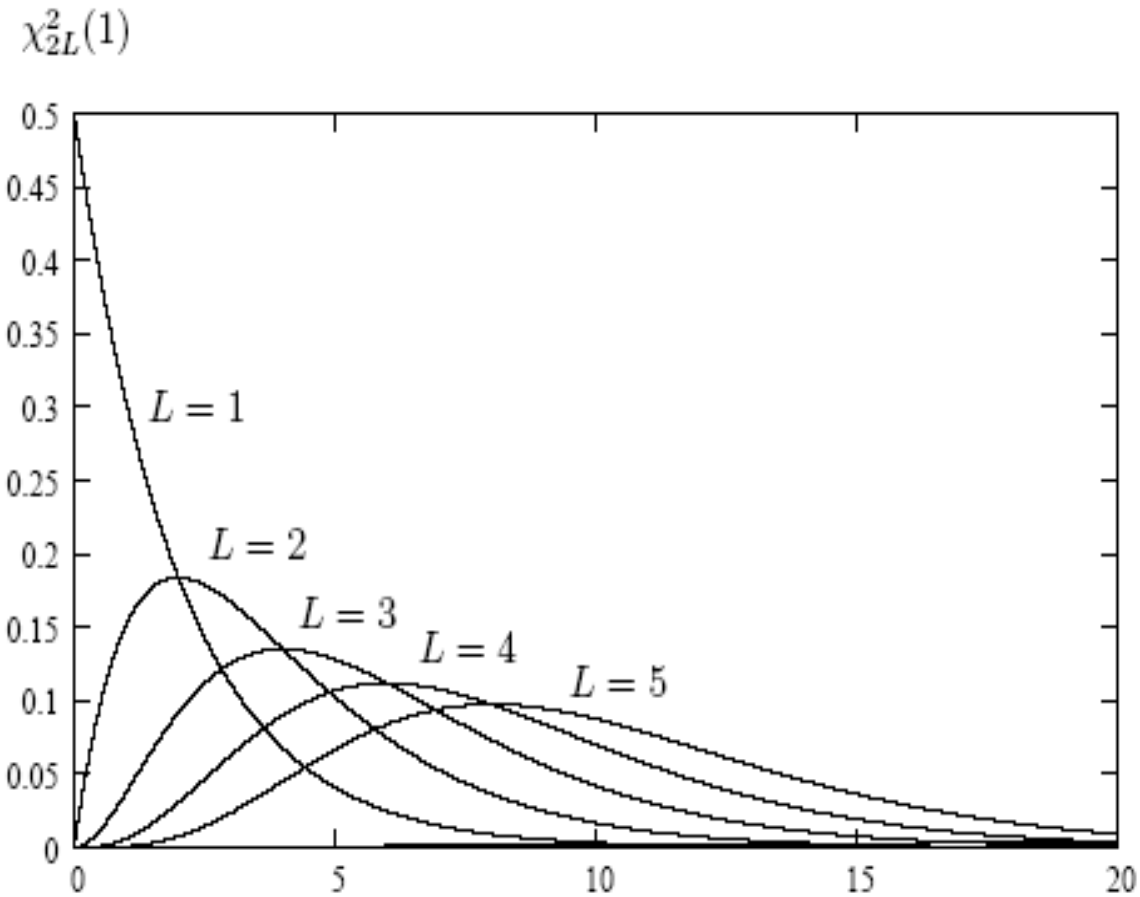
$$\tilde{y} = \frac{\mathbf{h}^*}{\|\mathbf{h}\|}\mathbf{y}$$

is a sufficient statistic (match filtering.)

Reduces to scalar detection problem:

$$\tilde{y} = \|\mathbf{h}\|x + \tilde{w}$$

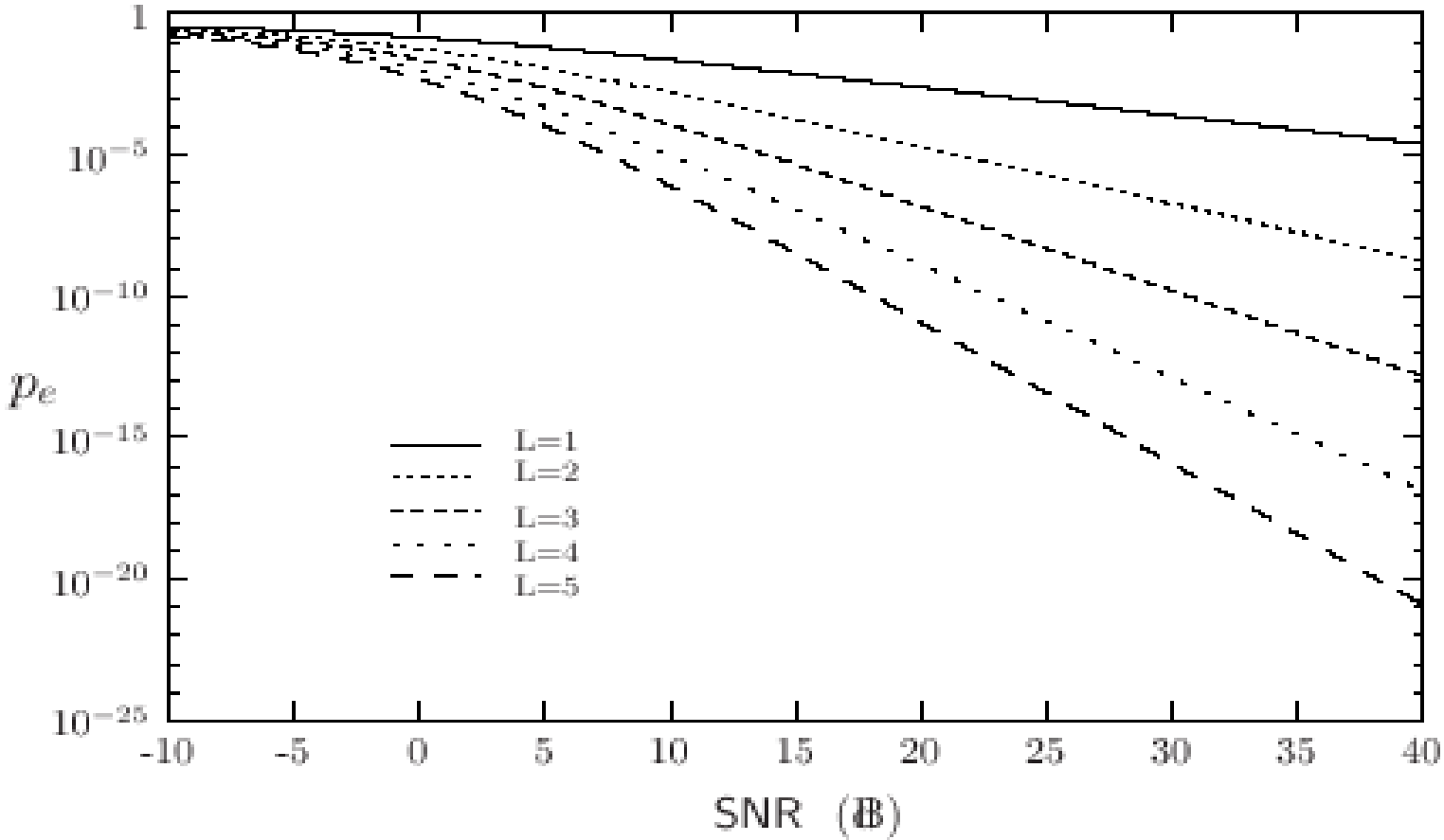
Deep Fades Become Rarer



$$P\left(\|\mathbf{h}\|^2 < \epsilon\right) \approx \frac{1}{L!}\epsilon^L$$

$$\begin{aligned} p_e &\approx P\left(\|\mathbf{h}\|^2 < \frac{1}{\mathsf{SNR}}\right) \\ &\approx \frac{1}{L!}\frac{1}{\mathsf{SNR}^L} \end{aligned}$$

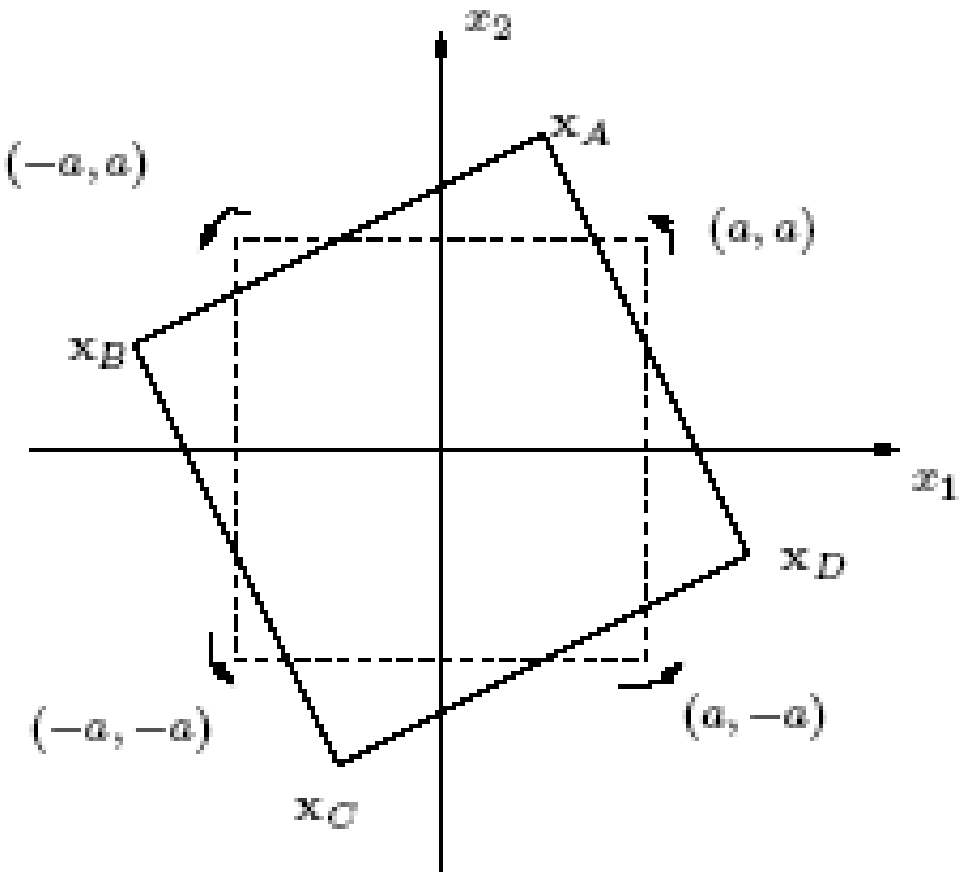
Performance



Beyond Repetition Coding

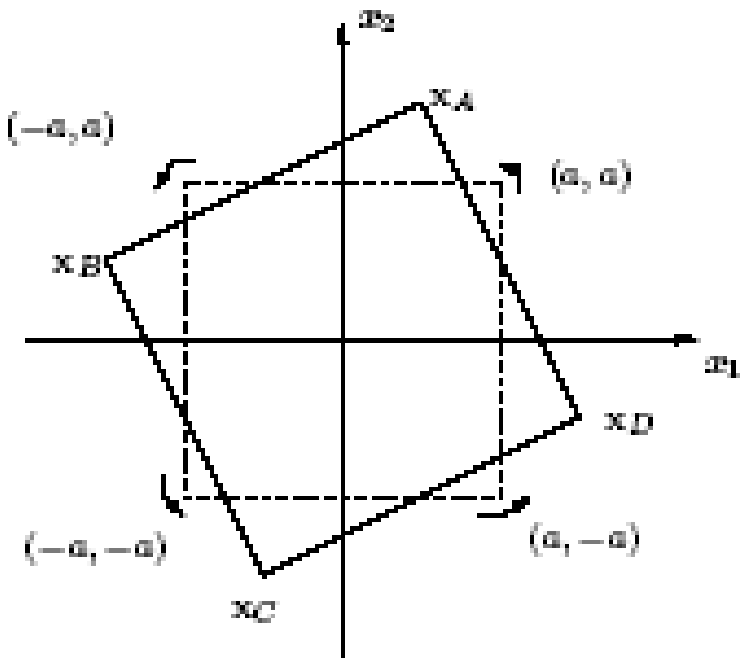
- Repetition coding gets full diversity, but sends only one symbol every L symbol times: does not exploit fully the degrees of freedom in the channel.
- How to do better?

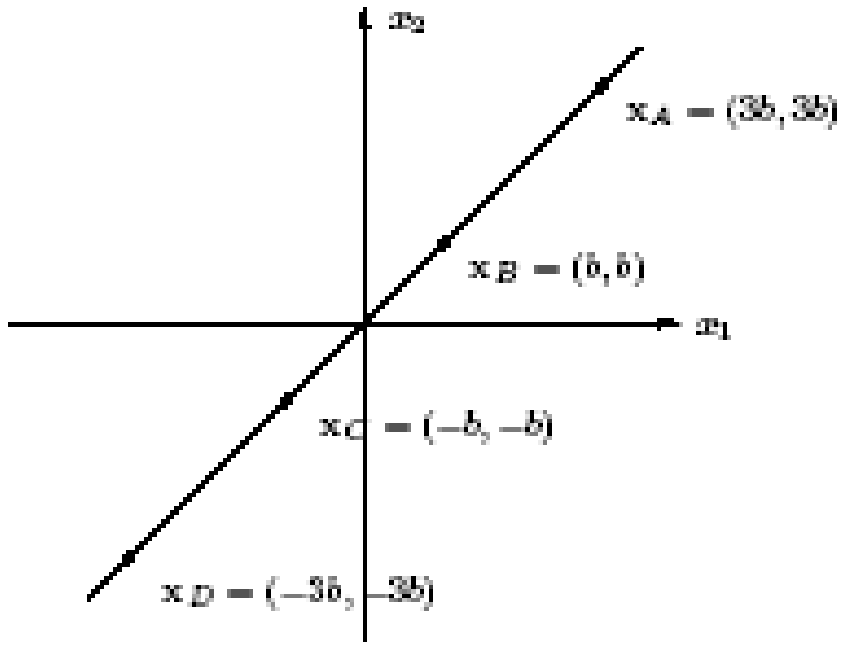
Example: Rotation code (L=2)



$$\mathbb{P}\{\mathbf{x}_A \to \mathbf{x}_B | h_1, h_2\} = Q\left(\frac{\|\mathbf{u}_A - \mathbf{u}_B\|}{2\sqrt{N_0/2}}\right) = Q\left(\sqrt{\mathrm{SNR}/2 \cdot (|h_1|^2|d_1|^2 + |h_2|^2|d_2|^2)}\right)$$

Rotation vs Repetition Coding





Product Distance

$$\mathcal{P}\left\{\mathbf{x}_A \to \mathbf{x}_B | h_1, h_2\right\} = Q\left(\frac{\mathsf{SNR}}{2}\left[|d_1|^2|h_1|^2 + |d_2|^2|h_2|\right]\right)$$

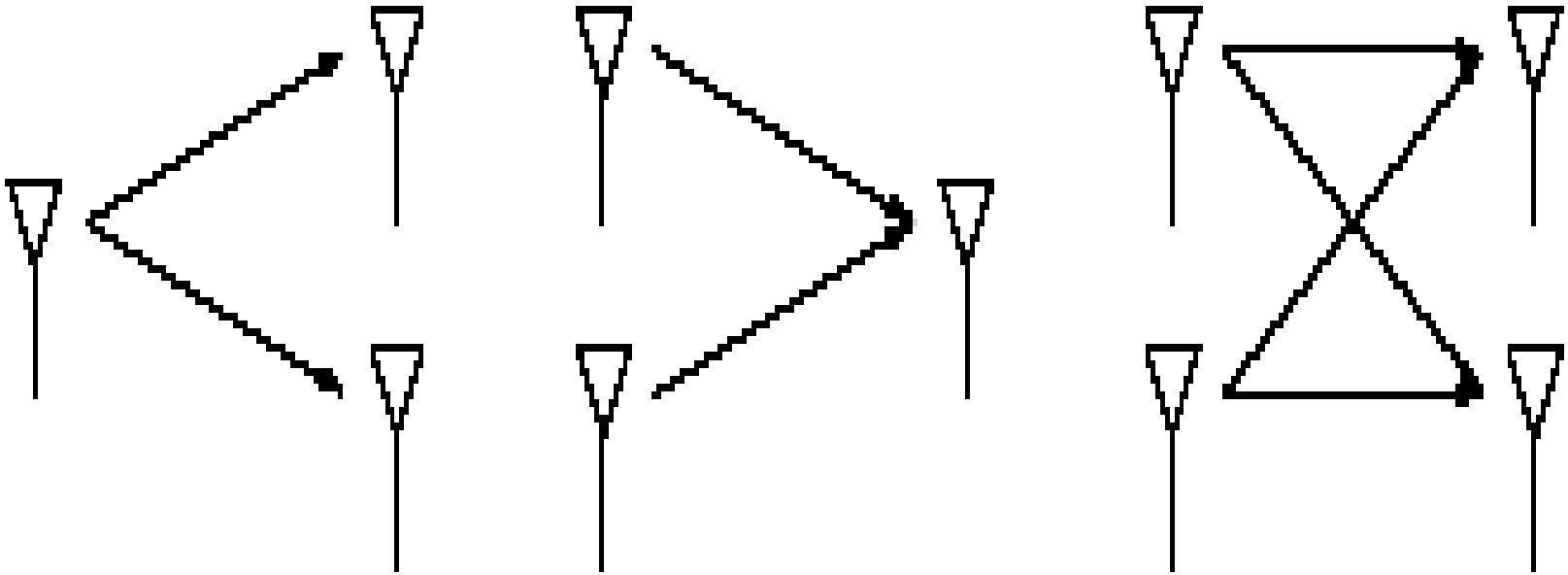
$$\begin{aligned} & \mathcal{P}\left\{\mathbf{x}_A \to \mathbf{x}_B\right\} \\ \approx\ & \mathcal{P}\left\{|d_1|^2|h_1|^2 < \frac{1}{\mathsf{SNR}} \quad \& \quad |d_2|^2|h_2|^2 < \frac{1}{\mathsf{SNR}}\right\} \\ \approx\ & \frac{1}{|d_1|^2|d_2|^2}\mathsf{SNR}^{-2} \end{aligned}$$

$$\text{product distance } = |d_1||d_2|.$$

Choose the rotation angle to maximize the worst-case product distance to all the other codewords:

$$\theta^* = \frac{1}{2}\tan^{-1} 2.$$

Antenna Diversity



Receive

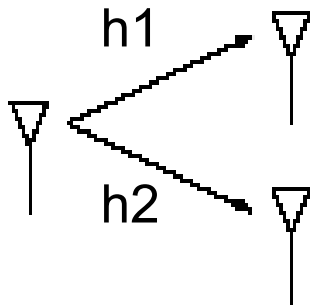
Transmit

Both

Receive Diversity

$$\mathbf{y} = x\mathbf{h} + \mathbf{w}$$

Same as repetition coding in time diversity, except that there is a further power gain.



Optimal reception is via match filtering (receive beamforming).

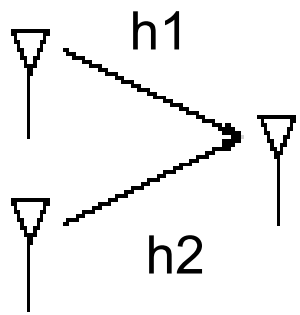
Transmit Diversity

$$y = \mathbf{h}^* \mathbf{x} + w$$

If transmitter knows the channel, send:

$$\mathbf{x} = x \frac{\mathbf{h}}{\|\mathbf{h}\|}.$$

maximizes the received SNR by in-phase addition of signals at the receiver (transmit beamforming).



Reduce to scalar channel:

$$y = \|\mathbf{h}\| x + w,$$

same as receive beamforming.

What happens if transmitter does not know the channel?

Space-time Codes

- Transmitting the same symbol simultaneously at the antennas doesn't work.
- Using the antennas one at a time and sending the same symbol over the different antennas is like repetition coding.
- More generally, can use any time-diversity code by turning on one antenna at a time.

Alamouti Scheme

Over two symbol times:

$$\begin{bmatrix} y[1] & y[2] \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \end{bmatrix} \begin{bmatrix} u_1 & -u_2^* \\ u_2 & u_1^* \end{bmatrix} + \begin{bmatrix} w[1] & w[2] \end{bmatrix}.$$

$$\begin{bmatrix} y[1] \\ y[2]^* \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} w[1] \\ w[2]^* \end{bmatrix}$$

Projecting onto the two columns of the H matrix yields:

$$r_i = \|\mathbf{h}\| u_i + w_i, \qquad i = 1, 2.$$

- double the symbol rate of repetition coding.
- 3dB loss of received SNR compared to transmit beamforming.

Space-time Code Design

A space-time code is a set of matrices $\{\mathbf{X}_i\}$.

Full diversity is achieved if all pairwise differences have full rank.

Coding gain determined by the determinants of $(\mathbf{X}_i - \mathbf{X}_j)(\mathbf{X}_i - \mathbf{X}_i)^*$.

Time-diversity codes have diagonal matrices and the determinant reduces to squared product distances.

Cooperative Diversity

- Different users can form a distributed antenna array to help each other in increasing diversity.
- Distributed versions of space-time codes may be applicable.
- Interesting characteristics:
  - Users have to exchange information and this consumes bandwidth.
  - Operation typically in half-duplex mode
  - Broadcast nature of the wireless medium can be exploited.

Frequency Diversity

$$y[m] = \sum_{\ell=0}^{L-1} h_\ell x[m-\ell + w[m]$$

Resolution of multipaths provides diversity.

Full diversity is achieved by sending one symbol every L symbol times.

But this is inefficient (like repetition coding).

Sending symbols more frequently may result in intersymbol interference.

Challenge is how to mitigate the ISI while extracting the inherent diversity in the frequencyp-selective channel.

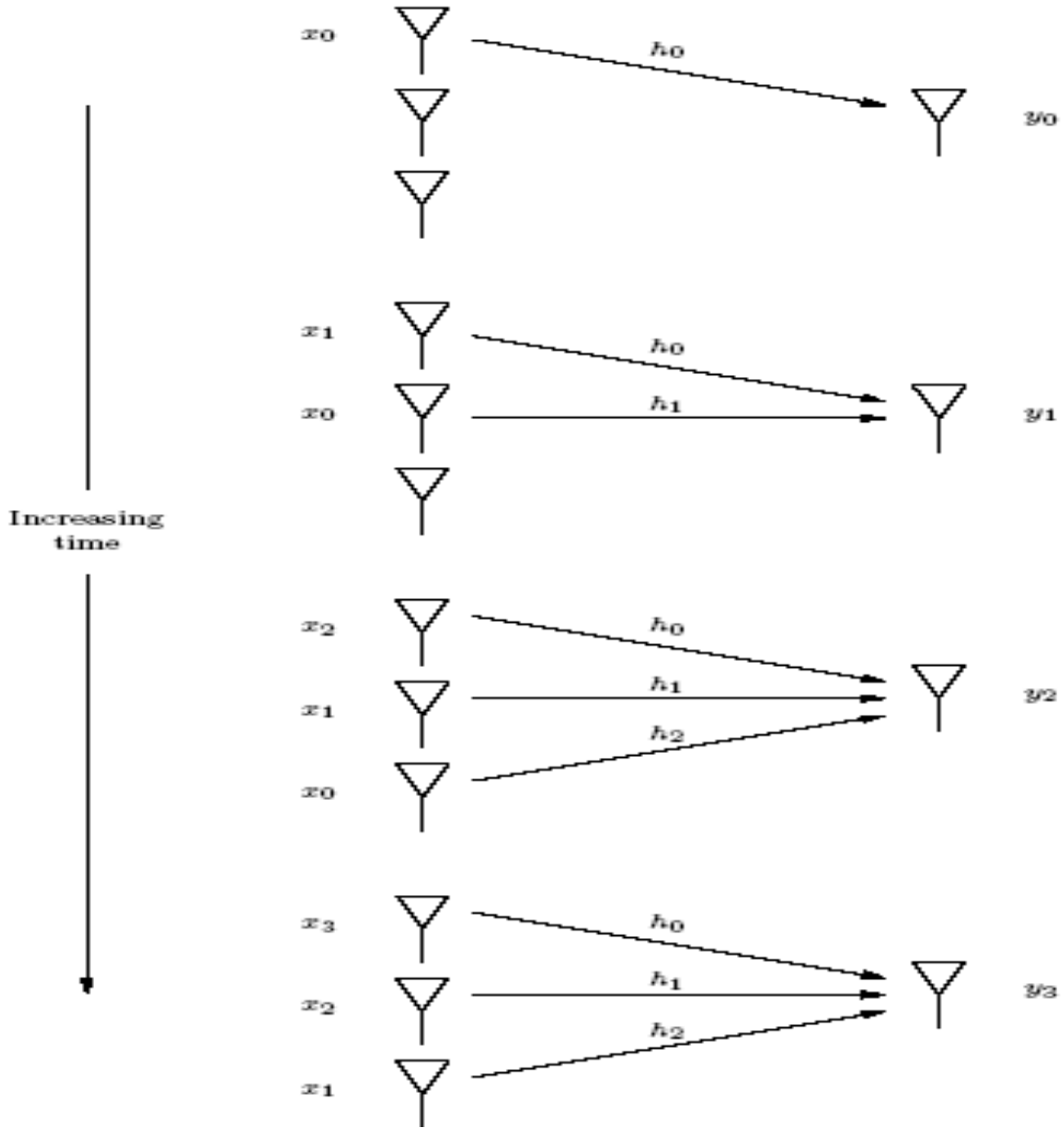
Approaches

- Time-domain equalization (eg. GSM)
- Direct-sequence spread spectrum (eg. IS-95 CDMA)
- Orthogonal frequency-division multiplexing OFDM (eg. 802.11a)

ISI Equalization

- Suppose a sequence of uncoded symbols are transmitted.
- Maximum likelihood sequence detection is performed using the Viterbi algorithm.
- Can full diversity be achieved?

Reduction to Transmit Diversity



MLSD Achieves Full Diversity

Space-time code matrix for input sequence $x[0], \ldots, x[N+L-1]$:

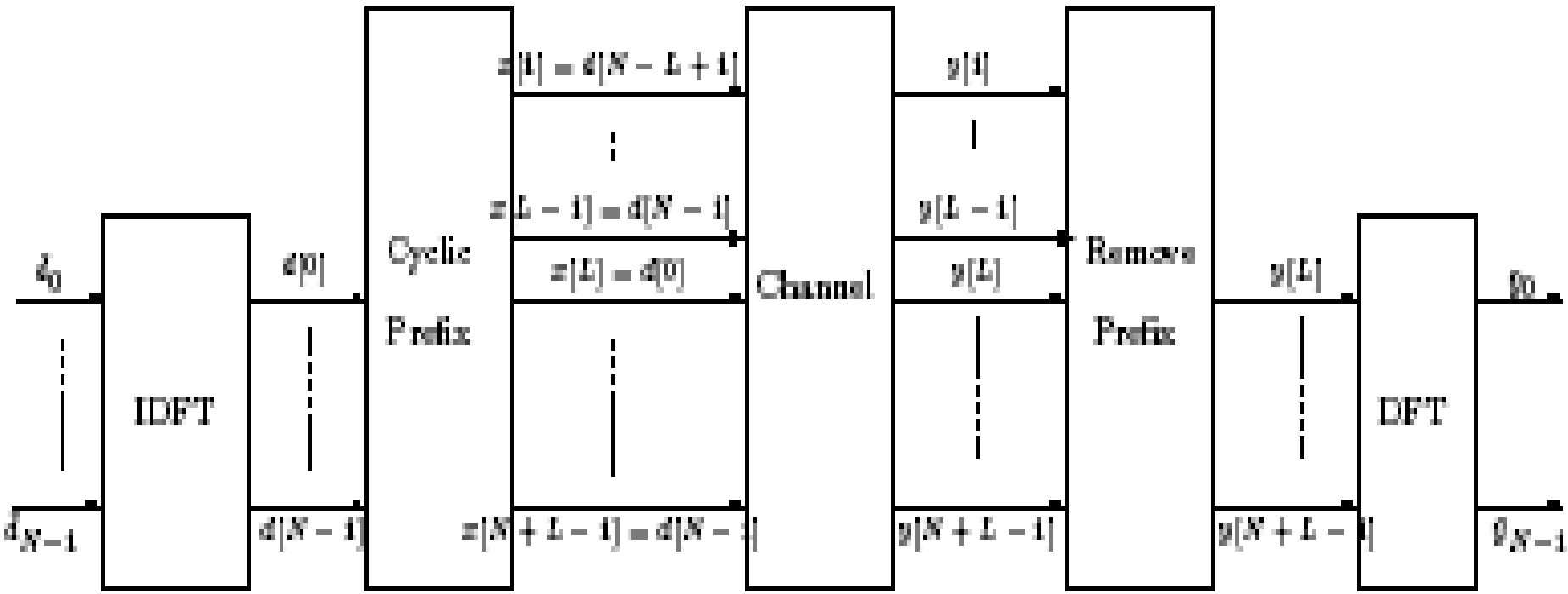
$$\mathbf{X} = \begin{bmatrix} x[1] & x[2] & \cdot & \cdot & \cdot & x[N] & \cdot & \cdot & x[N+L-1] \\ 0 & x[1] & x[2] & \cdot & \cdot & \cdot & x[N] & \cdot & x[N+L-2] \\ 0 & 0 & x[1] & x[2] & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & x[1] & x[2] & \cdot & \cdot & x[N] \end{bmatrix}$$

Difference matrix for two sequences first differing at $m^* \leq N$:

$$\mathbf{X}_A - \mathbf{X}_B = \begin{bmatrix} 0 & \cdot & 0 & x_A[m^*] - x_B[m^*] & & \cdot & & \cdot & & \cdot \\ 0 & \cdot & \cdot & & 0 & & x_A[m^*] - x_B[m^*] & \cdot & & \cdot \\ 0 & \cdot & \cdot & & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot & & \cdot & \cdot & & \cdot \\ 0 & \cdot & \cdot & & \cdot & & \cdot & 0 & x_A[m^*] - x_B[m^*] & \cdot \end{bmatrix}$$

is full rank.

OFDM



OFDM

OFDM transforms the communication problem into the frequency domain:

$$\tilde{y}_n = \tilde{h}_n \tilde{d}_n + \tilde{w}_n, \quad n = 0, \ldots, N_c - 1.$$

a bunch of non-interfering sub-channels, one for each sub-carrier.

$$\tilde{h}_n = H_b\left(\frac{nW}{N_c}\right)$$

Can apply time-diversity techniques.

Channel Uncertainty

- In fast varying channels, tap gain measurement errors may have an impact on diversity combining performance
- The impact is particularly significant in channel with many taps each containing a small fraction of the total received energy. (eg. Ultra-wideband channels)

3. Capacity of Wireless Channels

Information Theory

- So far we have only looked at uncoded or simple coding schemes.
- Information theory provides a fundamental characterization of coded performance.
- It succintly identifies the impact of channel resources on performance as well as suggests new and cool ways to communicate over the wireless channel.
- It provides the basis for the modern development of wireless communication.

Capacity of AWGN Channel

Capacity of AWGN channel

$$\begin{aligned} C_{\text{awgn}} &= \log(1+\text{SNR}) \quad \text{bits/s/Hz} \\ &= W\log(1+\text{SNR}) \quad \text{bits/s} \end{aligned}$$

If average transmit power constraint is $\bar{P}$ watts and noise psd is N_0 watts/Hz,

$$C_{\text{awgn}} = W\log\left(1+\frac{\bar{P}}{N_0 W}\right) \quad \text{bits/s.}$$

Power and Bandwidth Limited Regimes

$$C_{\text{awgn}} = W \log(1 + \text{SNR}) = W \log\left(1 + \frac{\bar{P}}{N_0 W}\right)$$

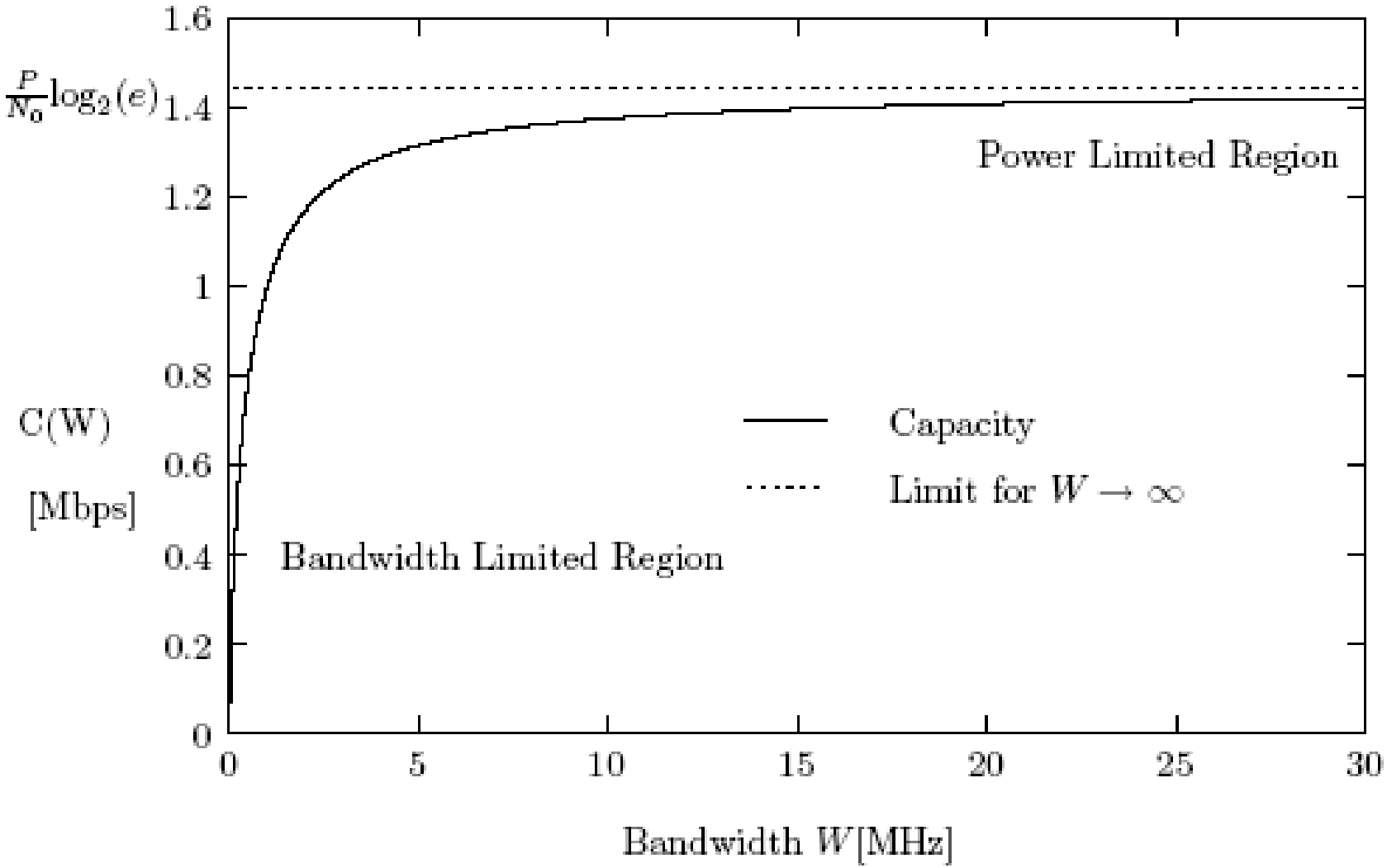
Bandwidth limited regime SNR $\gg 1$: capacity logarithmic in power.

Power limited regime SNR $\ll 1$: capacity linear in power.

As $W \to \infty$:

$$C_{\text{awgn}} \to \frac{\bar{P}}{N_0} \log_2 e \quad \text{bits/s}$$

Equivalently, minimum $\mathcal{E}_b/N_0 = -1.59$ dB.


1.6
$\frac{P}{N_0}\log_2(e)$
1.4
1.2
1
0.8
C(W)
[Mbps]
0.6
0.4
0.2
0
0
5
10
15
20
25
30
Power Limited Region
Capacity
Limit for $W \to \infty$
Bandwidth Limited Region
Bandwidth W[MHz]

Frequency-selective AWGN Channel

$$y[m] \sum_{\ell} h_\ell x[m-\ell] + w[m]$$

h_ℓ's are time-invariant.

OFDM converts it into a *parallel channel*:

$$\tilde{y}_n = \tilde{h}_n \tilde{d}_n + \tilde{w}_n, \qquad n = 1, \ldots, N_c.$$

$$C_{N_c} = \sum_{n=0}^{N_c-1} \log\left(1 + \frac{P_n^* |\tilde{h}_n|^2}{N_0}\right),$$

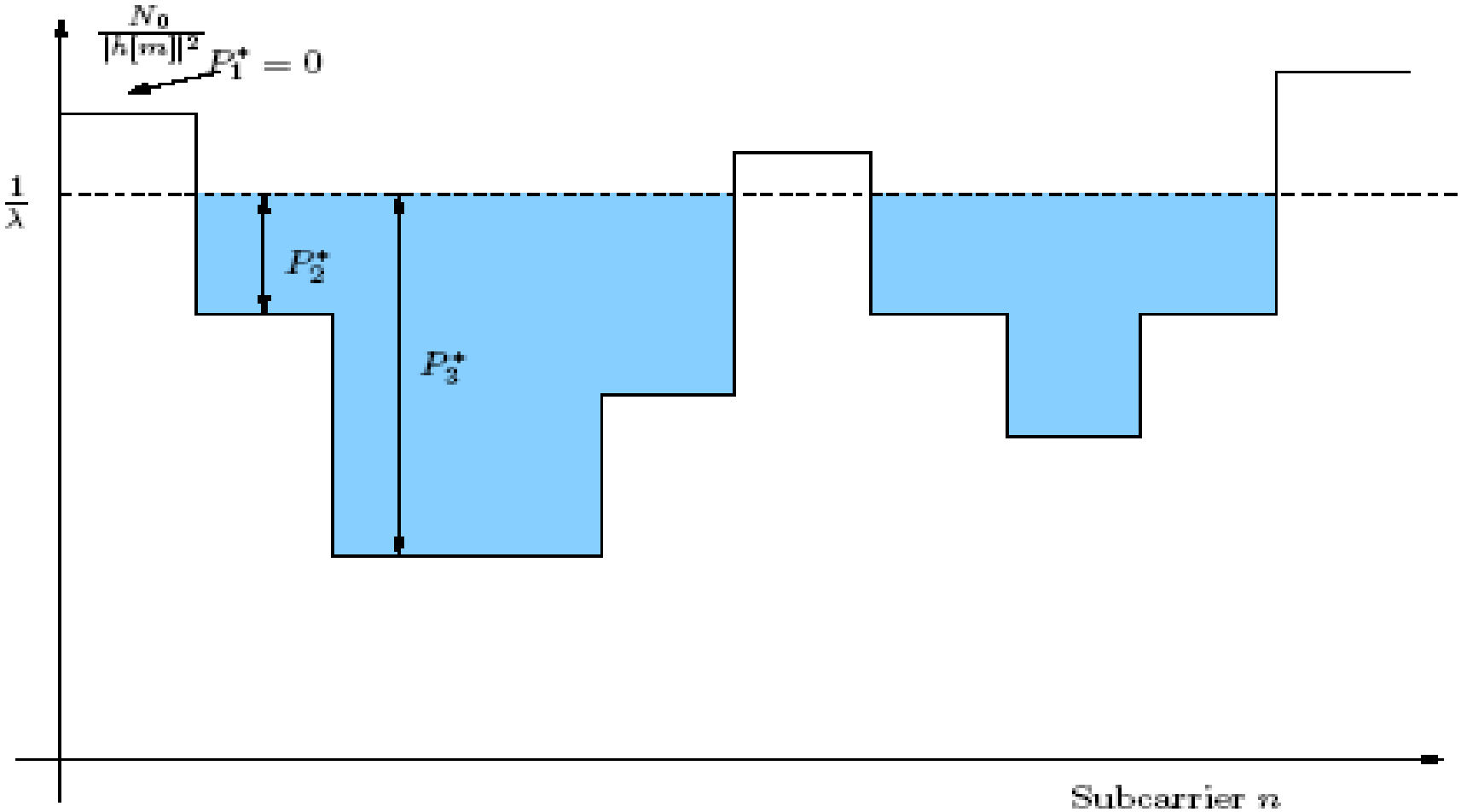
where P_n^* is the waterfilling allocation:

$$P_n^* = \left(\frac{1}{\lambda} - \frac{N_0}{|\tilde{h}_n|^2}\right)^+$$

with λ chosen to meet the power constraint.

Can be achieved with separate coding for each sub-carrier.

Waterfilling in Frequency Domain



Slow Fading Channel

$$y[m] = hx[m] + w[m]$$

h random.

There is no definitive capacity: for any target rate R, there is a probability that the channel cannot support that rate no matter how much coding is done.

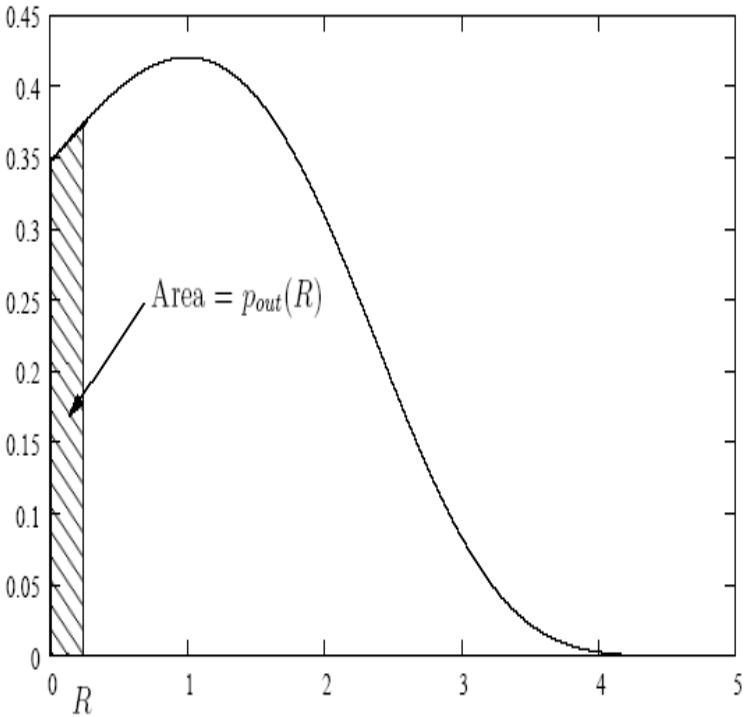
Outage probability $p_{\text{out}}(R) = \mathcal{P}\left\{\log(1 + |h|^2 \text{SNR}) < R\right\}$

$$\epsilon\text{-outage capacity } C_\epsilon = p_{\text{out}}^{-1}(\epsilon)$$

中断概率等于e的时候，R的值——中断率为e时的系统容量

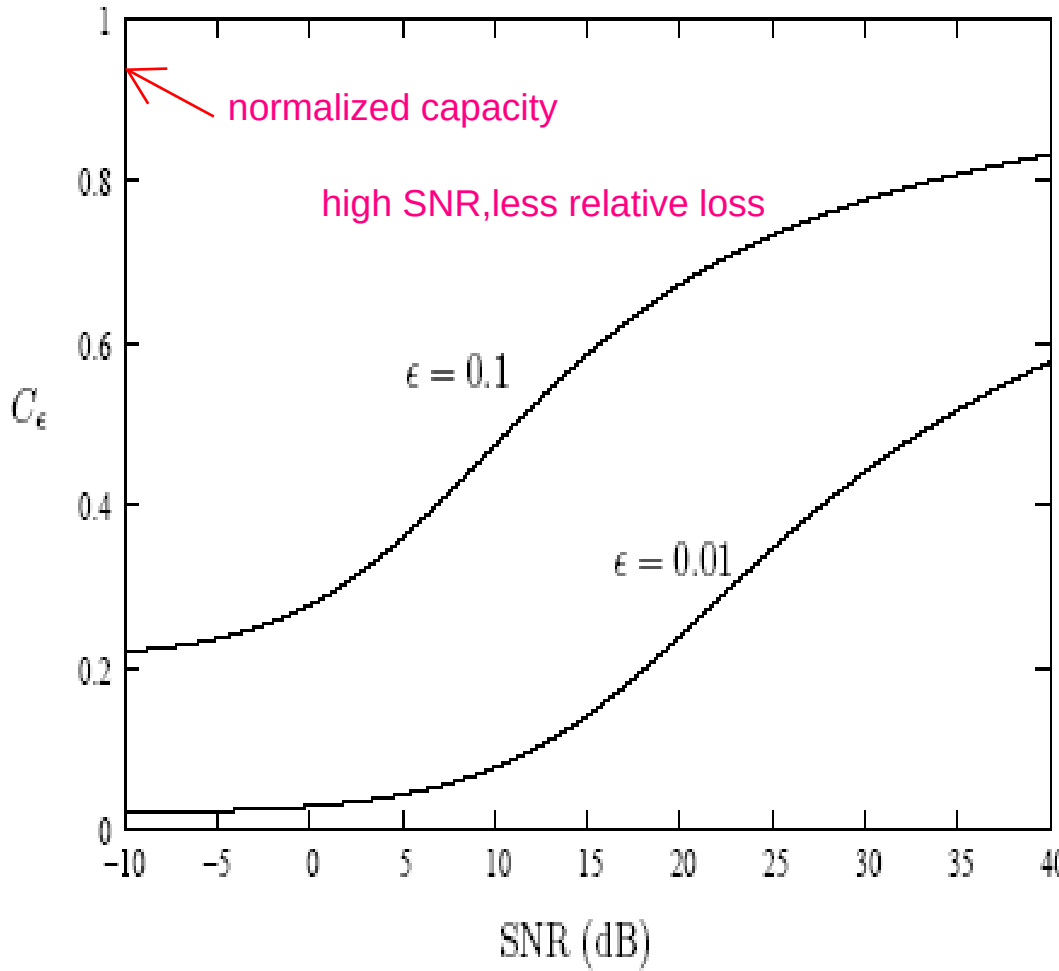
Outage for Rayleigh Channel

Pdf of log(1+|h|²SNR)



$$p_{\text{out}}(R) \approx \frac{2^R - 1}{\text{SNR}}$$

Outage cap. as fraction of AWGN cap.

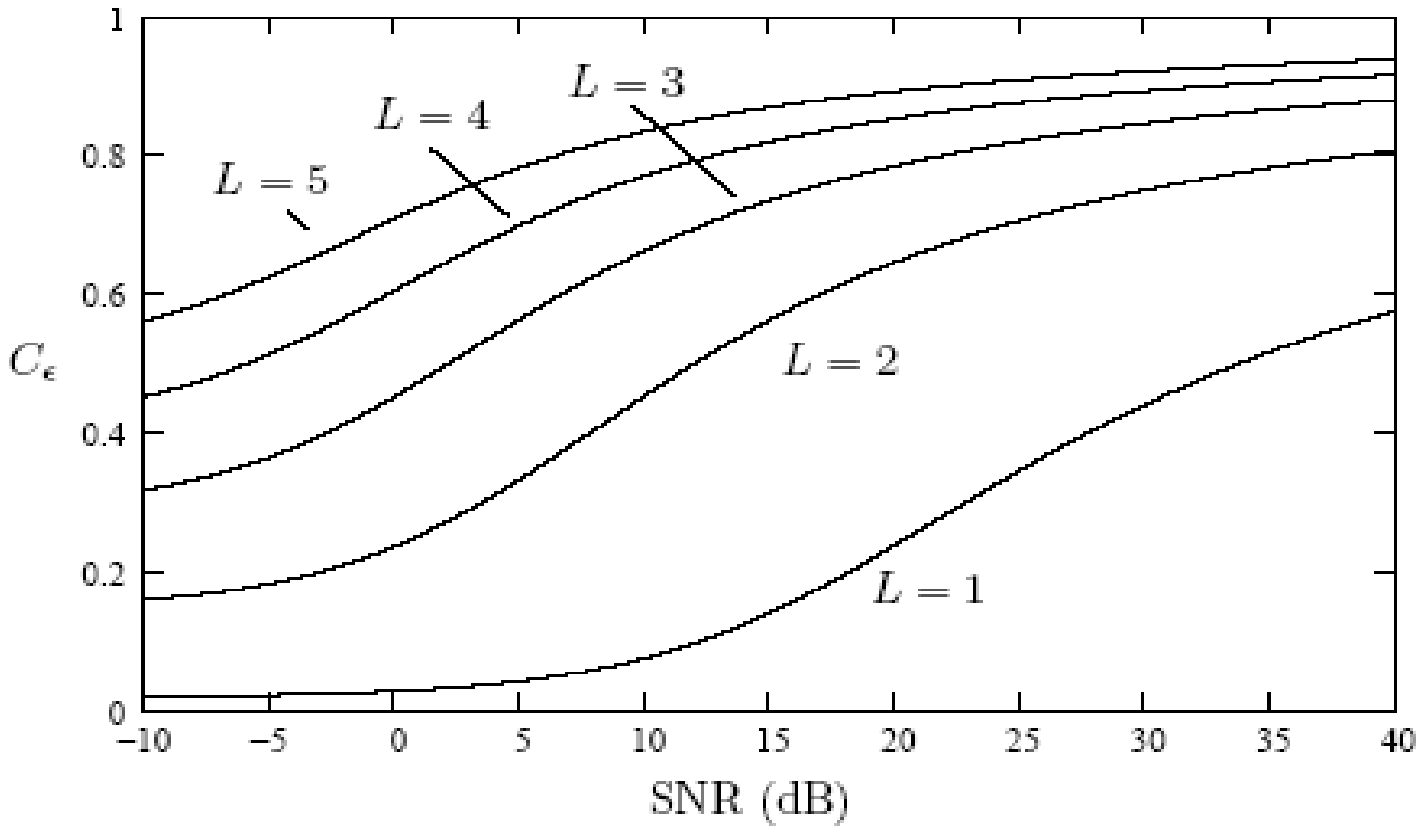


coding只用于对抗高斯噪声，对于瑞利衰落没有效果

Receive Diversity

$$p_{\text{out}}(R) = \mathcal{P}\left\{\log\left(1 + \|\mathbf{h}\|^2 \text{SNR}\right) < R\right\}$$

引入diversity，性能明显提高



Transmit Diversity

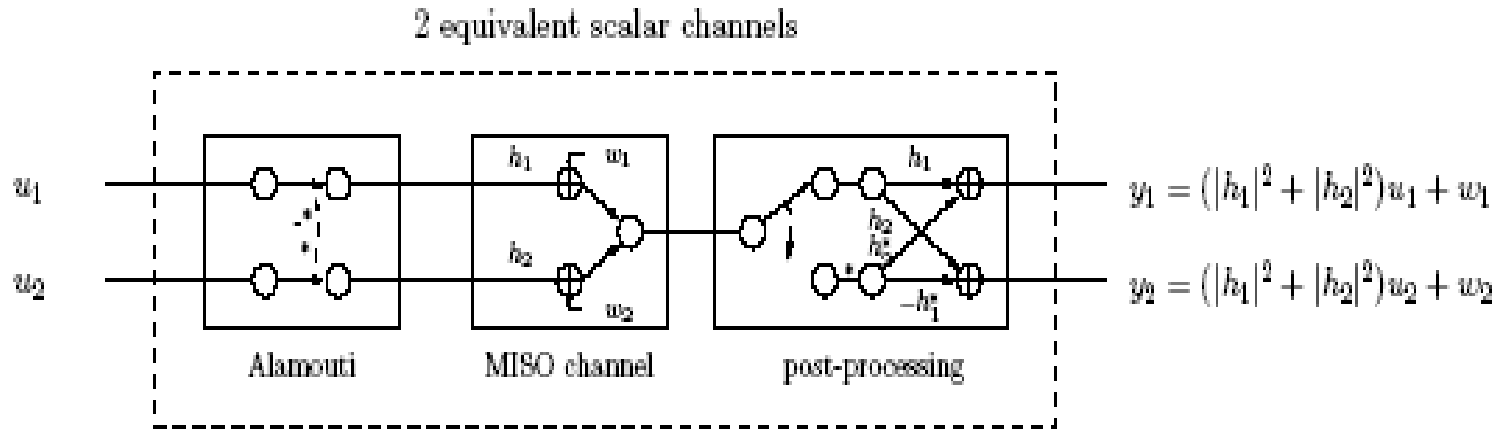
Transmit beamforming:

$$p_{\text{out}}(R) = \mathcal{P}\left\{\log\left(1 + \|\mathbf{h}\|^2 \text{SNR}\right) < R\right\}$$

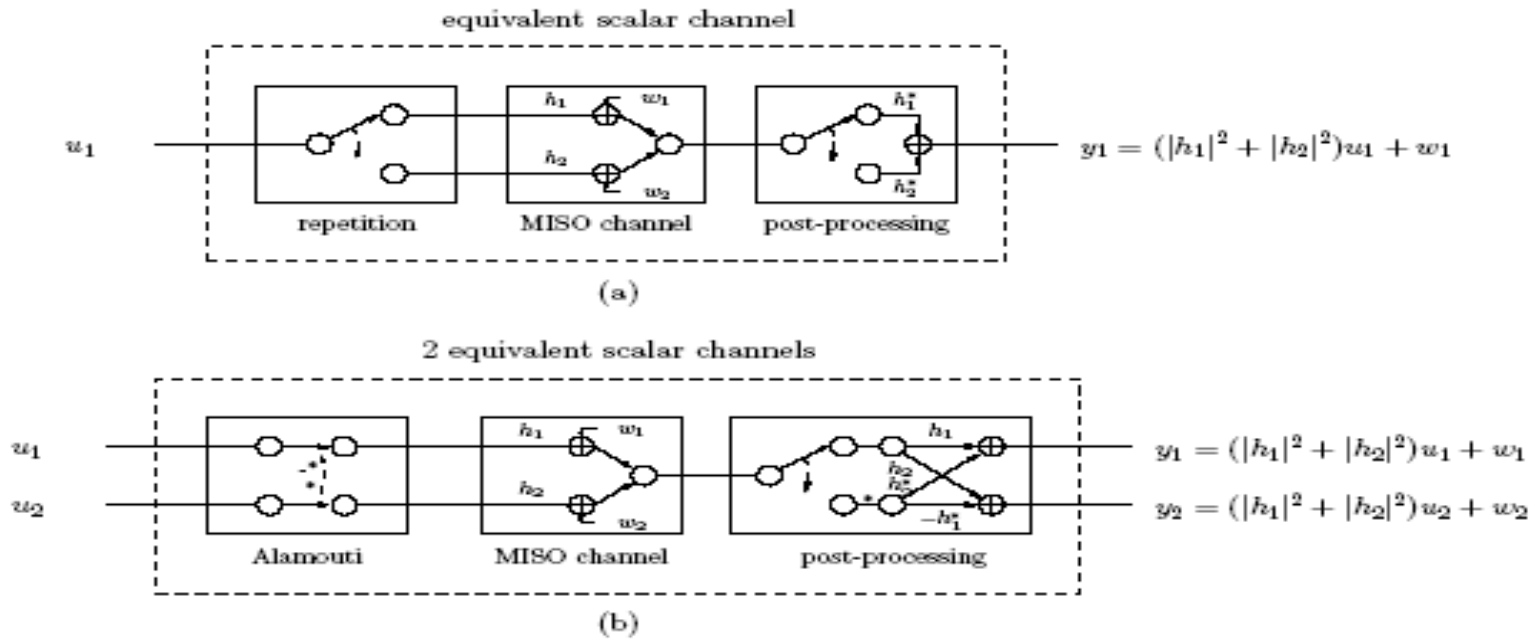
Alamouti (2 Tx):

$$p_{\text{out}}(R) = \mathcal{P}\left\{\log\left(1 + \|\mathbf{h}\|^2 \frac{\text{SNR}}{2}\right) < R\right\}$$

Alamouti is optimal among all schemes that radiate energy isotropically.



Repetition vs Alamouti



Repetition: $p_{\text{out}}(R) = \mathcal{P}\left\{\frac{1}{2}\log\left(1 + \|\mathbf{h}\|^2 \text{SNR}\right) < R\right\}$

Alamouti: $p_{\text{out}}(R) = \mathcal{P}\left\{\log\left(1 + \|\mathbf{h}\|^2 \frac{\text{SNR}}{2}\right) < R\right\}$

Significant loss at high SNR; negligible loss at low SNR.

Time Diversity

$$y_\ell = h_\ell x_\ell + w_\ell, \qquad \ell = 1, \ldots, L$$

Coding done over L coherence blocks, each of many symbols.

This is a parallel channel. If transmitter knows the channel, can do waterfilling.

Without channel knowledge,

$$p_{\text{out}}(R) = \mathcal{P}\left\{\frac{1}{L}\sum_{\ell=1}^{L} \log\left(1 + |h_\ell|^2 \text{SNR}\right) < R\right\}$$

Note: coding across sub-channels is now necessary.

Fast Fading Channel

Channel with L-fold time diversity:

$$p_{\text{out}}(R) = \mathcal{P}\left\{\frac{1}{L}\sum_{\ell=1}^{L}\log\left(1+|h_\ell|^2\text{SNR}\right) < R\right\}$$

As $L \to \infty$,

$$\frac{1}{L}\sum_{\ell=1}^{L}\log\left(1+|h_\ell|^2\text{SNR}\right) \to \mathcal{E}[\log(1+|h|^2\text{SNR})]$$

Fast fading channel has a definite capacity:

$$C = \mathcal{E}[\log(1+|h|^2\text{SNR})]$$

Tolerable delay $\gg$ coherence time.

Waterfilling Capacity

Suppose now transmitter has full channel knowledge.

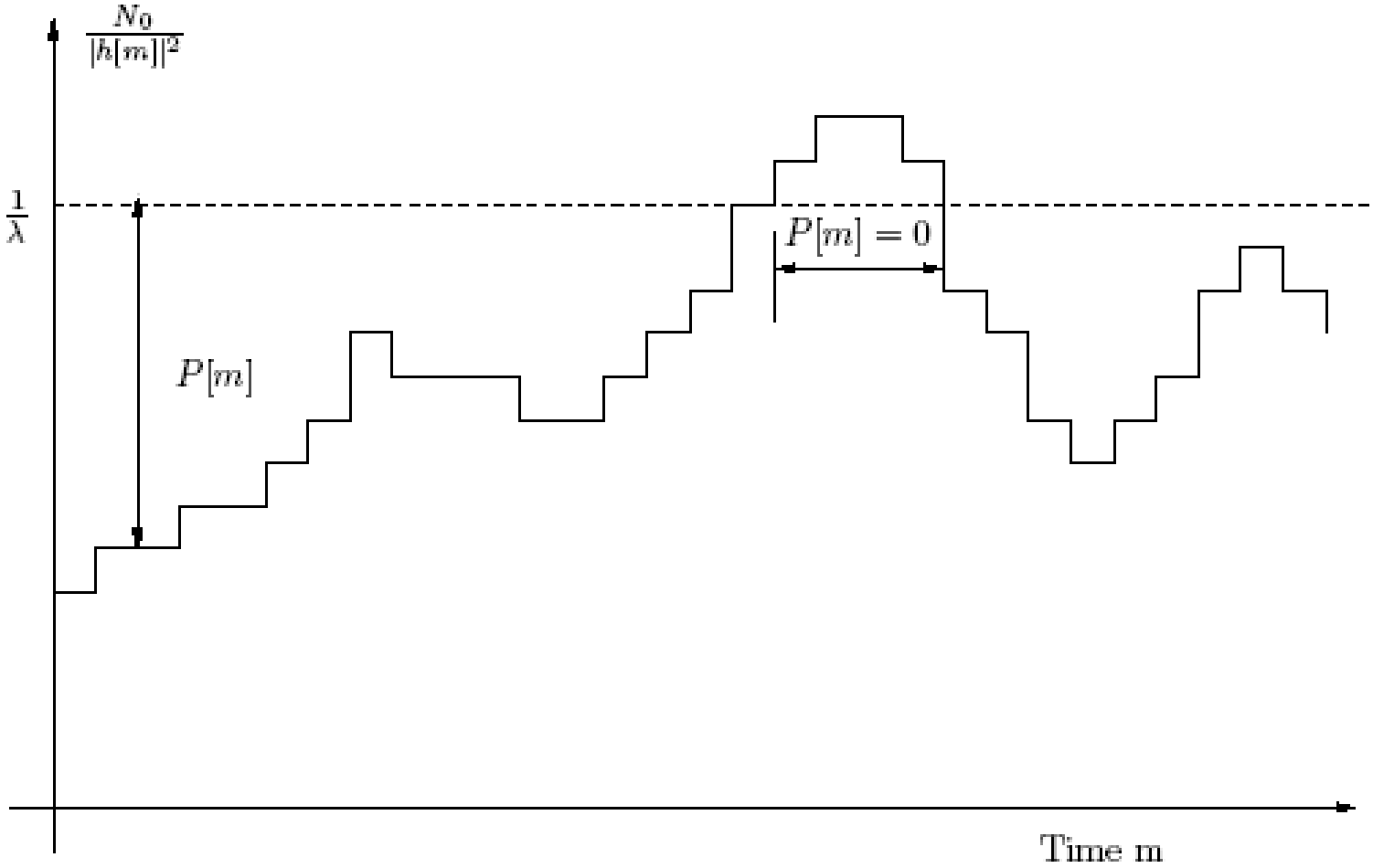
$$C = \mathcal{E}\left[\log\left(1 + \frac{P^*(h)|h|^2}{N_0}\right)\right]$$

where

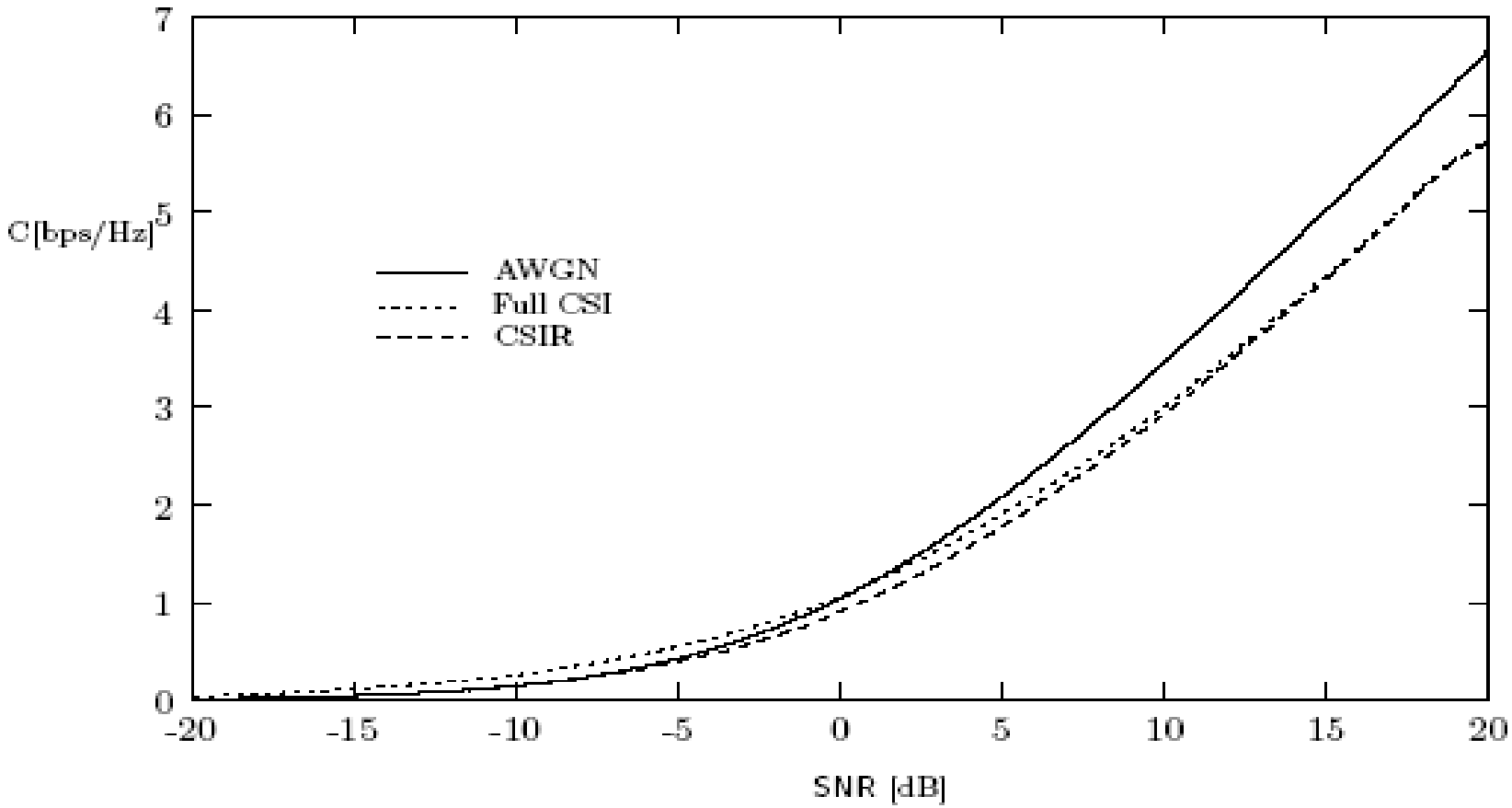
$$P^*(h) = \left(\frac{1}{\lambda} - \frac{N_0}{|h|^2}\right)^+.$$

is the waterfilling power allocation as a function of the fading state. and λ is chosen to satisfy the average power constraint.

Transmit More when Channel is Good

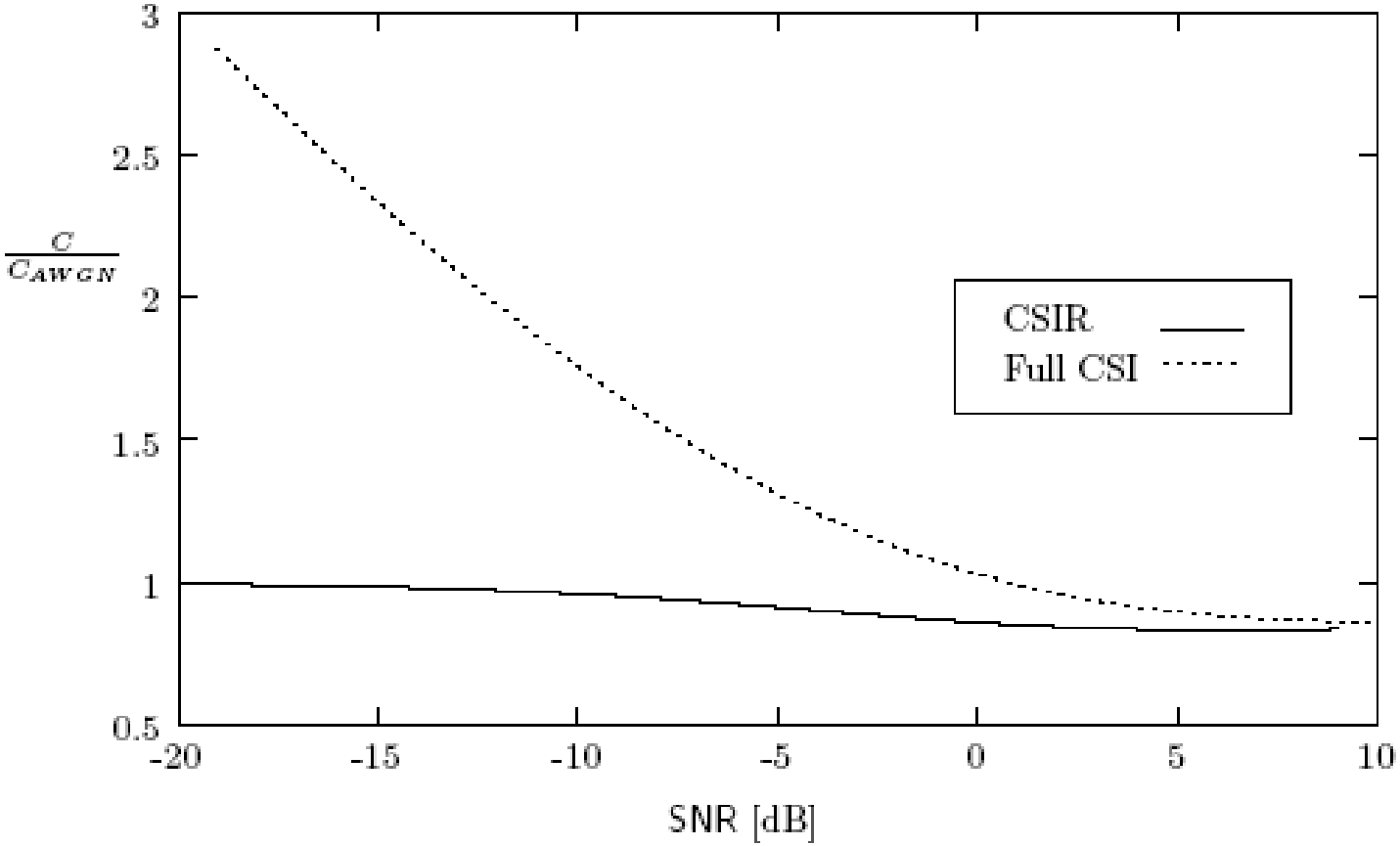


Performance



At high SNR, waterfilling does not provide any gain.

Performance: Low SNR



Waterfilling povides a significant power gain at low SNR.

Summary

- A slow fading channel is a source of unreliability: very poor outage capacity. Diversity is needed.
- A fast fading channel with only receiver CSI has a capacity close to that of the AWGN channel: only a small penalty results from fading. Delay is long compared to channel coherence time.
- A fast fading channel with full CSI can have a capacity greater than that of the AWGN channel: fading now provides more opportunities for performance boost.
- The idea of opportunistic communication is even more powerful in multiuser situations, as we will see.

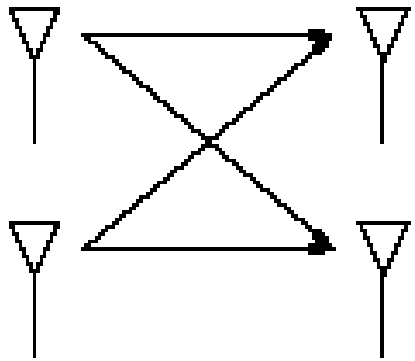
4. MIMO I: Spatial Multiplexing and Channel Modeling

Main Story

- So far we have only considered single-input multi-output (SIMO) and multi-input single-output (MISO) channels.
- They provide diversity and power gains but no degree-of-freedom (d.o.f.) gain.
- D.o.f gain is most useful in the high SNR regime.
- MIMO channels have a potential to provide d.o.f gain.
- We would like to understand how the d.o.f gain depends on the physical environment and come up with statistical models that capture the properties succinctly.
- We start with deterministic models and then progress to statistical ones.

MIMO Capacity via SVD

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

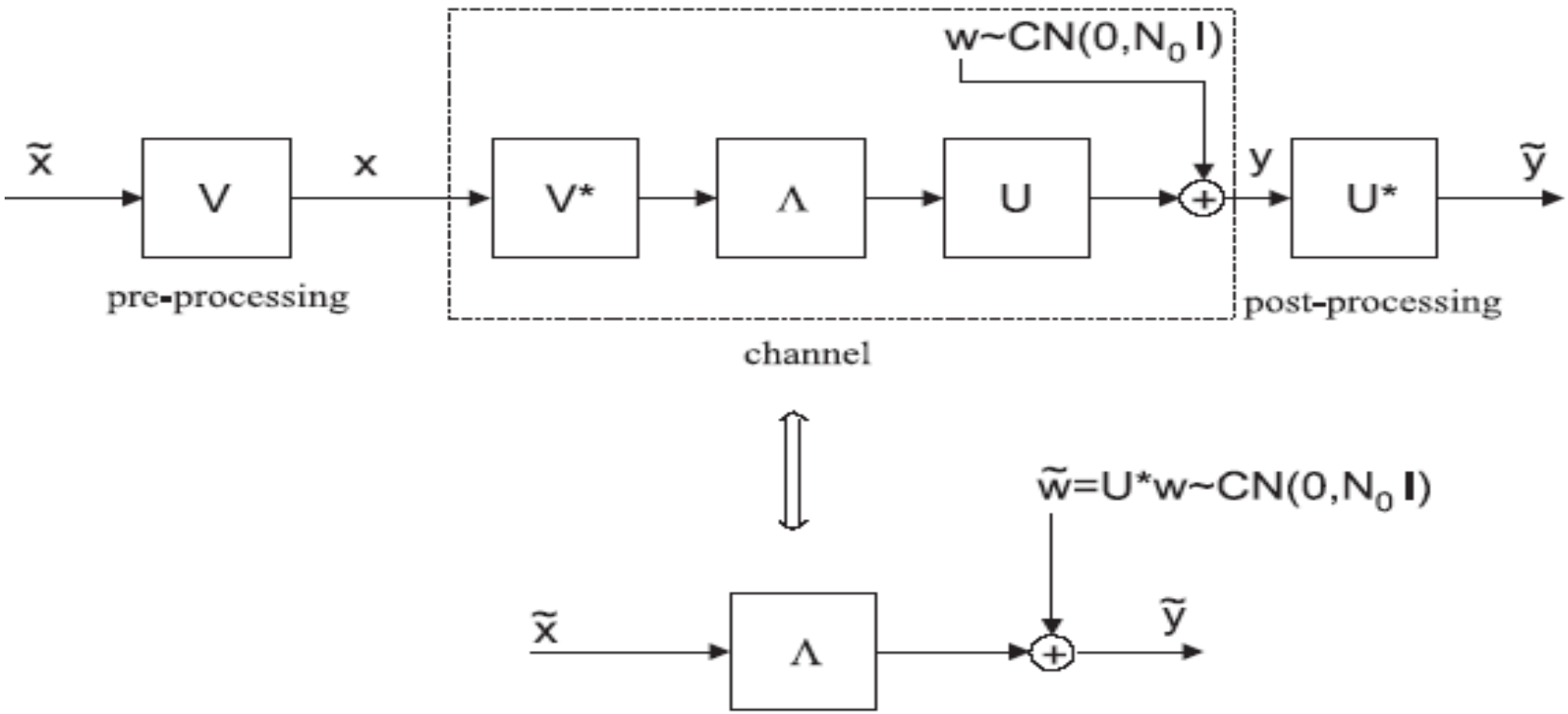


$\mathbf{H}$ is n_r by n_t, fixed channel matrix. SVD:

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$$

$\mathbf{U}, \mathbf{V}$ are unitary matrices and $\mathbf{\Lambda}$ real diagonal (singular values).

Spatial Parallel Channel



Capacity is achieved by waterfilling over the eigenmodes of H.

Rank and Condition Number

At high SNR, equal power allocation is optimal:

$$C \approx \sum_{i=1}^{k} \log\left(1 + \frac{P\lambda_i^2}{kN_0}\right) \approx k \log \mathsf{SNR} + \sum_{i=1}^{k} \log\left(\frac{\lambda_i^2}{k}\right)$$

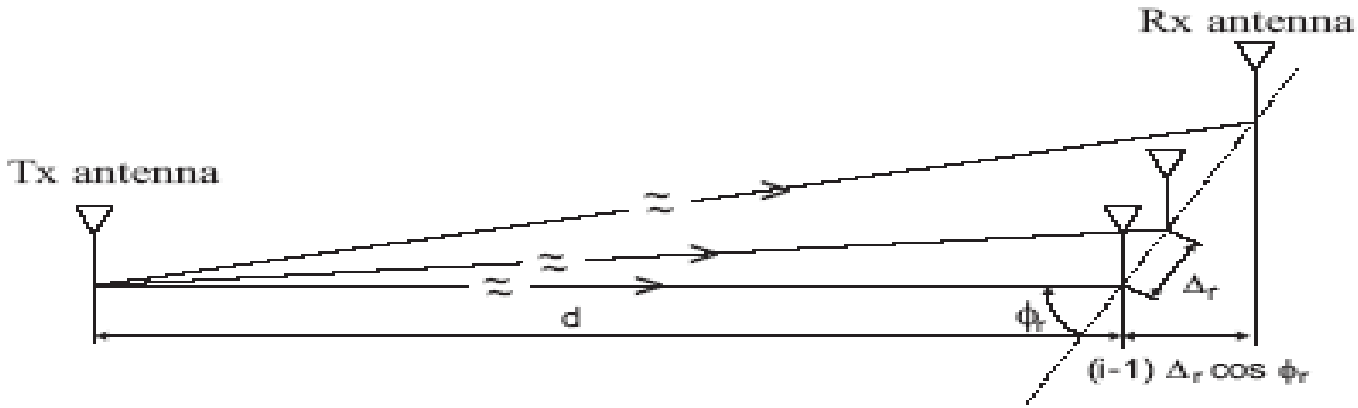
where k is the number of nonzero λ_i^2's, i.e., the rank of $\mathbf{H}$.

The closer the condition number:

$$\frac{\max_i \lambda_i}{\min_i \lambda_i}$$

to 1, the higher the capacity.

Example 1: SIMO, Line-of-sight

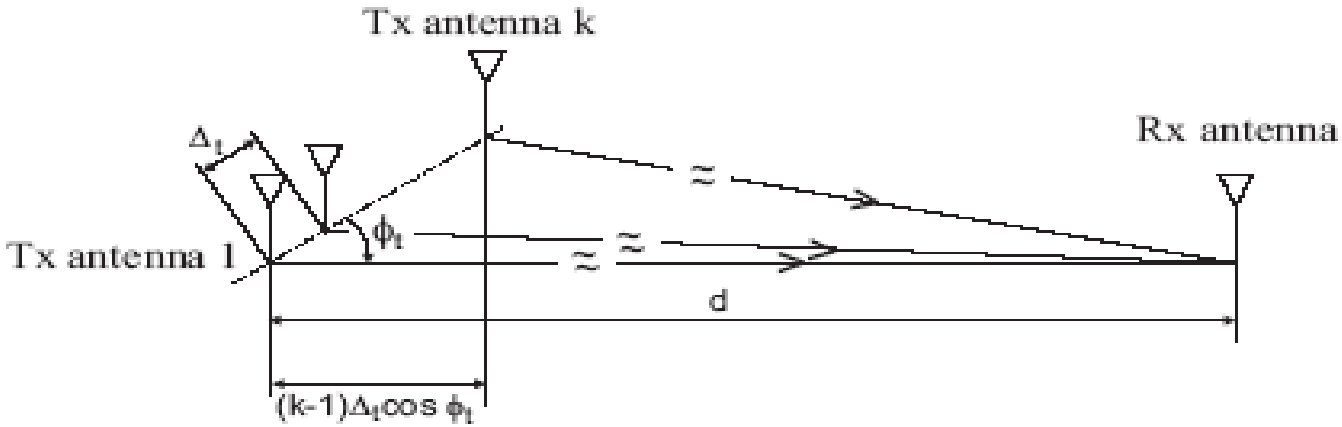


$$\mathbf{y} = \mathbf{h}x + \mathbf{w}$$

$\mathbf{h}$ is along the receive spatial signature in the direction $\Omega := \cos\phi$:

$$\mathbf{e_r}(\Omega) := \frac{1}{\sqrt{n_r}} \begin{bmatrix} 1 \\ \exp(-j2\pi\Delta_r\Omega) \\ \exp(-j2\pi 2\Delta_r\Omega) \\ \cdot \\ \cdot \\ \exp(-j2\pi(n_r-1)\Delta_r\Omega) \end{bmatrix}$$

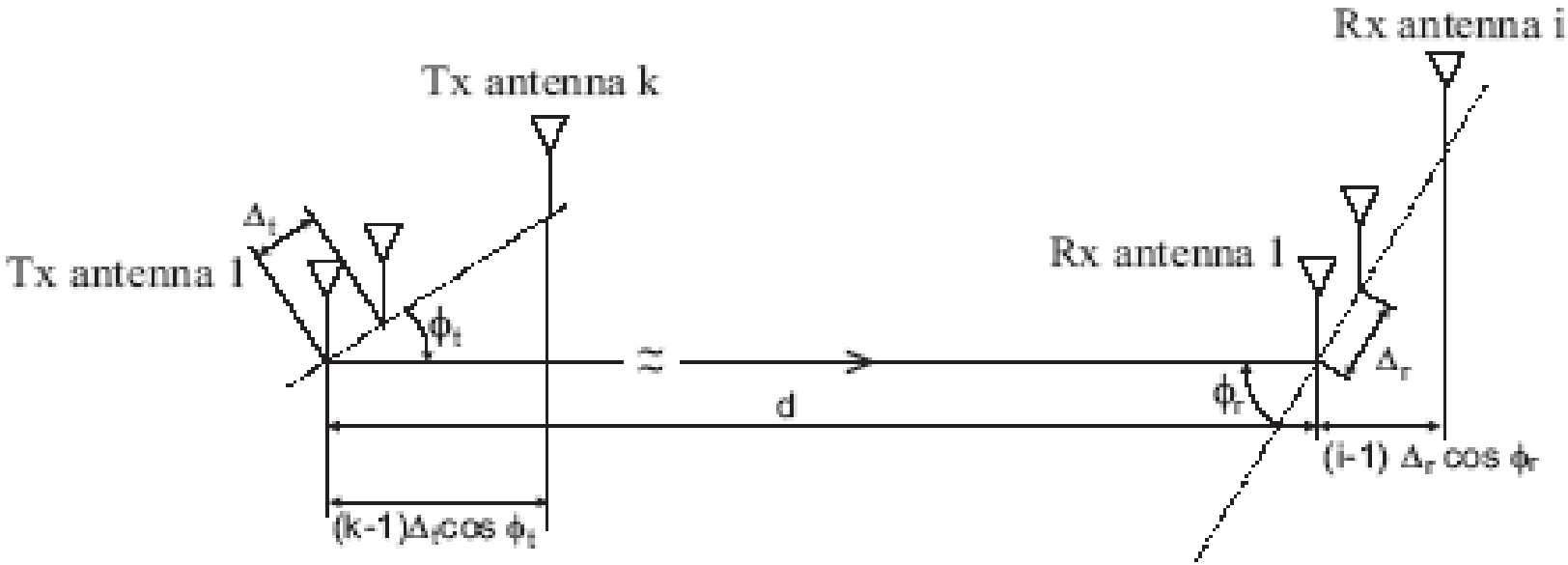
Example 2: MISO, Line-of-Sight



$\mathbf{h}$ is along the transmit spatial signature in the direction $\Omega := \cos\phi$:

$$\mathbf{e_t}(\Omega) := \frac{1}{\sqrt{n_t}} \begin{bmatrix} 1 \\ \exp(-j2\pi\Delta_t\Omega) \\ \exp(-j2\pi 2\Delta_t\Omega) \\ \cdot \\ \cdot \\ \exp(-j2\pi(n_t-1)\Delta_t\Omega) \end{bmatrix}$$

Example 3: MIMO, Line-of-Sight

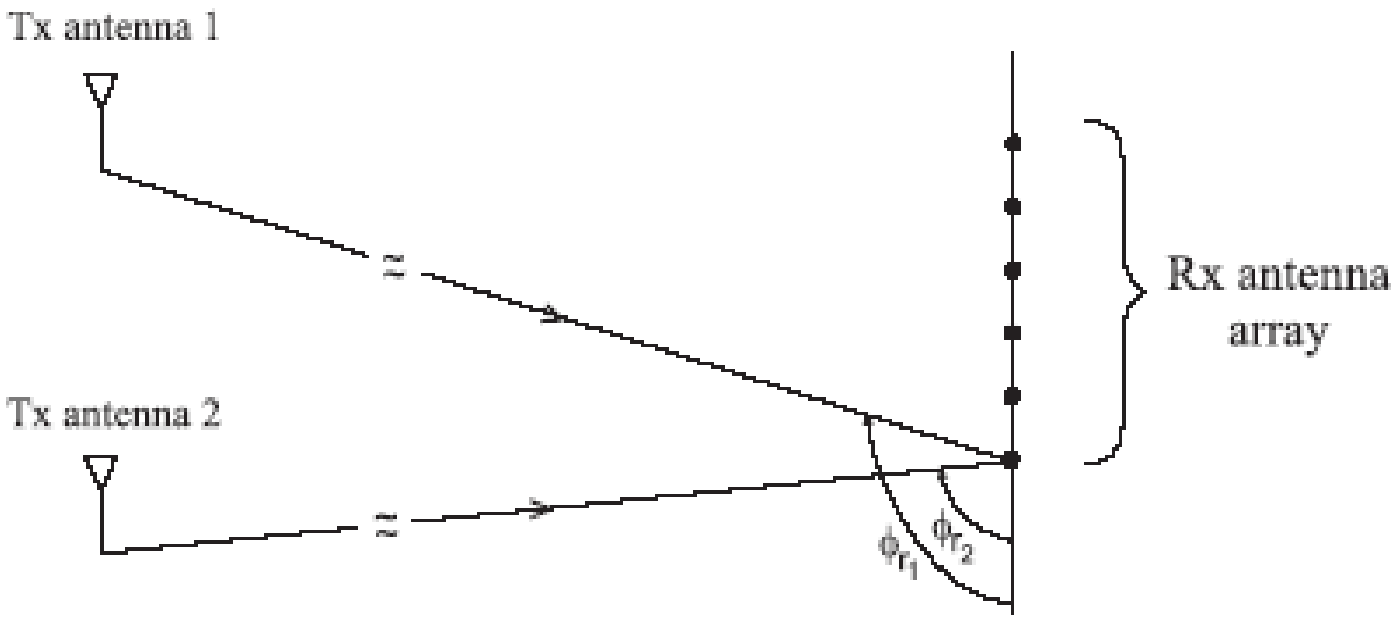


$$\mathbf{H} = \beta \cdot \mathbf{e_r}\left(\Omega_r\right) \mathbf{e_t}\left(\Omega_t\right)^*$$

Rank 1, only one degree of freedom.

No spatial multiplexing gain.

Example 4: MIMO, Tx Antennas Apart



$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2]$$

$\mathbf{h}_i$ is the receive spatial signature from Tx antenna i along direction $\Omega_i = \cos\phi_{ri}$:

$$\mathbf{h}_i = \beta_i \cdot \mathbf{e_r}\left(\Omega_i\right).$$

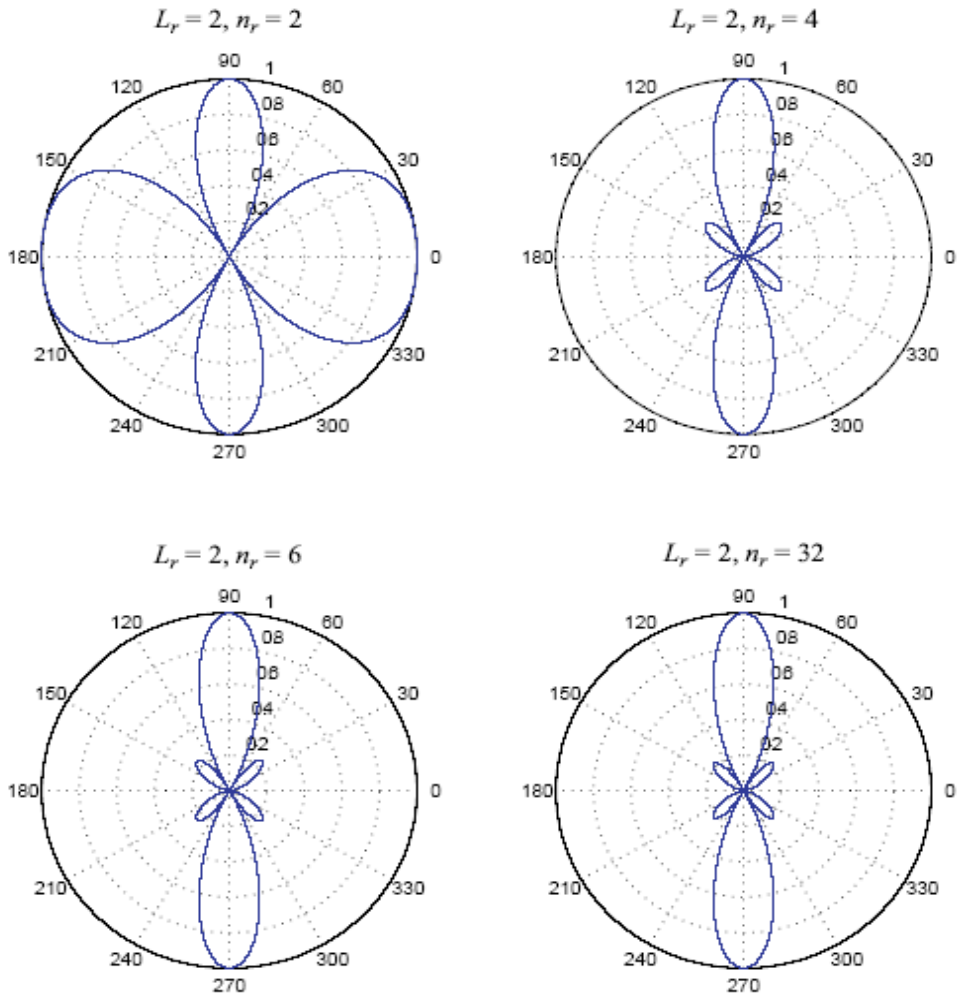
Condition number depends on $|\mathbf{e_r}\left(\Omega_1\right))^* \mathbf{e_r}\left(\Omega_2\right)|$.

Beamforming Patterns

The receive beamforming pattern associated with $\mathbf{e}_r(\Omega)$:

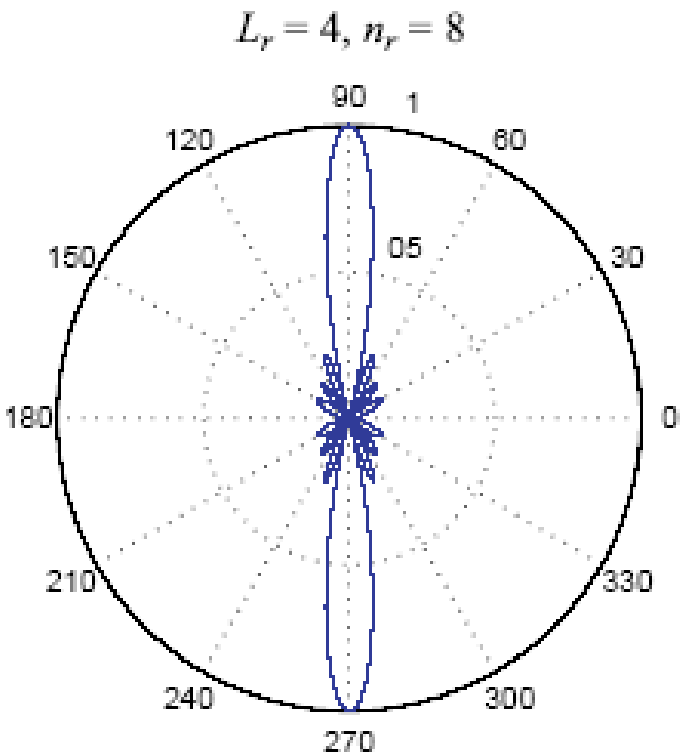
$$B_r(\phi) := |\mathbf{e}_r(\Omega)^* \mathbf{e}_r(\cos\phi)|$$

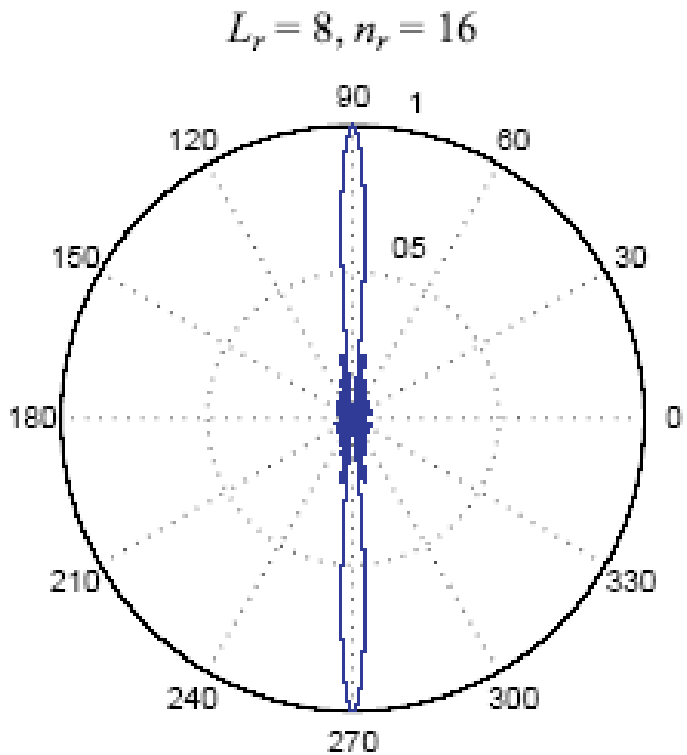
L_r is the length of the antenna array, normalized to the carrier wavelength.



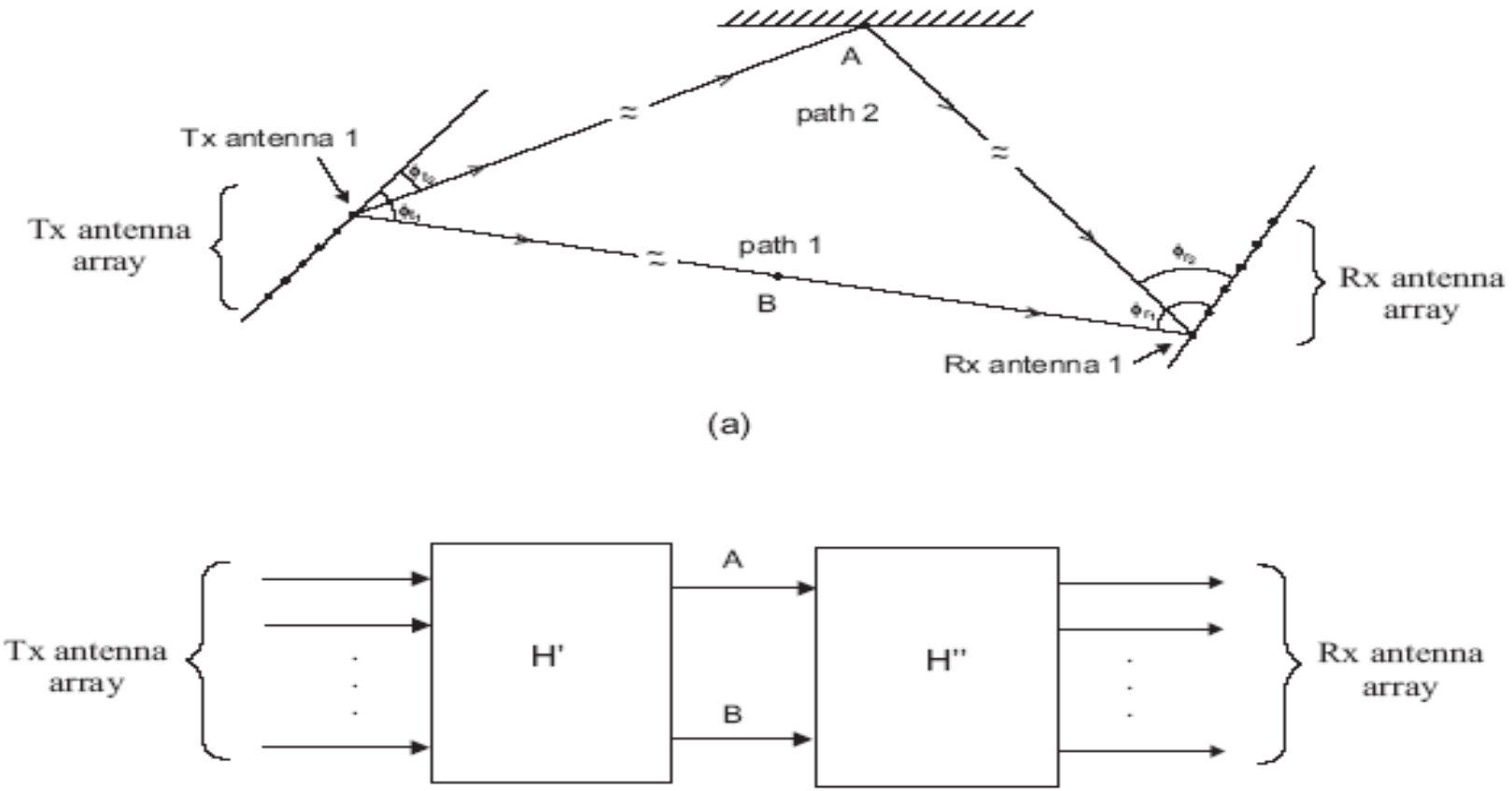
Angular Resolution

Antenna array of length L_r provides angular resolution of $1/L_r$: paths that arrive at angles closer is not very distinguishable.





Example 5: Two-Path MIMO



- A scattering environment provides multiple degrees of freedom even when the antennas are close together.

Channel Modeling in Angular Domain

Represent transmitted and received signal in the angular instead of spatial domain.

Transmit angular basis (orthonormal):

$$\mathcal{S}_t := \left\{ \mathbf{e_t}(0), \mathbf{e_t}\left(\frac{1}{L_t}\right), \ldots, \mathbf{e_t}\left(\frac{n_t - 1}{L_t}\right) \right\}$$

Receive angular basis (orthonormal):

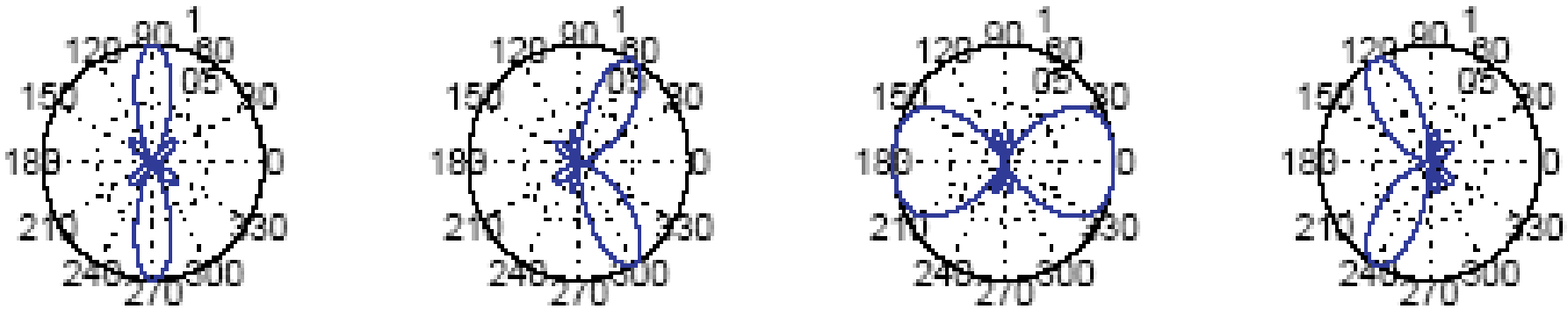
$$\mathcal{S}_r := \left\{ \mathbf{e_r}(0), \mathbf{e_r}\left(\frac{1}{L_r}\right), \ldots, \mathbf{e_r}\left(\frac{n_r - 1}{L_r}\right) \right\}$$

Channel matrix in angular domain:

$$\mathbf{H^a} := \mathbf{U}_r^* \mathbf{H} \mathbf{U}_t$$

where $\mathbf{U}_r$ and $\mathbf{U}_t$ are the receive and transmit change of basis matrices.

Angular Basis



(a) $L_r = 2, n_r = 4$

- The angular transformation decomposes the received (transmit) signals into components arriving (leaving) in different directions.

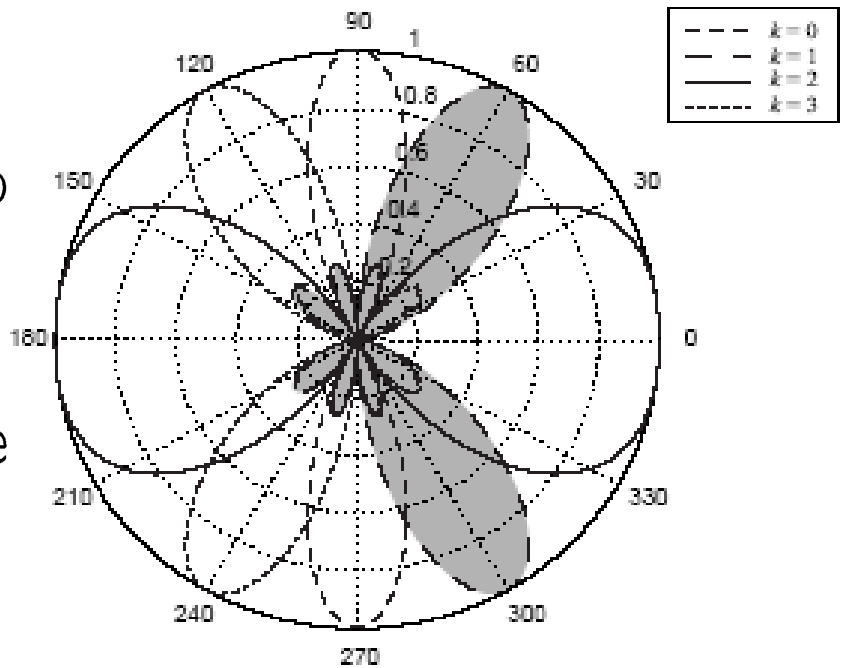
Statistical Modeling in Angular Domain

The transmit basis splits the outgoing paths into bins $\{\mathcal{T}_l\}$ of angular width $1/L_t$

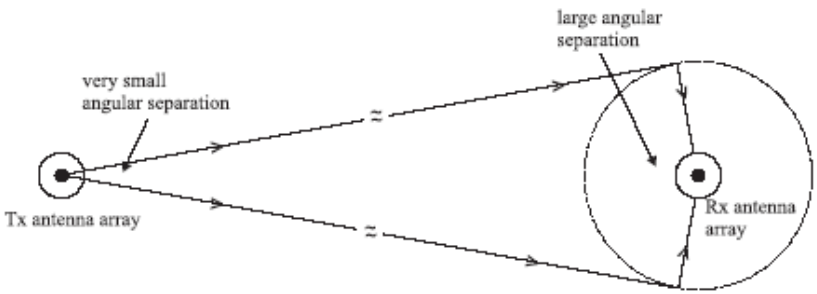
The receive basis splits the incoming paths into bins $\{\mathcal{R}_k\}$ of angular $1/L_r$.

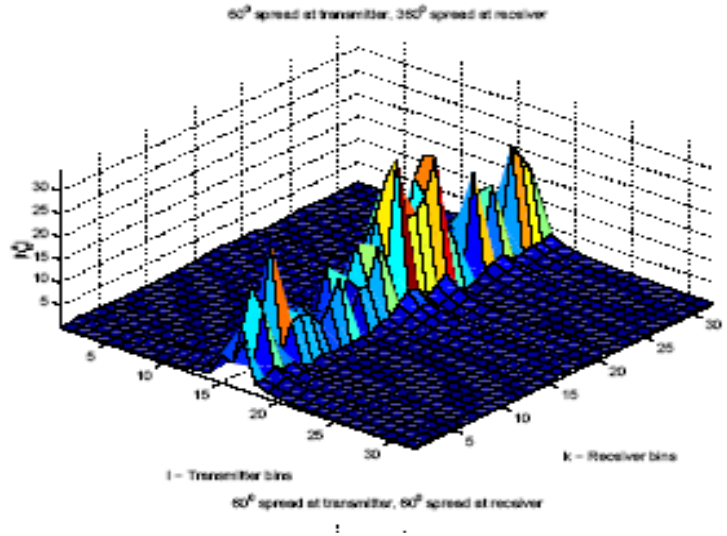
The $(k,l)^{\text{th}}$ entry of $\mathbf{H}^{\mathbf{a}}$ is (approximately) the aggregation of paths in $\mathcal{T}_l \cap \mathcal{R}_k$.

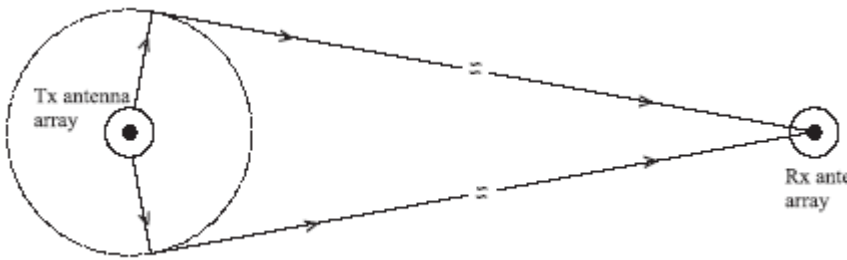
Can statistically model each entry as Gaussian.

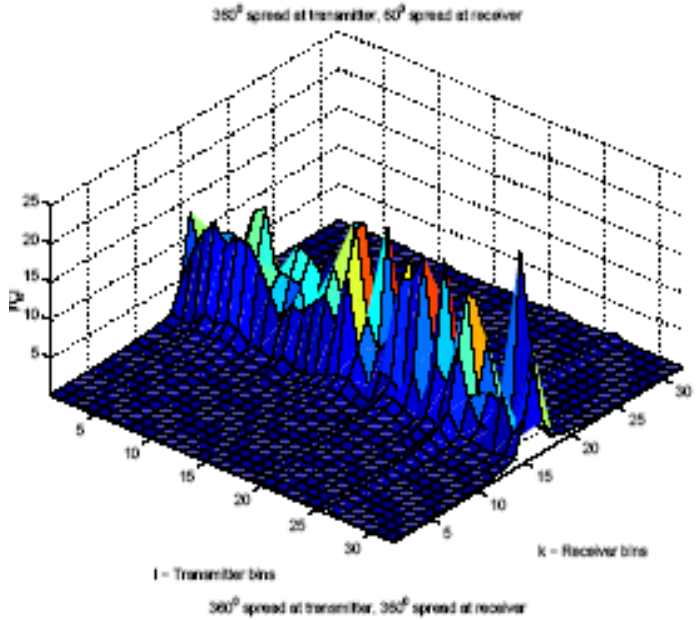


Examples

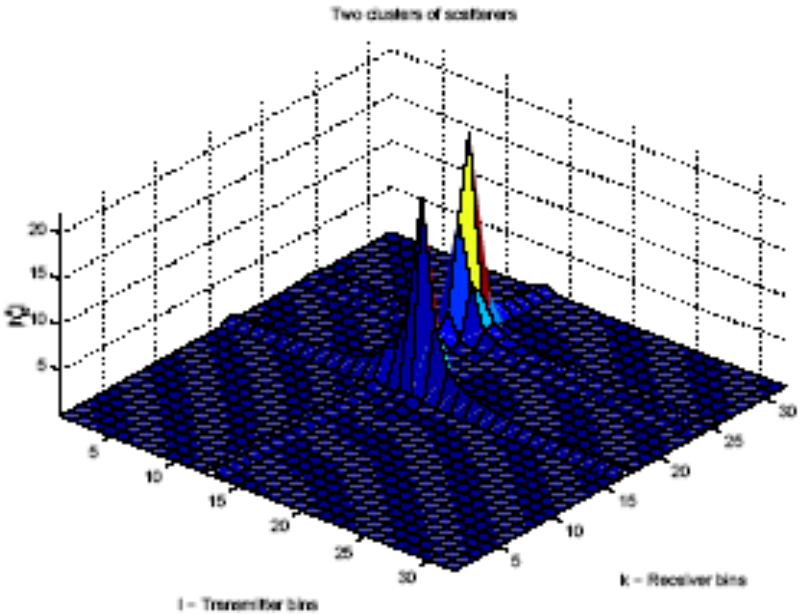

very small
angular separation
large angular
separation
Tx antenna array
Rx antenna
array

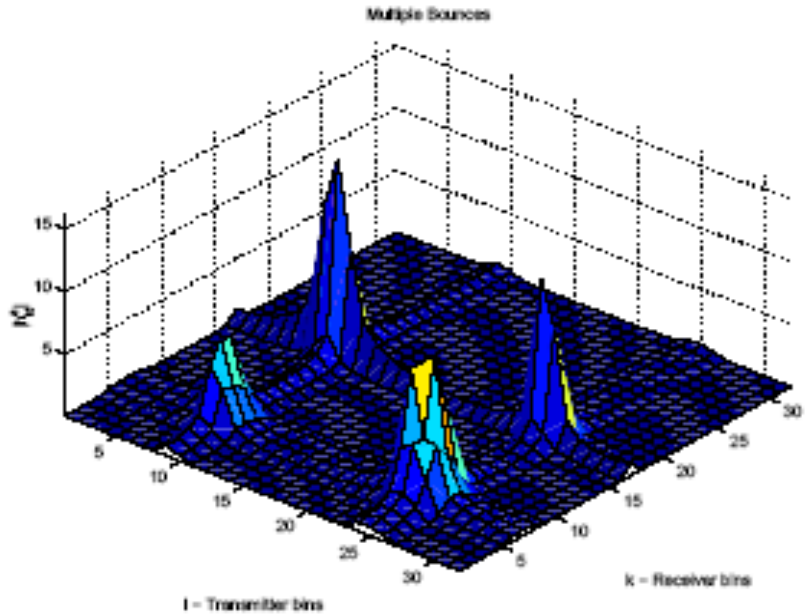

60° spread at transmitter, 360° spread at receiver
|h|
k – Receiver bins
l – Transmitter bins
60° spread at transmitter, 60° spread at receiver

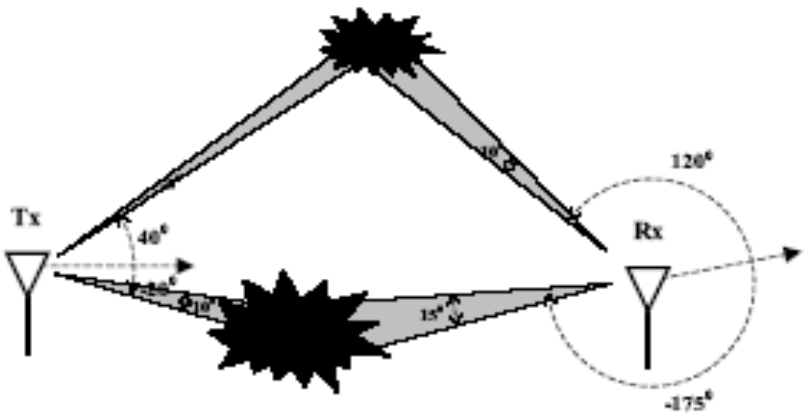

Tx antenna
array
Rx antenna
array

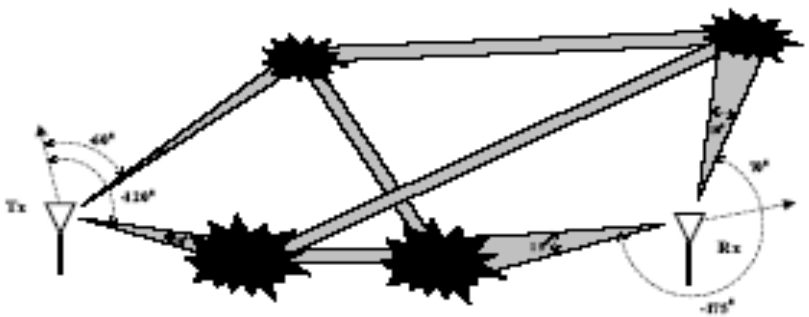

360° spread at transmitter, 60° spread at receiver
|h|
k – Receiver bins
l – Transmitter bins
360° spread at transmitter, 360° spread at receiver

More Examples

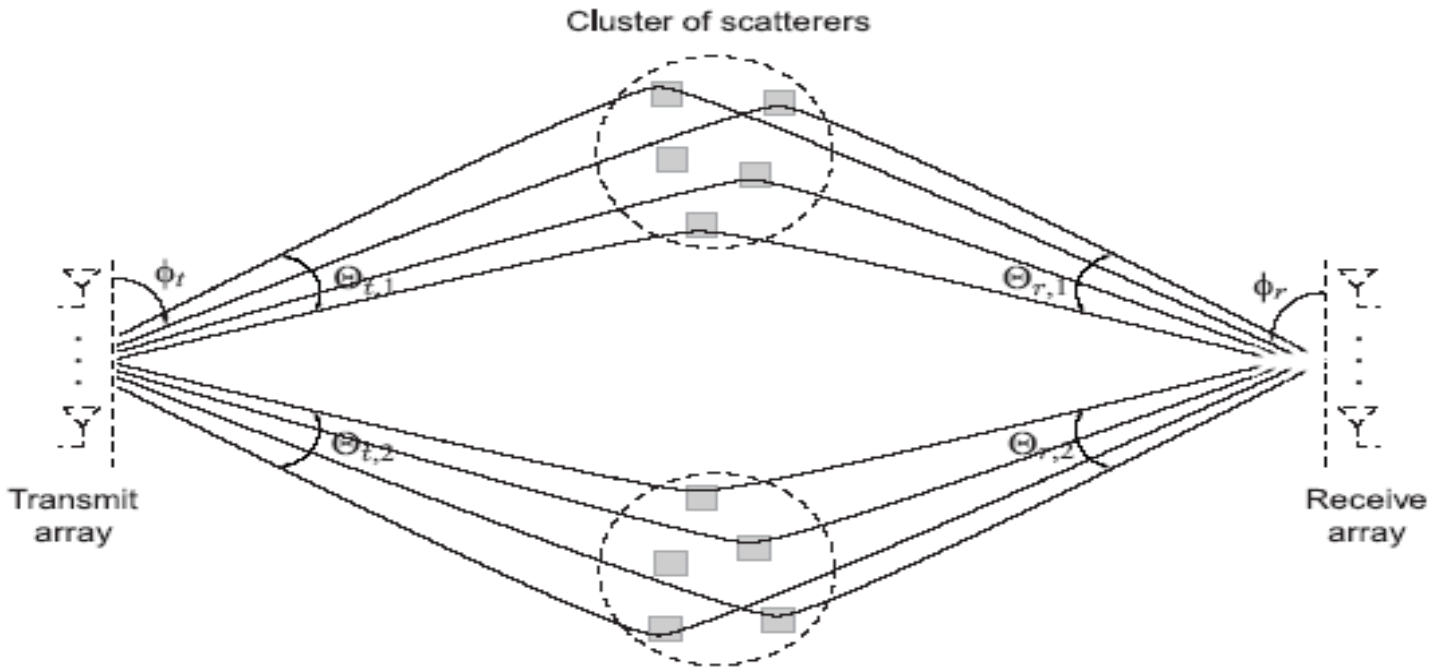








Clustered Model



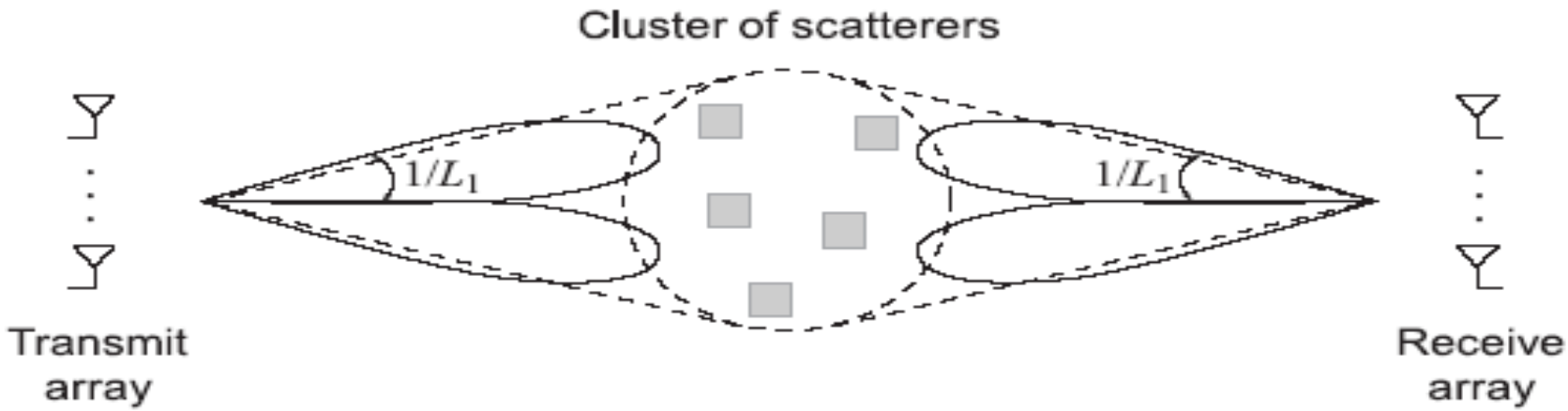
Number of dof:

$$\min\{L_t\Omega_t, L_r\Omega_r\}$$

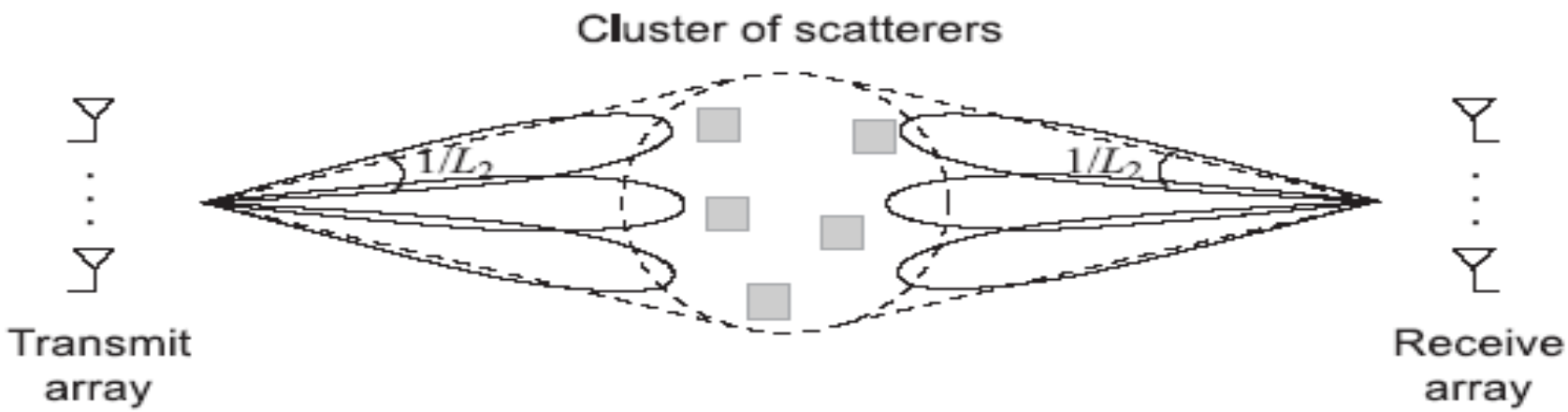
where

$$\Omega_t = \sum_i \Theta_{t,i}, \qquad \Omega_r = \sum_i \Theta_{r,i}$$

Dependency on Antenna Size

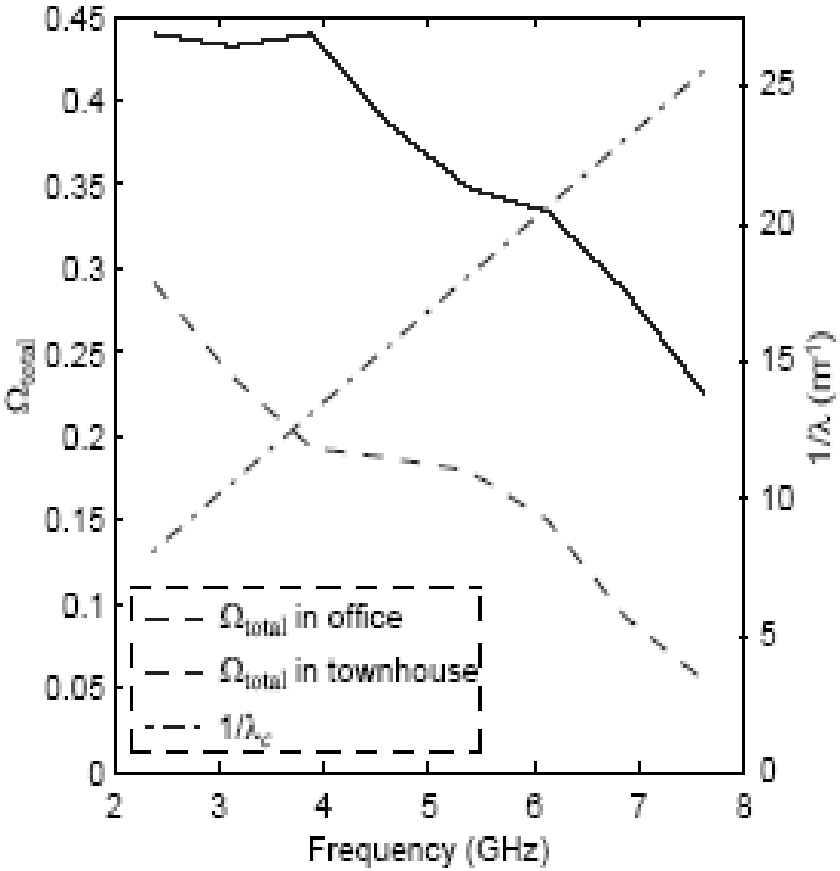


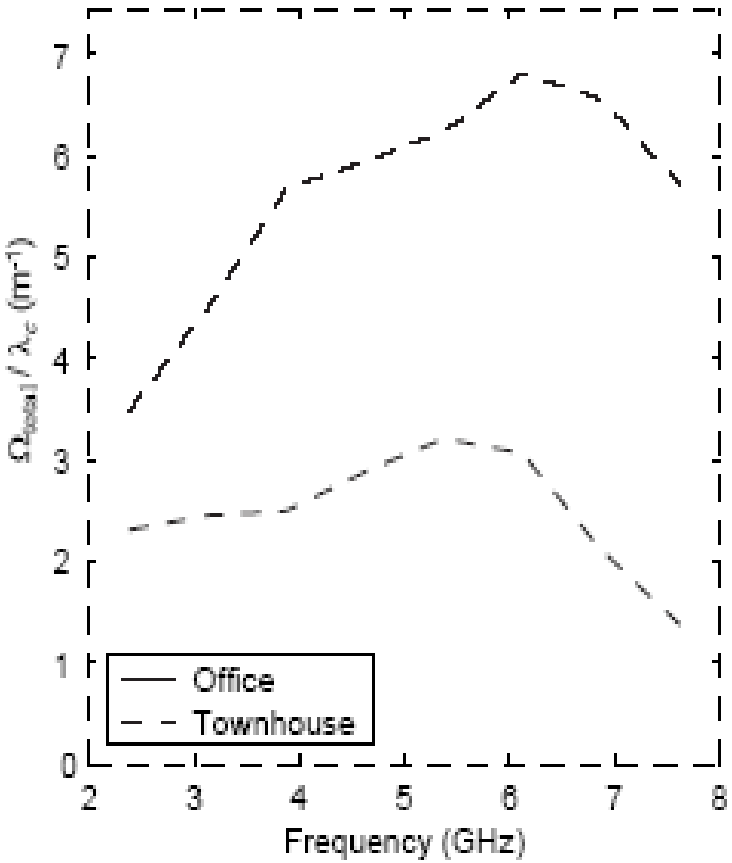
(a) Array length of L_1



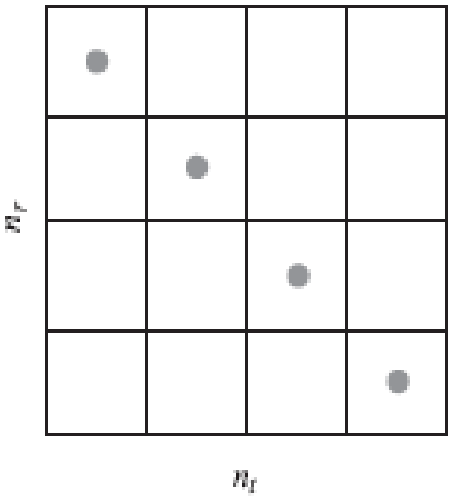
(b) Array length of $L_2 > L_1$

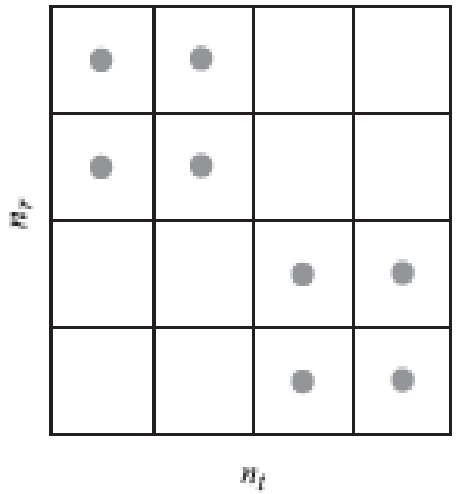
Dependency of dof on Carrier Frequency

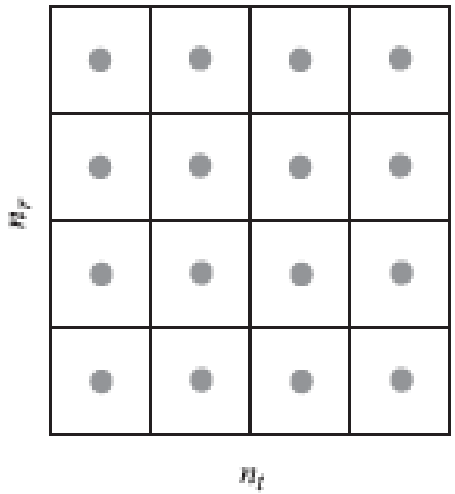




Diversity and Dof

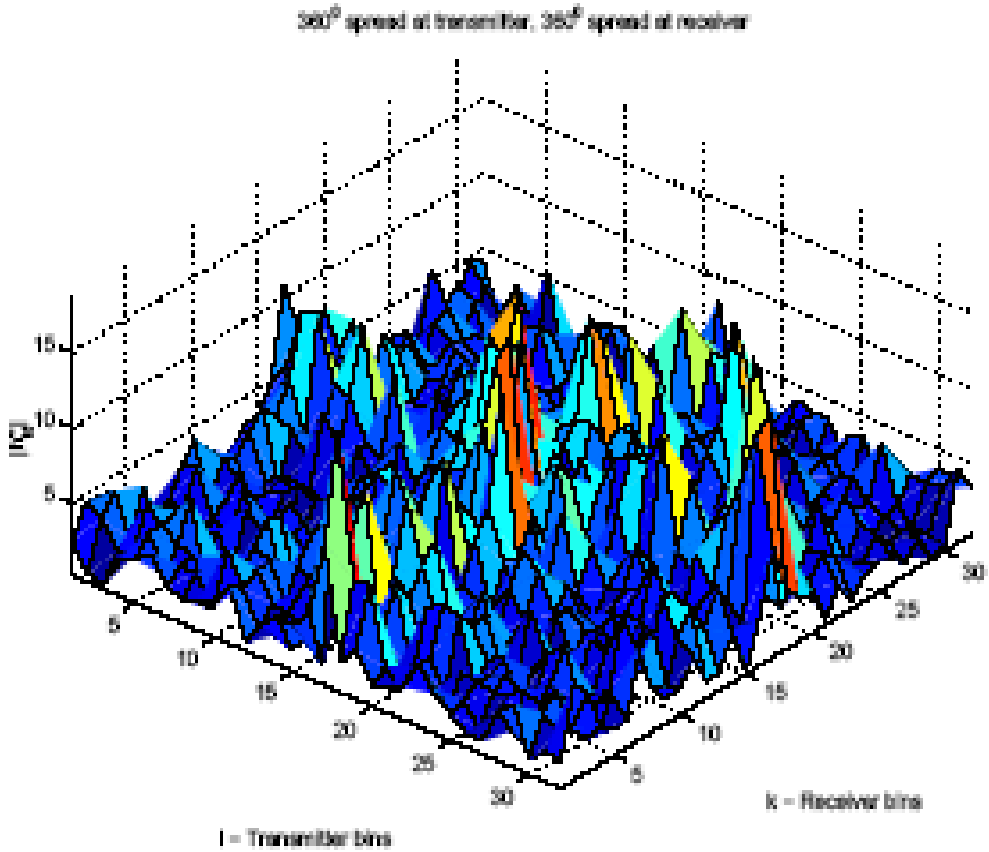






I.I.D. Rayleigh Model

Scatterers at all angles from Tx and Rx.



$$\mathbf{H} = \mathbf{U}_r \mathbf{H}^{\mathrm{a}} \mathbf{U}_t^*$$

$\mathbf{H}^{\mathrm{a}}$ i.i.d. Rayleigh $\Rightarrow$ $\mathbf{H}$ i.i.d. Rayleigh

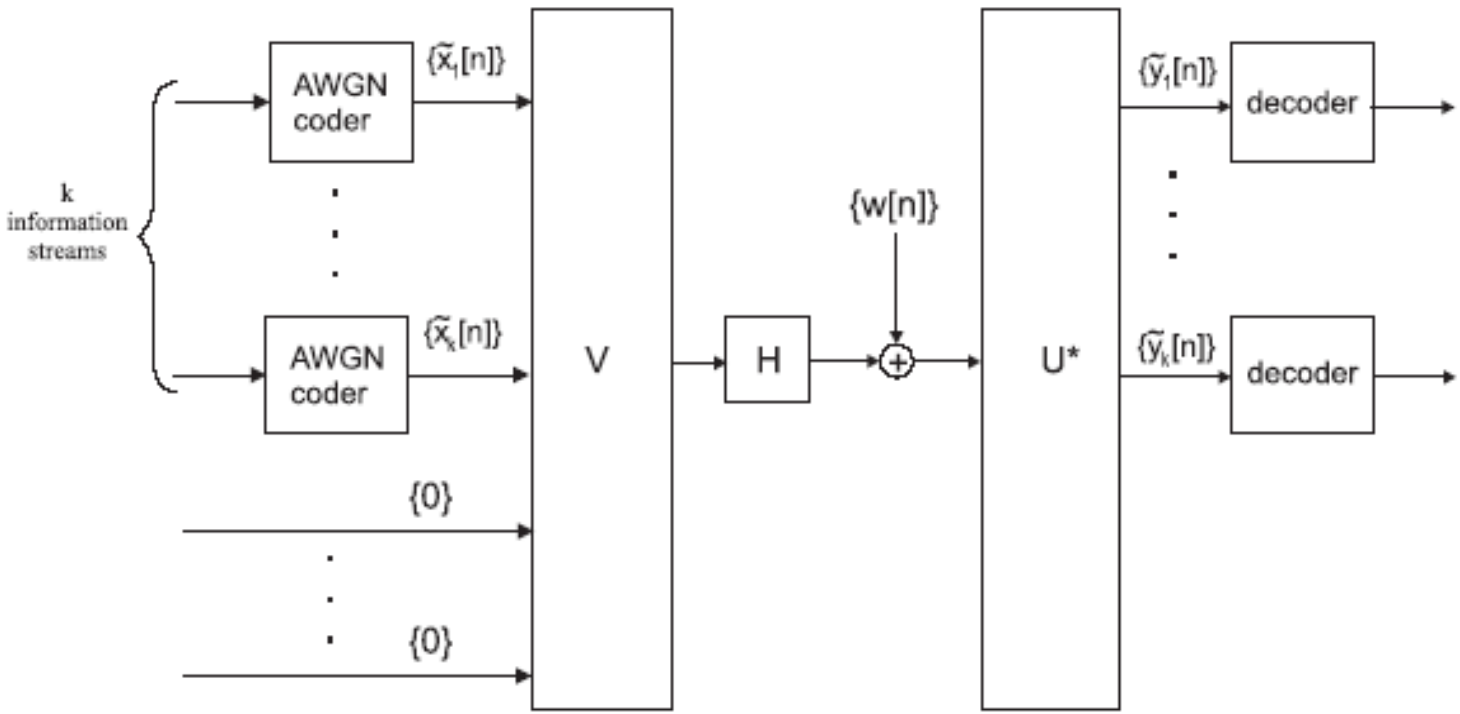
5. MIMO Capacity and Multiplexing Architectures

Outline

- Capacity of MIMO channels
- Nature of performance gains
- Receiver architectures for fast fading (V-BLAST family)
- Transceiver architecture for slow fading (D-BLAST)
- More on performance in slow fading in next section.

Transmitter and Receiver CSI

- Can decompose the MIMO channel into a bunch of orthogonal sub-channels.
- Can allocate power and rate to each sub-channel according to waterfilling



Receiver CSI Only

The channel matrix $\mathbf{H}$ and its singular values λ_i^2's are random and unknown to the transmitter.

Can only fix a power allocation strategy. Equal power allocation on all antennas is natural and optimal in many cases.

Slow fading:

$$p_{\text{out}}(R) = \mathcal{P}\left\{\sum_i \log\left(1 + \frac{P}{n_t}\frac{\lambda_i^2}{N_0}\right) < R\right\}$$

Fast fading:

$$C = \mathcal{E}\left[\sum_i \log\left(1 + \frac{P}{n_t}\frac{\lambda_i^2}{N_0}\right)\right]$$

Capacity

Can write:

$$\sum_i \log\left(1+\frac{P}{n_t}\frac{\lambda_i^2}{N_0}\right) = \log\det\left(\mathbf{I}+\frac{\mathsf{SNR}}{n_t}\mathbf{H}\mathbf{H}^*\right)$$

Slow fading:

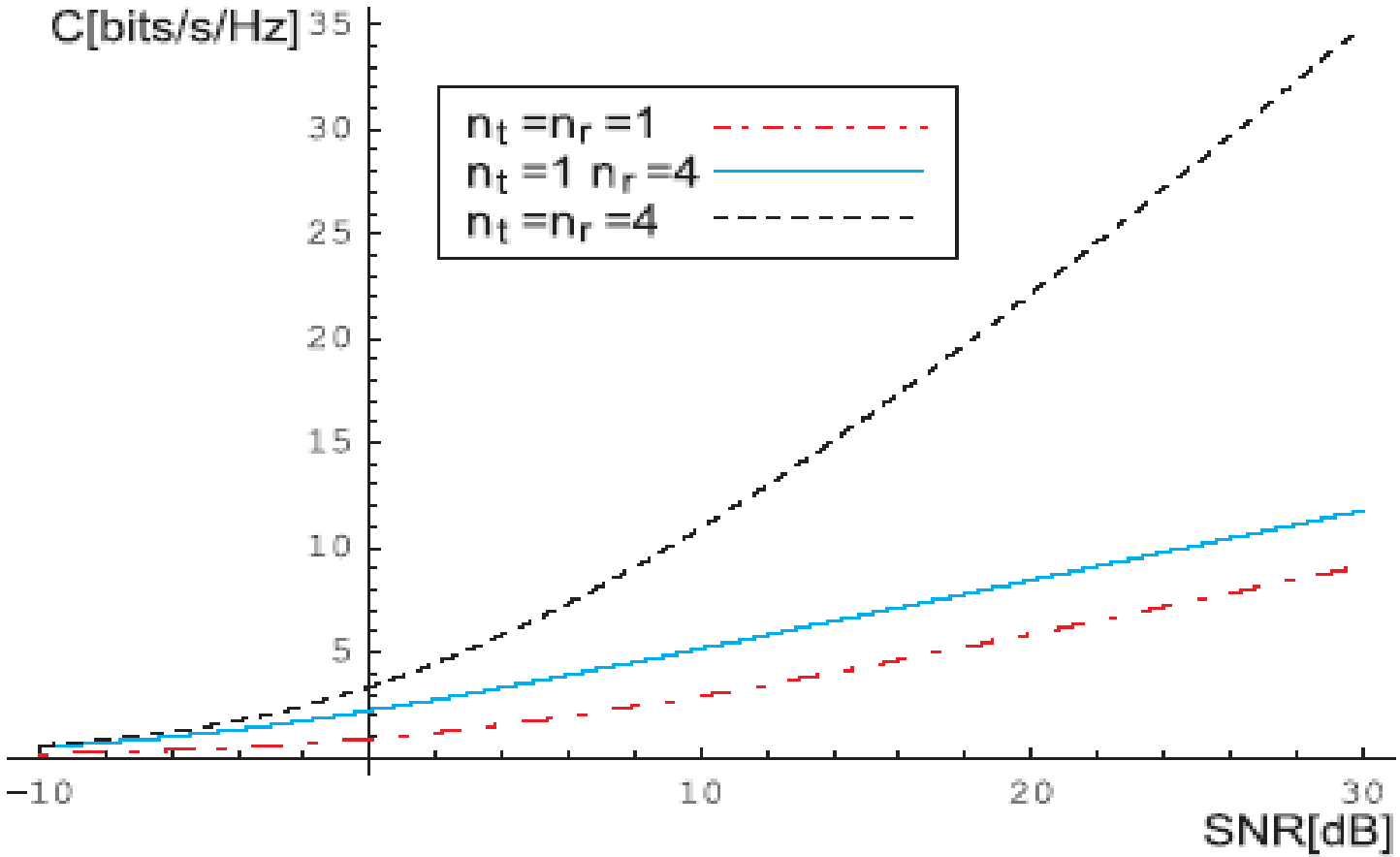
$$p_{\text{out}}(R) = \mathcal{P}\left\{\log\det\left(\mathbf{I}+\frac{\mathsf{SNR}}{n_t}\mathbf{H}\mathbf{H}^*\right) < R\right\}$$

Fast fading:

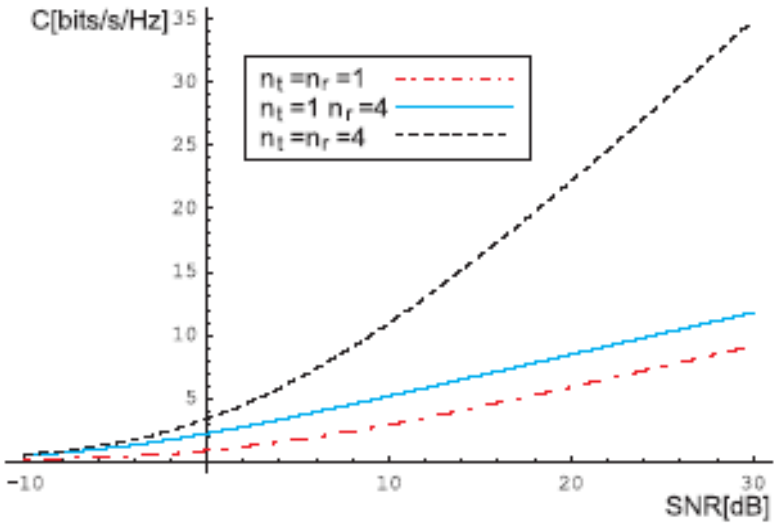
$$C = \mathcal{E}\left[\log\det\left(\mathbf{I}+\frac{\mathsf{SNR}}{n_t}\mathbf{H}\mathbf{H}^*\right)\right]$$

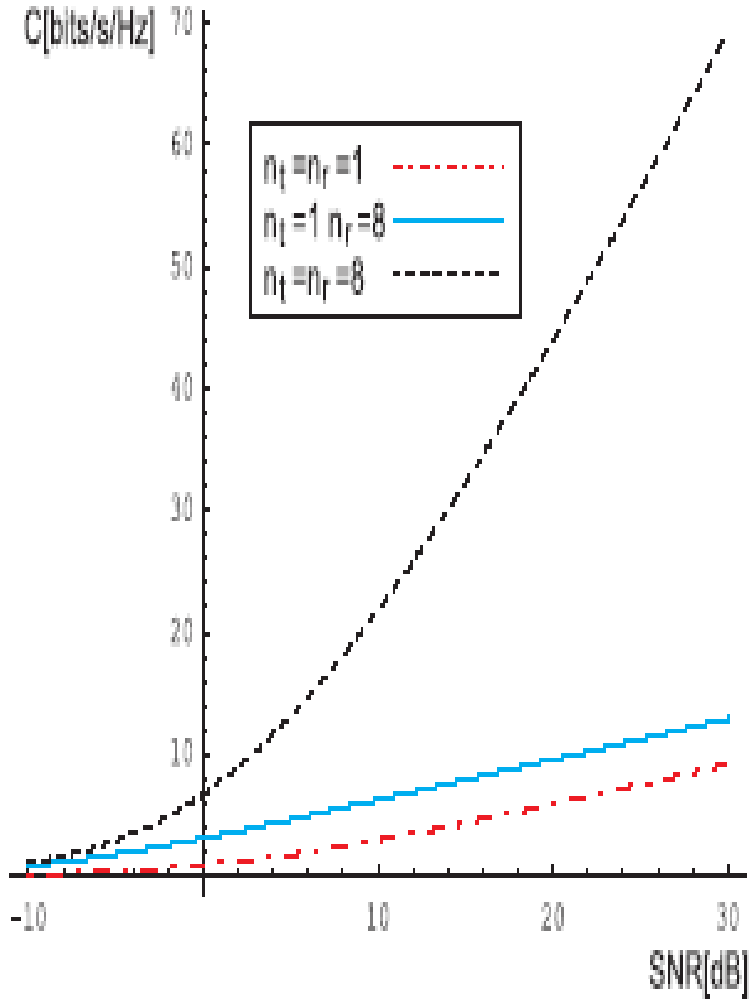
It is non-trivial to come up with capacity-achieving architectures.

Fast Fading Capacity

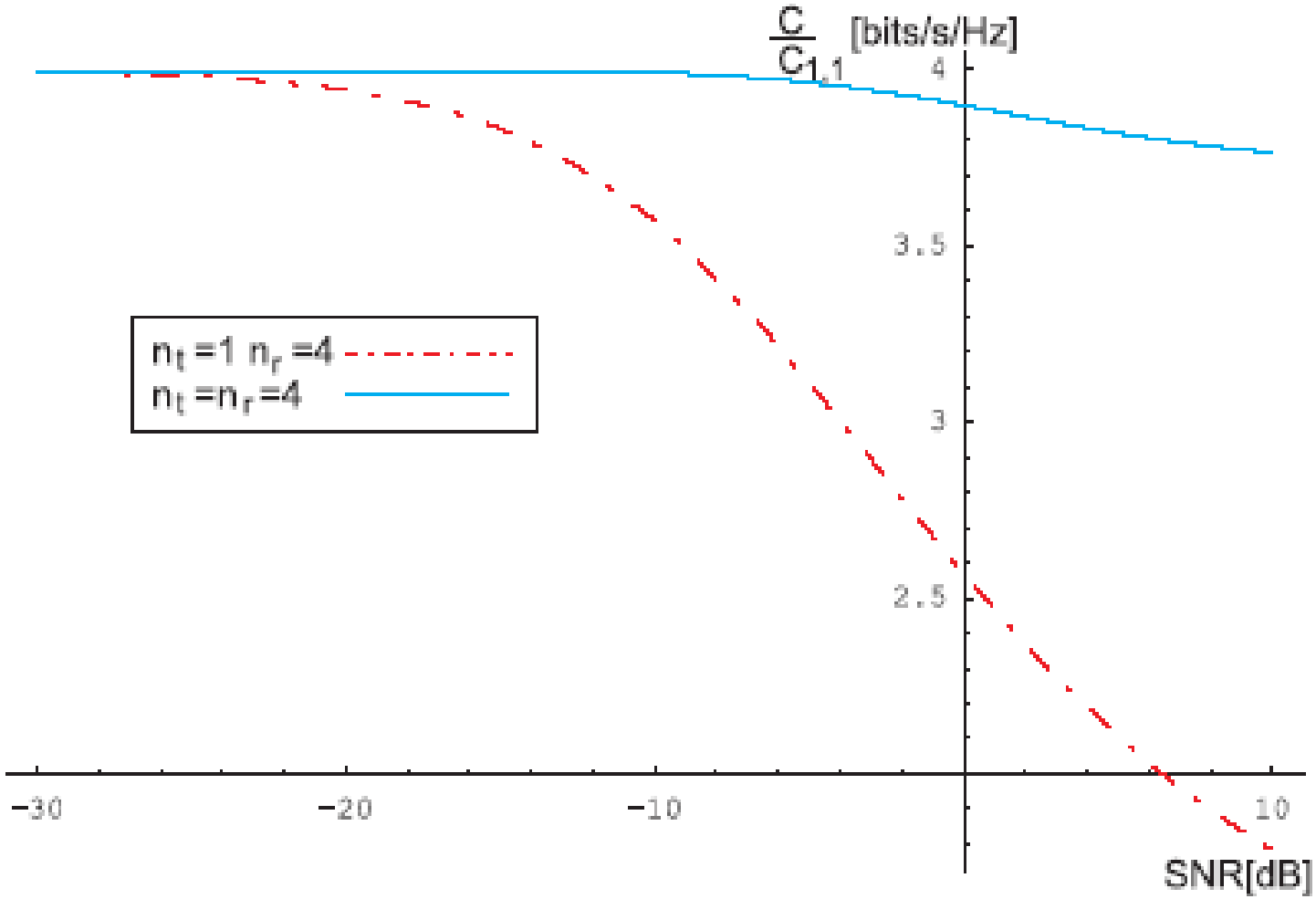


d.o.f. $\min(n_r, n_r)$ determines the high SNR slope.





Fast Fading Capacity: Low SNR



Nature of MIMO Performance Gain

At high SNR (d.o.f. limited) : min(n_t, n_r)-fold d.o.f. gain. MIMO is crucial.

At low SNR (power limited): n_r-fold power gain. Only need multiple receive antennas.

Question:

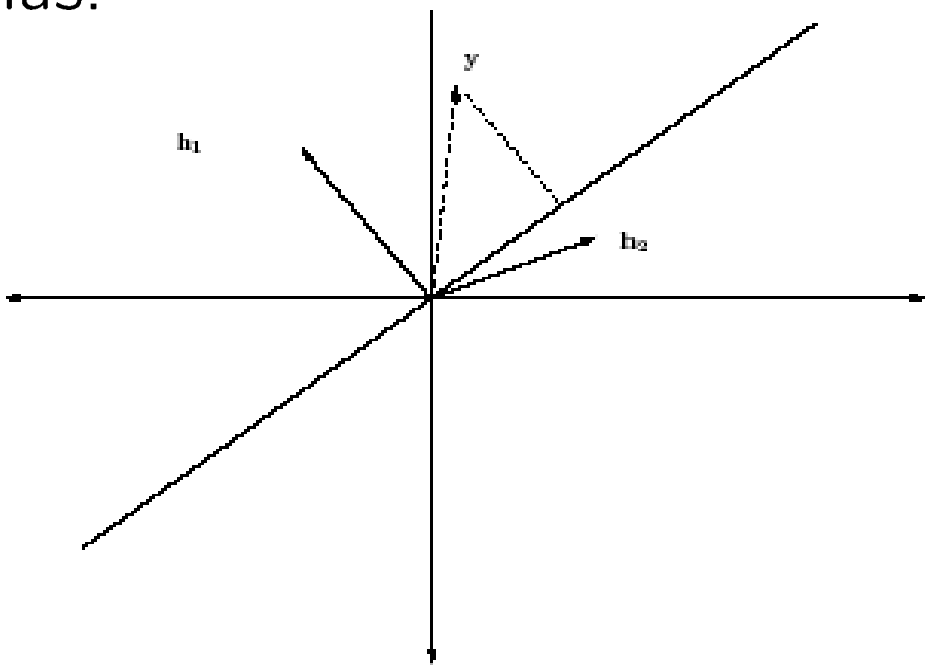
How to get the dof gain even when transmitter does not know the channel?

Interference Nulling

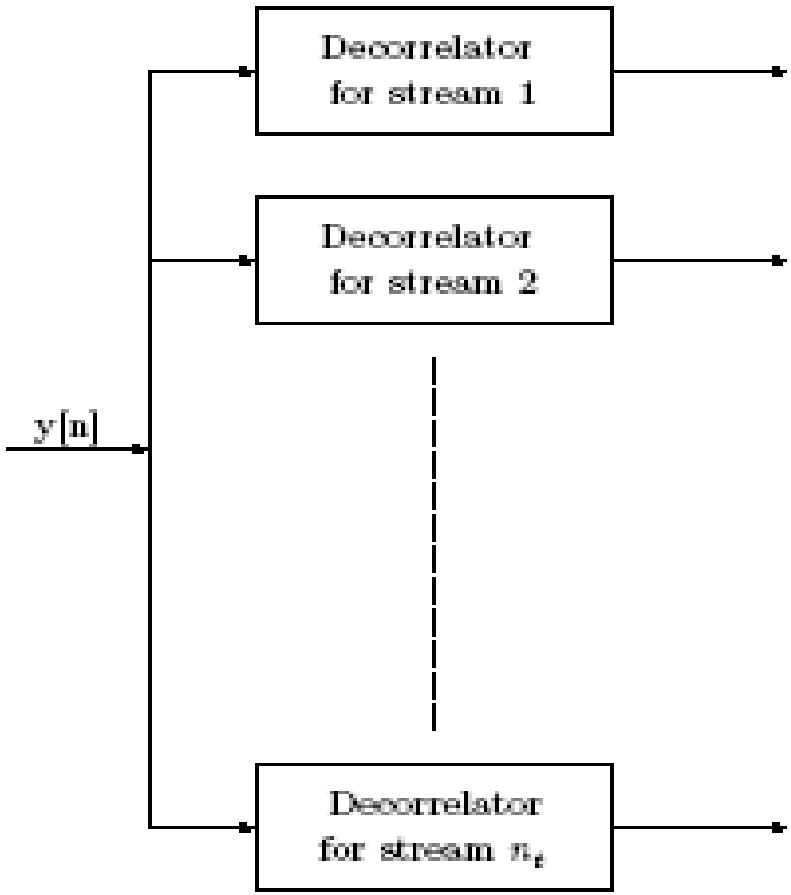
Focusing on Tx antenna 1:

$$\mathbf{y} = \mathbf{h}_1 x_1 + \sum_{i \neq 1} \mathbf{h}_i x_i + \mathbf{w}$$

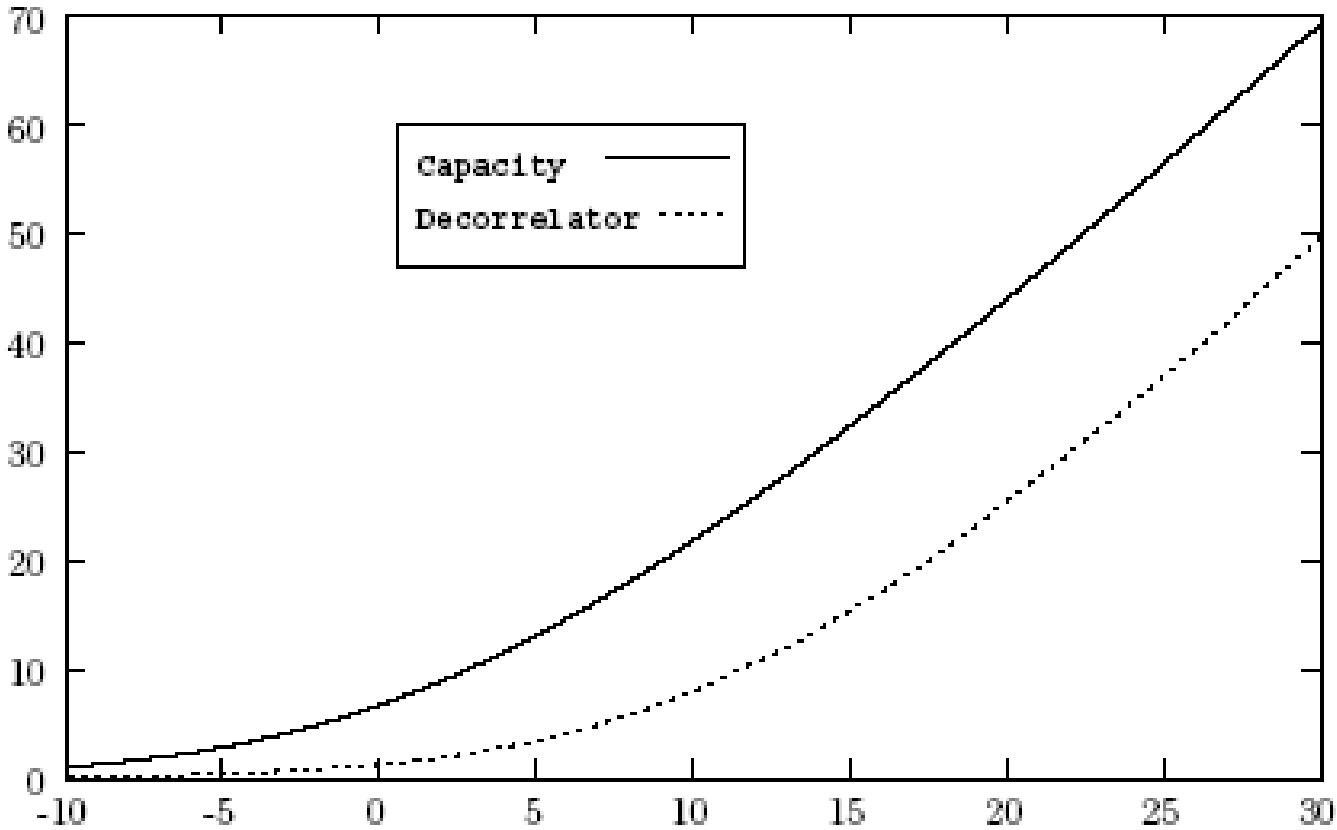
Simple strategy: null out the interference from other antennas.



Receiver Architecture I: Bank of Decorrelators



Bank of Decorrelators: Performance



$n_t = n_r = 8$, i.i.d. Rayleigh

Performance Gap of Decorrelator

Achieves the full d.of. $\min(n_t, n_r)$ of the MIMO channel.

But:

There is still a substantial constant gap at high SNR.

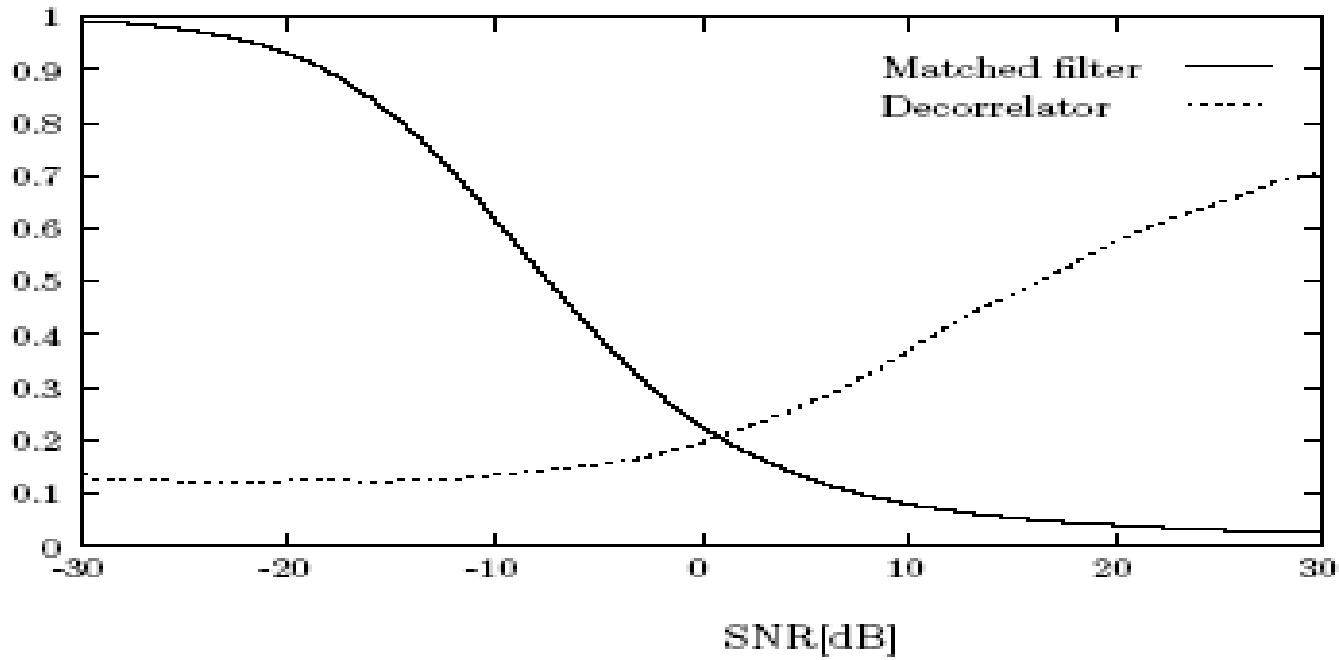
At moderate and low SNR, performance sucks.

Interference Nulling vs Match Filtering

$$\mathbf{y} = \mathbf{h}_1 x_1 + \sum_{i \neq 1} \mathbf{h}_i x_i + \mathbf{w}$$

Interference nulling: remove all interference at the expense of reducing the SNR

Match filtering: projecting onto $\mathbf{h}_1$ to maximize the SNR but SINR may be bad.



Optimal Linear Filter:MMSE

$$\mathbf{y} = \mathbf{h}_1 x_1 + \sum_{i \neq 1} \mathbf{h}_i x_i + \mathbf{w}$$

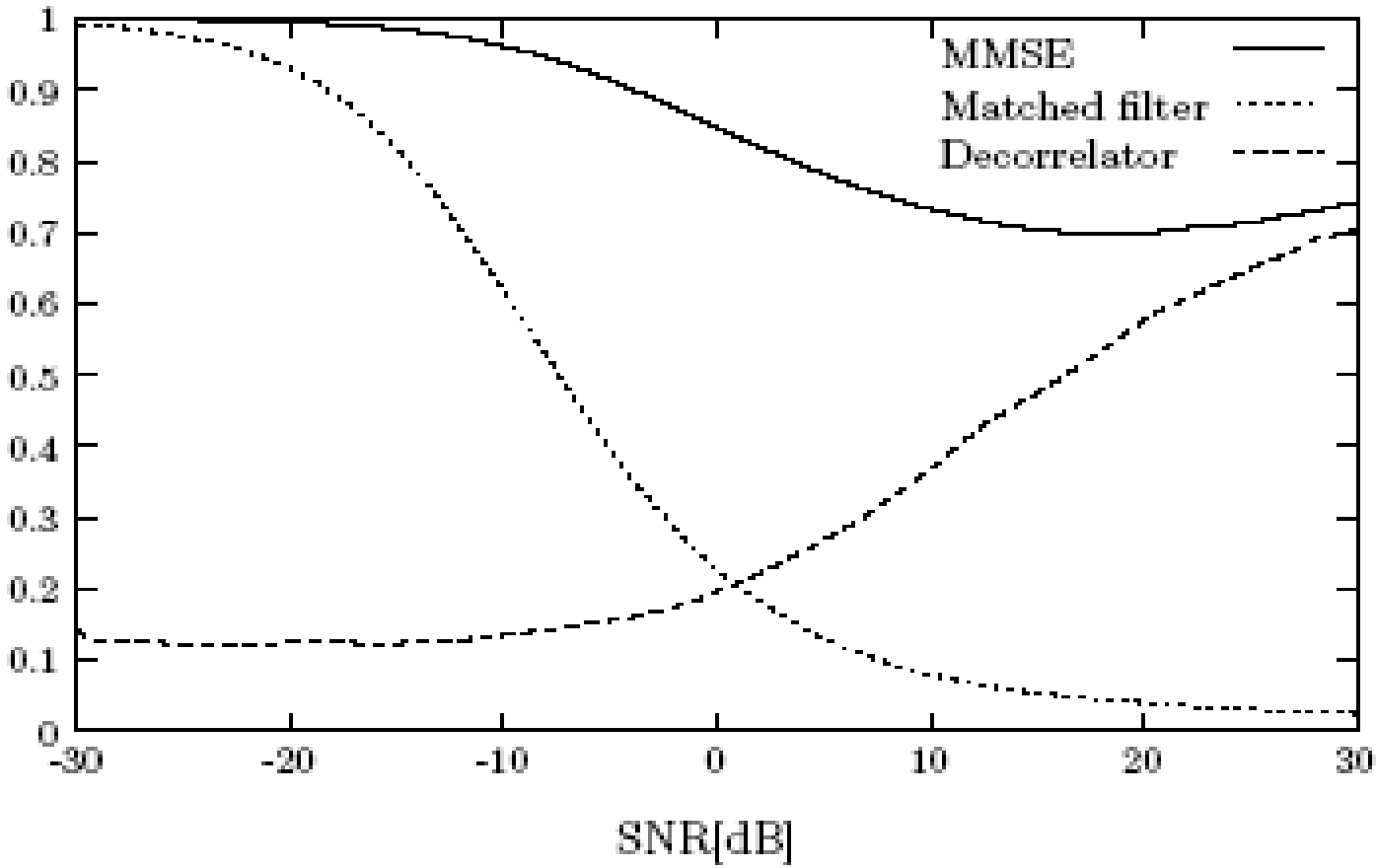
Seek a linear filter that maximizes the output SINR at all SNR.

Offers the optimal compromise between nulling and match filtering

It whitens the interference first and then match filter.

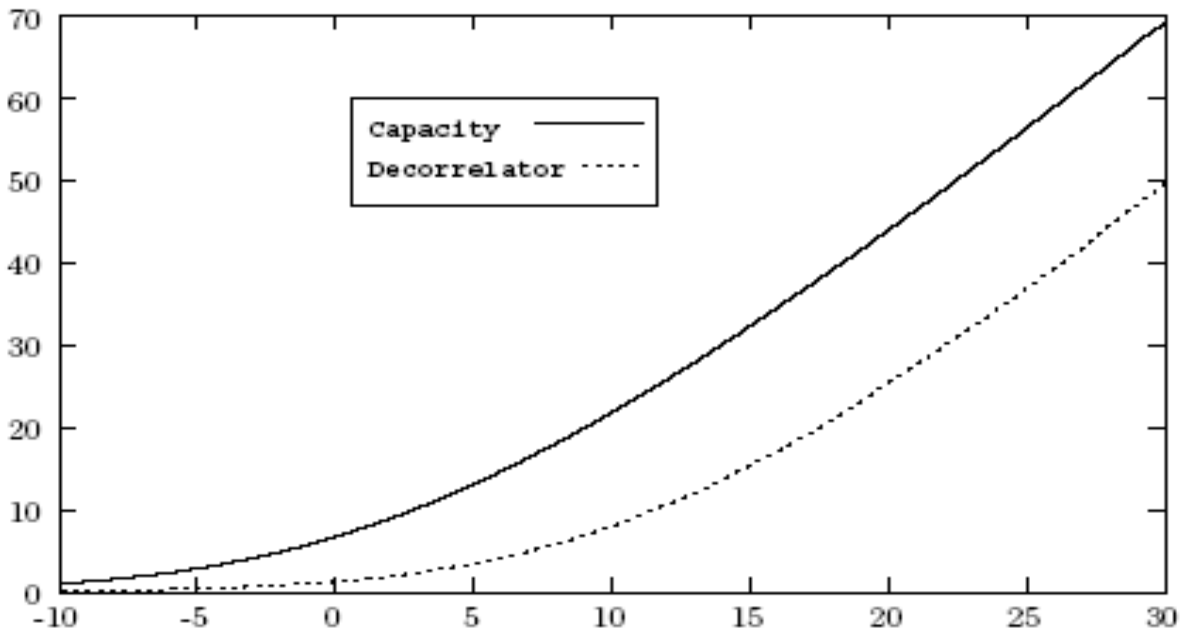
This is the linear MMSE filter.

Linear MMSE: Performance

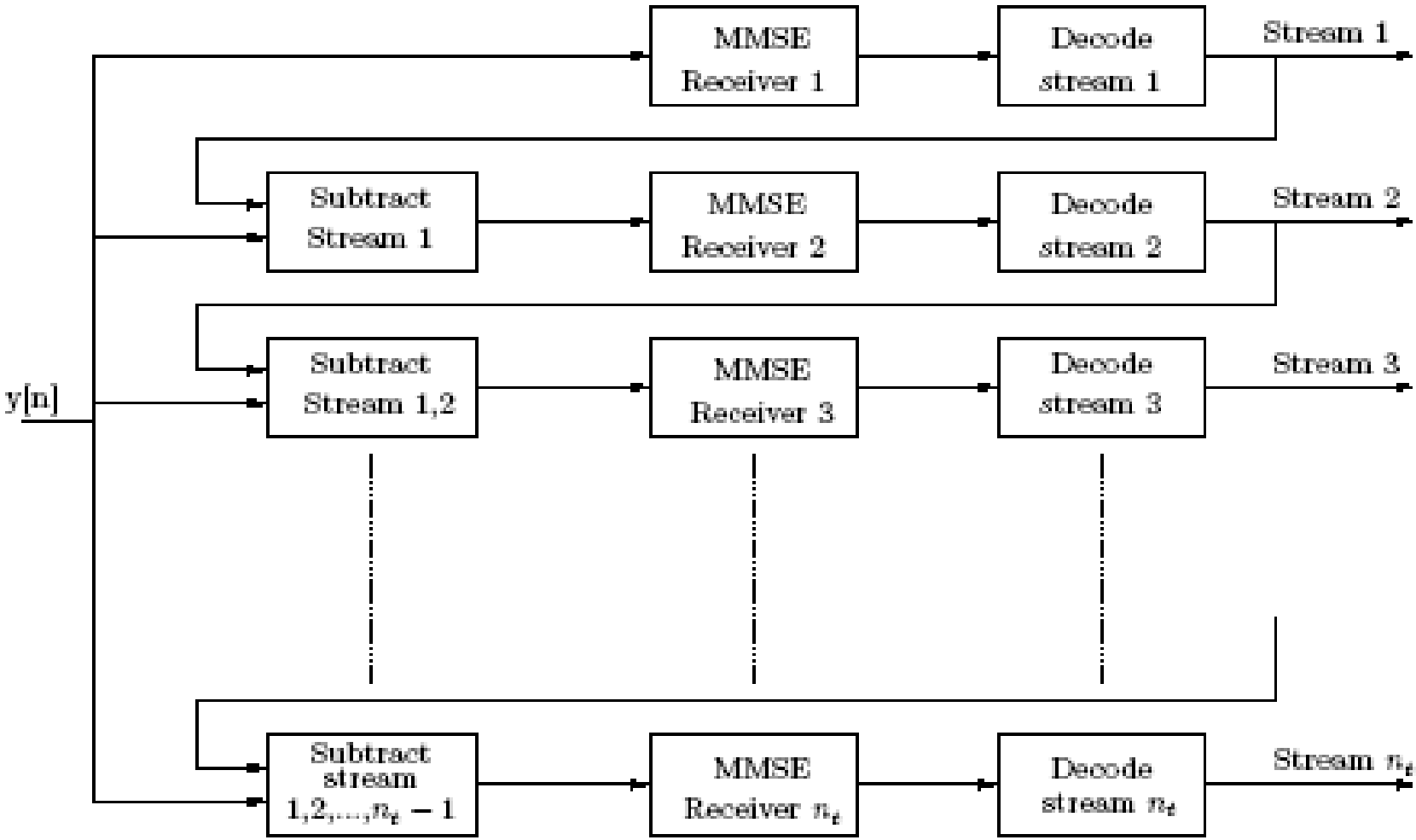


Gap at High SNR

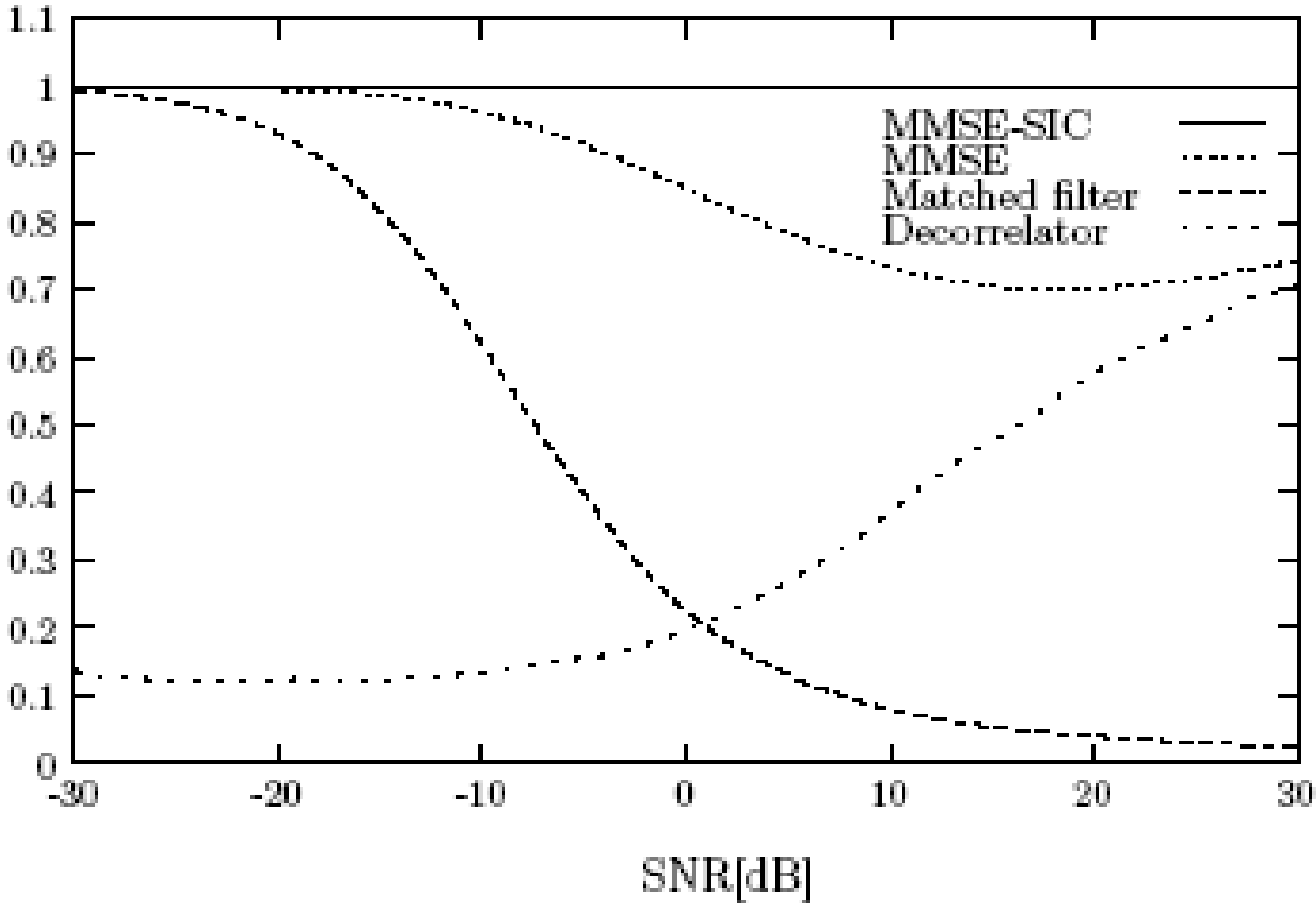
- MMSE improves the performance of decorrelator at moderate and low SNR.
- Does not remove the gap in performance at high SNR
- To remove that gap we have to go to non-linear receivers.



Successive Interference Cancellation



MMSE-SIC Achieves MIMO Capacity



Optimality of MMSE-SIC

Given a fixed channel $\mathbf{H}$,

$$\log\det\left(\mathbf{I}+\frac{\mathsf{SNR}}{n_t}\mathbf{H}\mathbf{H}^*\right)=\sum_{i=1}^{n_t}\log(1+\mathsf{SNR}_{\mathsf{mmse},i})$$

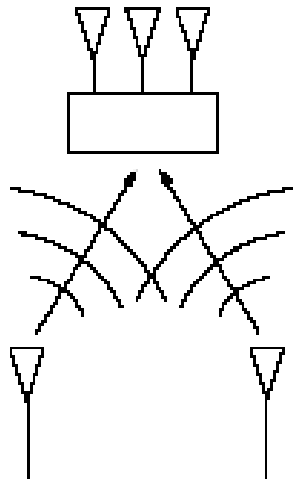
Why is MMSE-SIC optimal?

MMSE is information lossless at each stage.

The SIC architecture implements the chain rule of information.

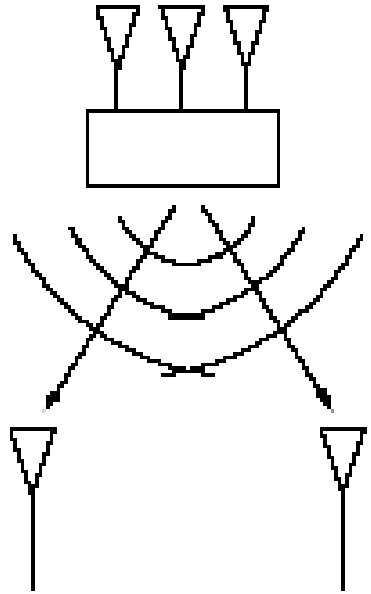
Uplink Architectures

- So far we have considered point-to-point communication.
- But since we are sending independent streams from each transmit antennas, we can use the receiver structures for the uplink with multiple users.
- This is called space-division multiple access (SDMA)
- Several simultaneous users can be supported.
- Linear MMSE also called receive beamforming.

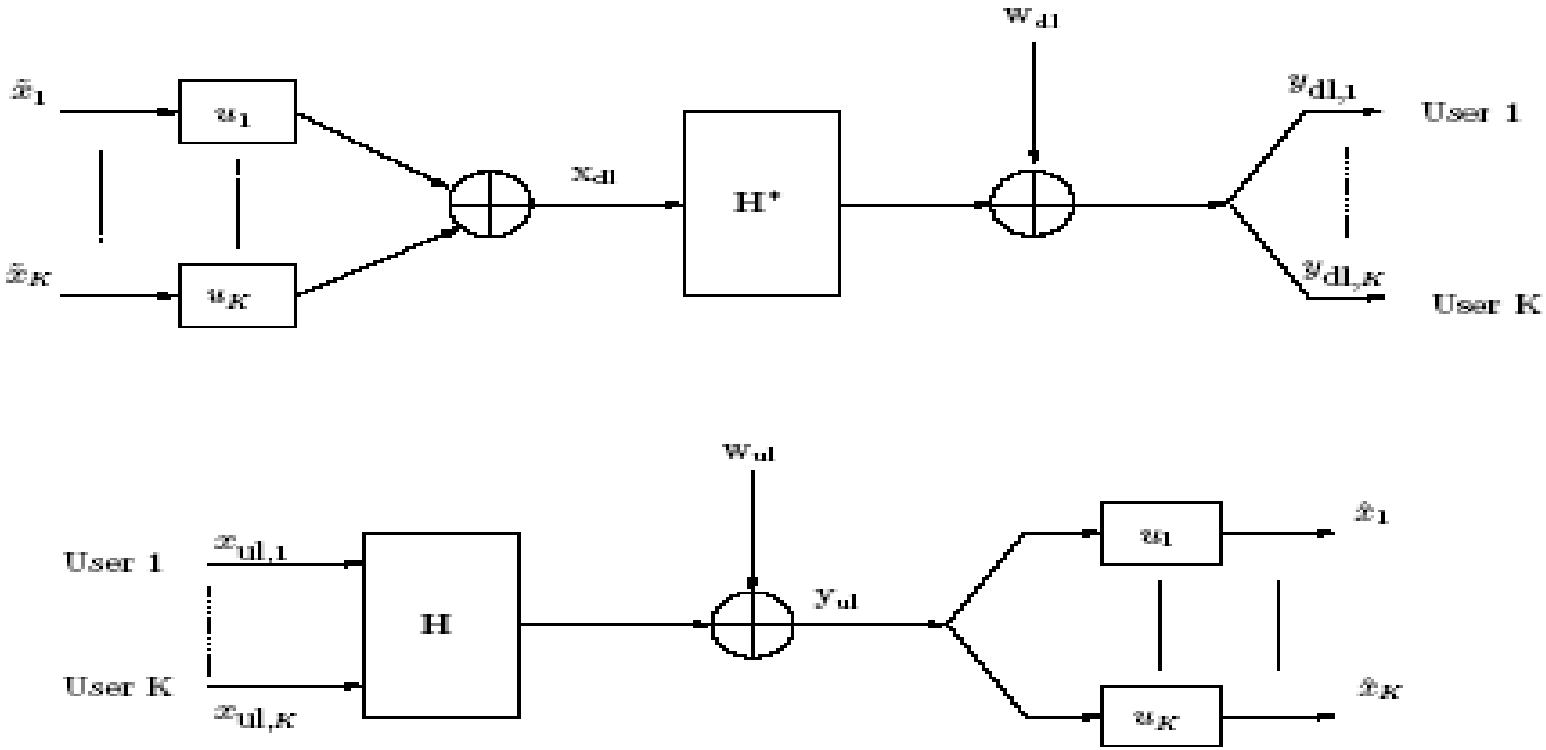


Downlink

- In the uplink, transmitters cannot cooperate, but receiver can jointly process the received signal at all the antennas.
- In the downlink, it is the receivers that cannot cooperate.
- If the transmitter does not track the channel, cannot do SDMA on the downlink.
- If it does, can use techniques reciprocal to the uplink.



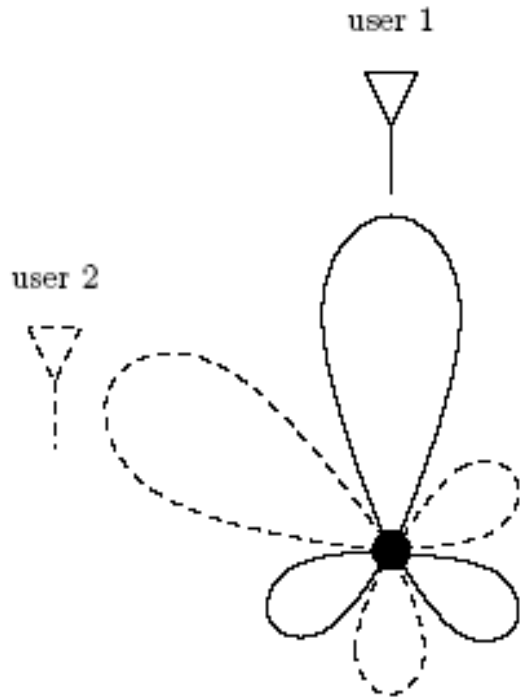
Uplink-Downlink Reciprocity



Downlink Transmit Beamforming

Can use transmit filter for user 1 that nulls out interference to other users.

More generally, can optimally balance the energy transferred to the users and the inter-user interference (downlink MMSE)



Example: ArrayComm

- SDMA overlay on Japan's PHS system, also a newer data system (iBurst)
- Up to 12 antennas at BS, with up to 4 users simultaneously in SDMA.
- Antennas also used to null out inter-cell interference, increasing frequency-reuse factor (from 1/8 to 1 in PHS)
- System is TDD.
- Channel is measured from pilot in uplink, and used in downlink transmit beamforming.

Uplink-Downlink Duality

- Linear receive beamforming strategies for the uplink map to linear transmit beamforming strategies in the downlink.
- But in the uplink we can improve performance by doing successive interference cancellation at the receiver
- Is there a dual to this strategy in the downlink?

Transmit Precoding

- In downlink transmit beamforming, signals for different users interfere with each other.
- A user is in general not able to decode information for other users to cancel them off.
- However, the transmitter knows the information to be transmitted for every user and can precode to cancel at the transmitter.

Symbol-by-Symbol Precoding

$$y = x + s + w$$

Interference s is known at the transmitter, not at the receiver.

Applications:

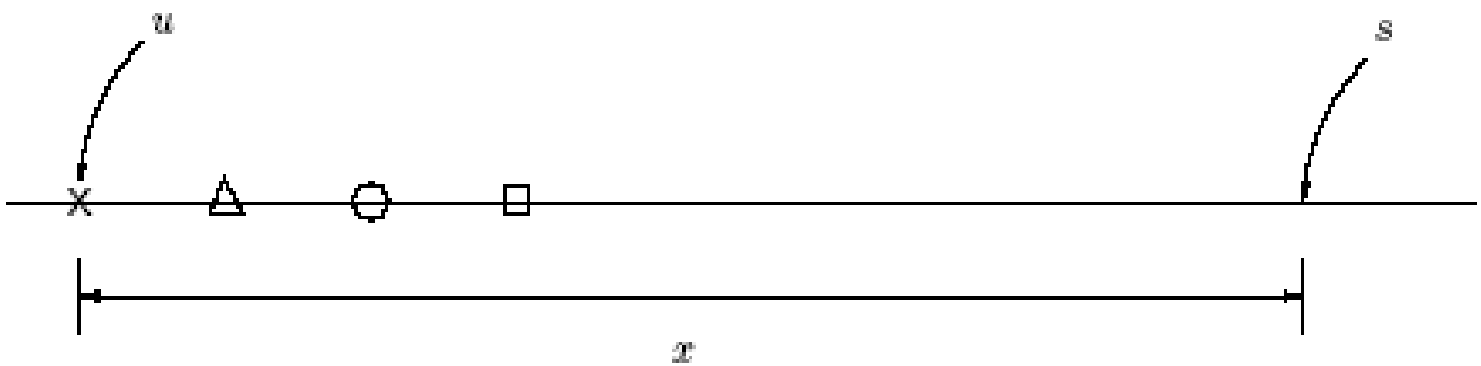
- downlink: s is signal for another user.
- information embedding: s is the host signal
- ISI precoding: s is the intersymbol interference

Naïve Pre-cancellation Strategy

Want to send point u in a 4-PAM constellation.

Transmit $x = u - s$ to pre-cancel the effect of s.

But this is very power inefficient if s is large.

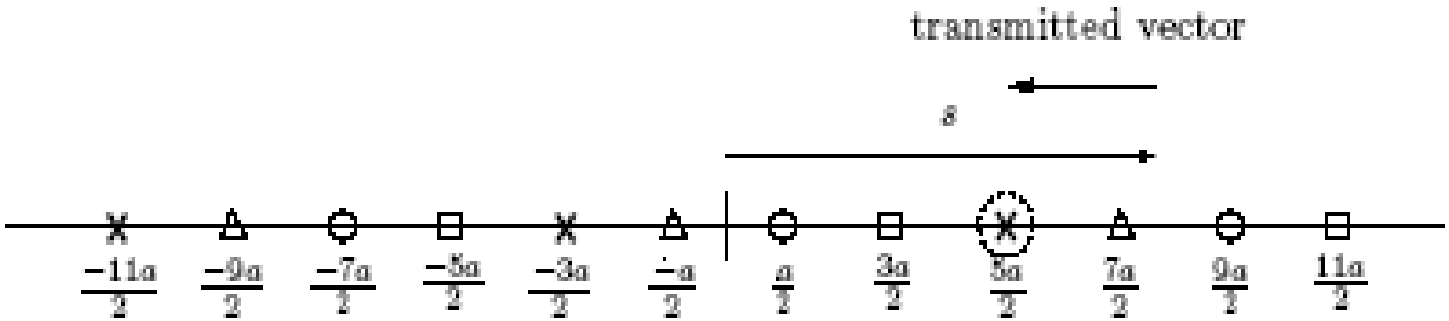


Tomlinson-Harashima Precoding

Replicate the PAM constellation to tile the whole real line.

Represent information u by an equivalence class of constellation points instead of a single point.

Given u and u , find the point in its equivalence class closest to s and transmit the difference.



Writing on Dirty Paper

- Can extend this idea to block precoding.
- Problem is to design codes which are simultaneously good source codes (vector quantizers) as well as good channel codes.
- Very active research area.
- Somewhat surprising, information theory guarantees that one can get to the capacity of the AWGN channel with the interference completely removed.

6. Diversity-Multiplexing Tradeoff

Slow Fading MIMO Channel

- So far we have emphasized the spatial multiplexing aspect of MIMO channels.
- But we also learnt that in slow fading scenario, diversity is an important thing.
- How do the two aspects interact?
- It turns out that you can get both in a slow fading channel but there is a fundamental tradeoff.
- We characterize the optimal tradeoff and find schemes that approach the optimal tradeoff.

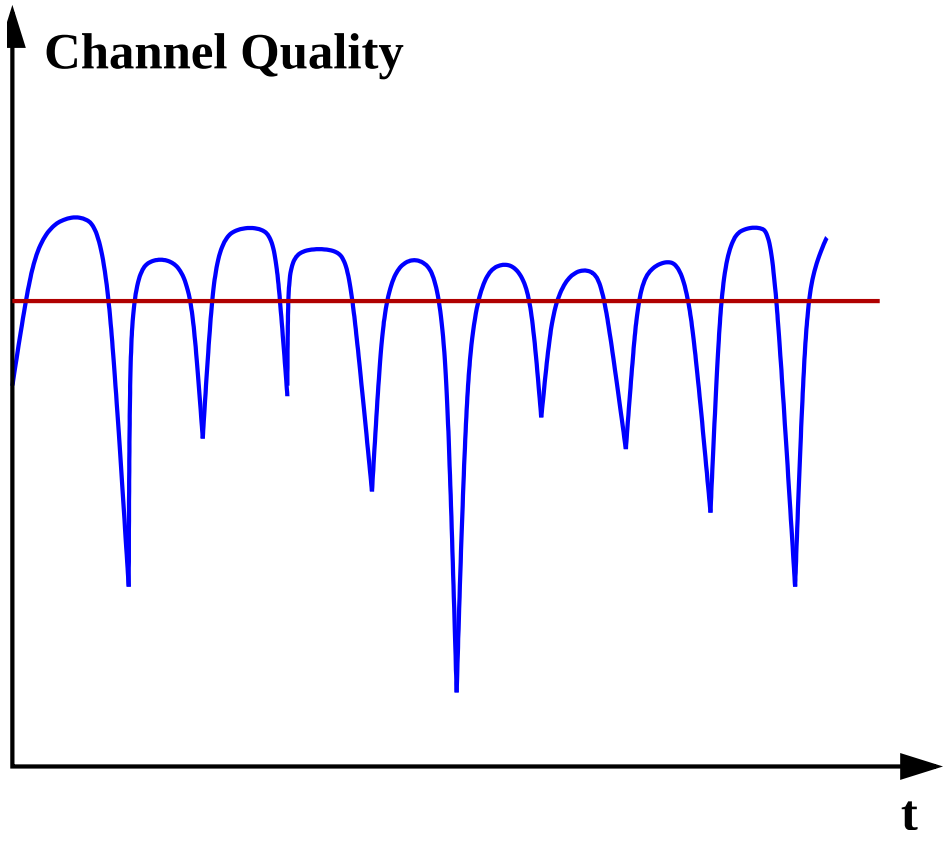
Diversity and Freedom

Two fundamental resources of a MIMO fading channel:

diversity

degrees of freedom

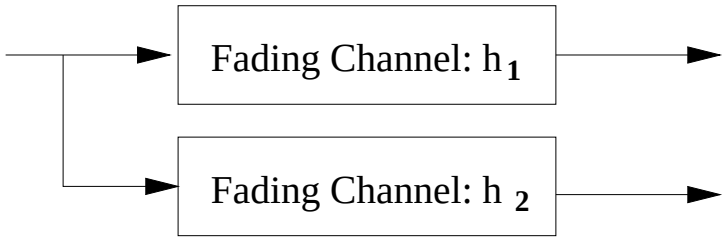
Diversity



A channel with more diversity has smaller probability in deep fades.

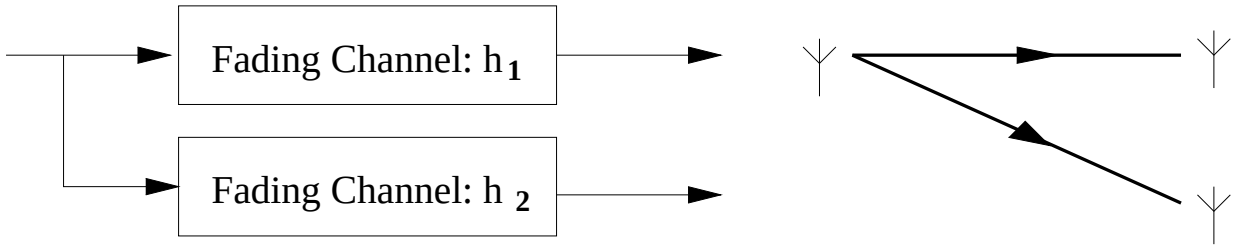
Diversity

Diversity



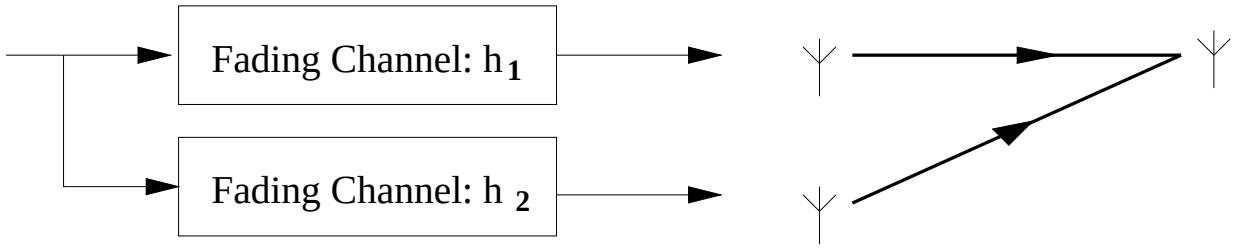
- Additional independent fading channels increase diversity.

Diversity



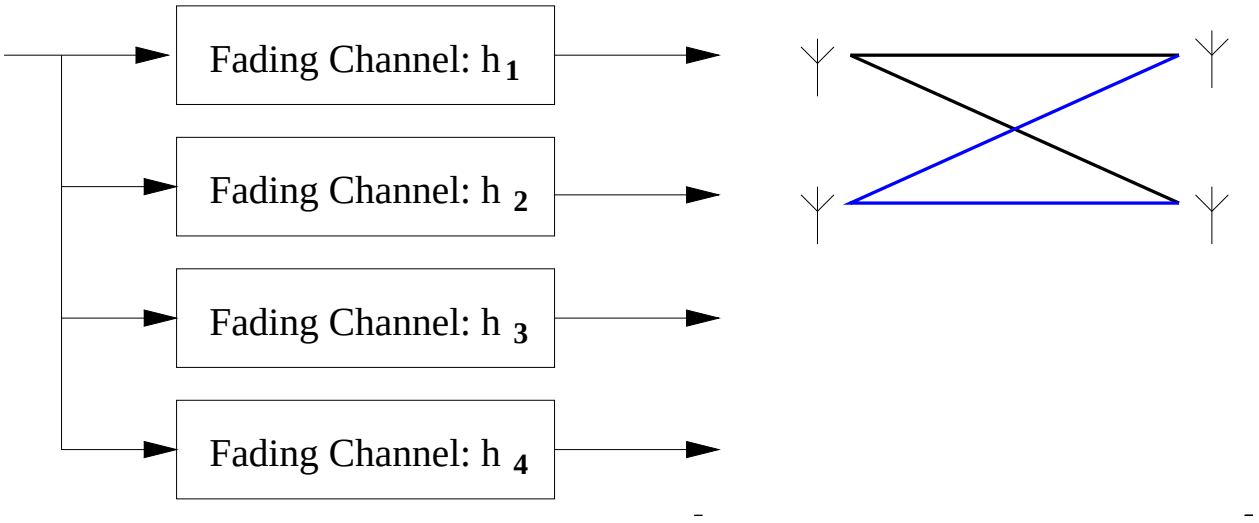
- Additional independent fading channels increase diversity.
- Spatial diversity

Diversity



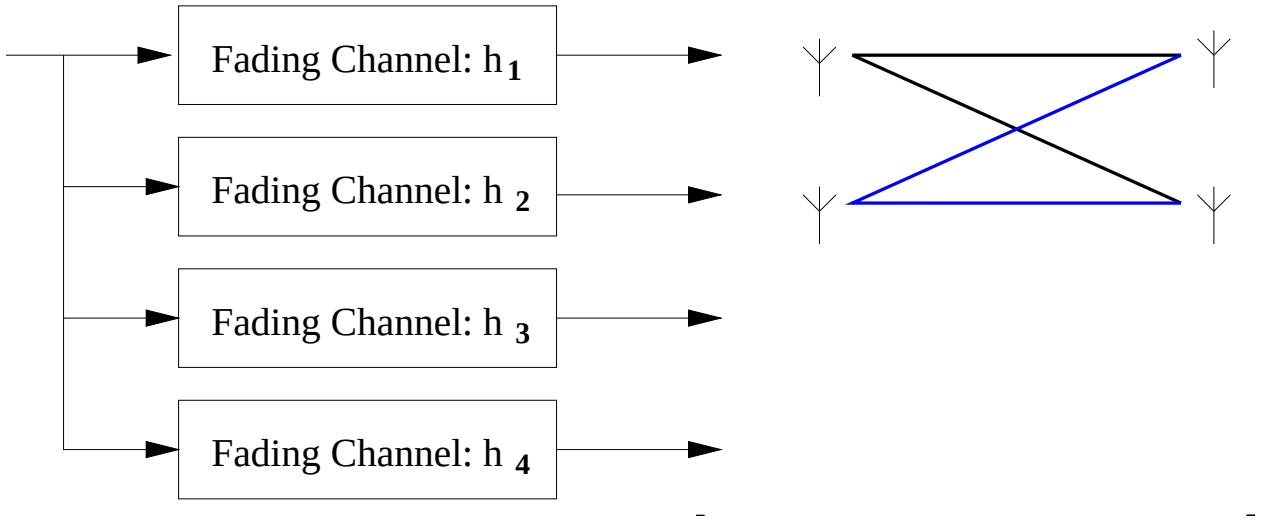
- Additional independent fading channels increase diversity.
- Spatial diversity: receive, transmit

Diversity



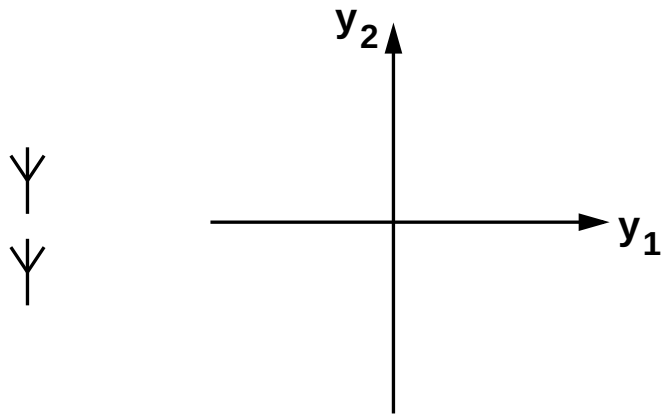
- Additional independent fading channels increase diversity.
- Spatial diversity: receive, transmit or both.

Diversity

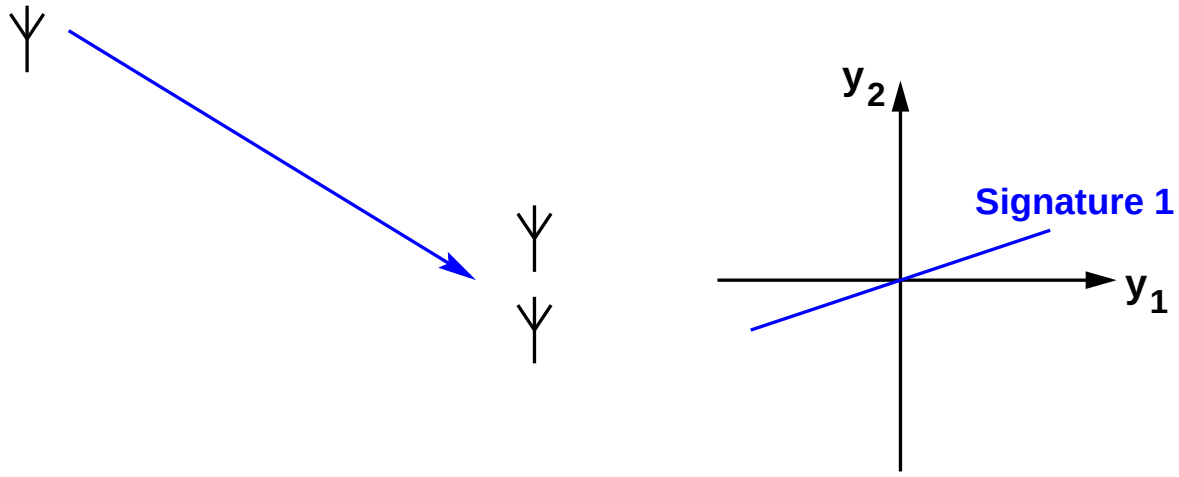


- Additional independent fading channels increase diversity.
- Spatial diversity: receive, transmit or both.
- For a m by n channel, diversity is mn.

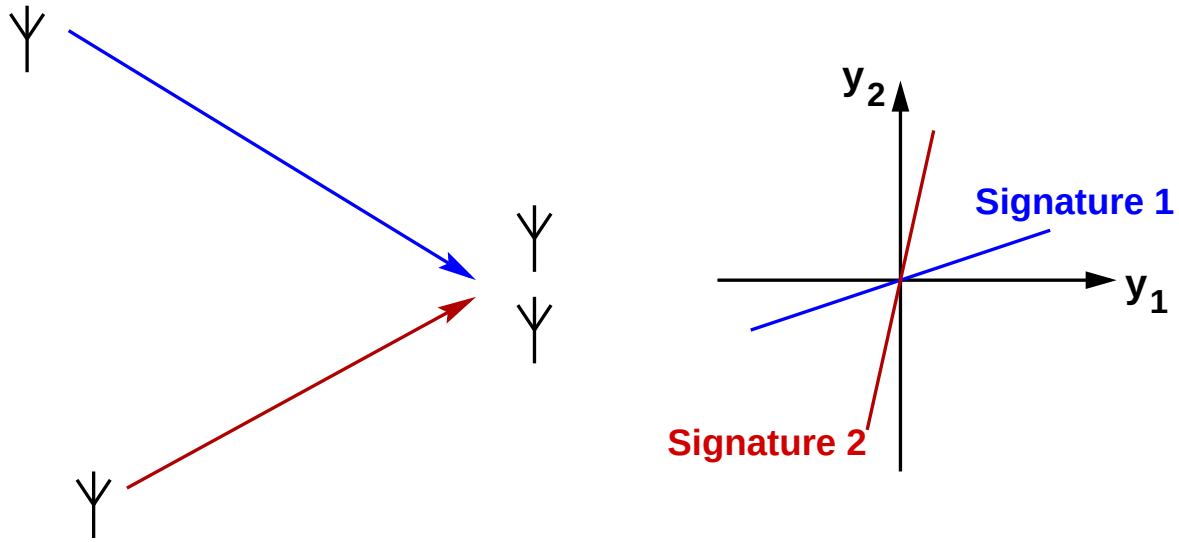
Degrees of Freedom



Degrees of Freedom

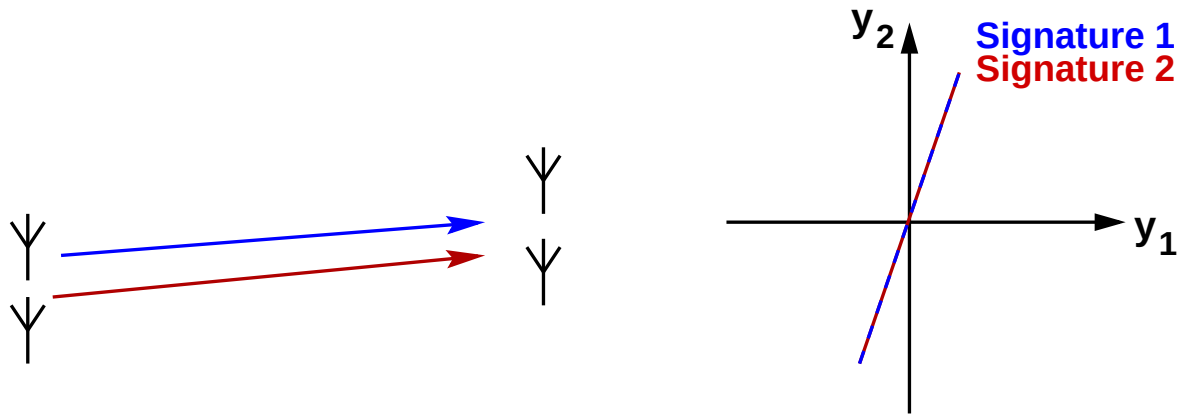


Degrees of Freedom



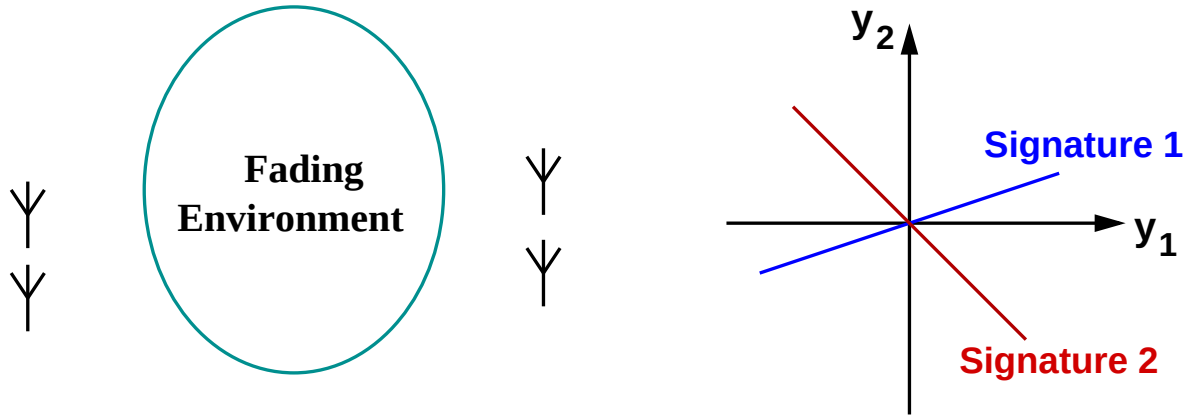
Signals arrive in multiple directions provide multiple degrees of freedom for communication.

Degrees of Freedom



Signals arrive in multiple directions provide multiple degrees of freedom for communication.

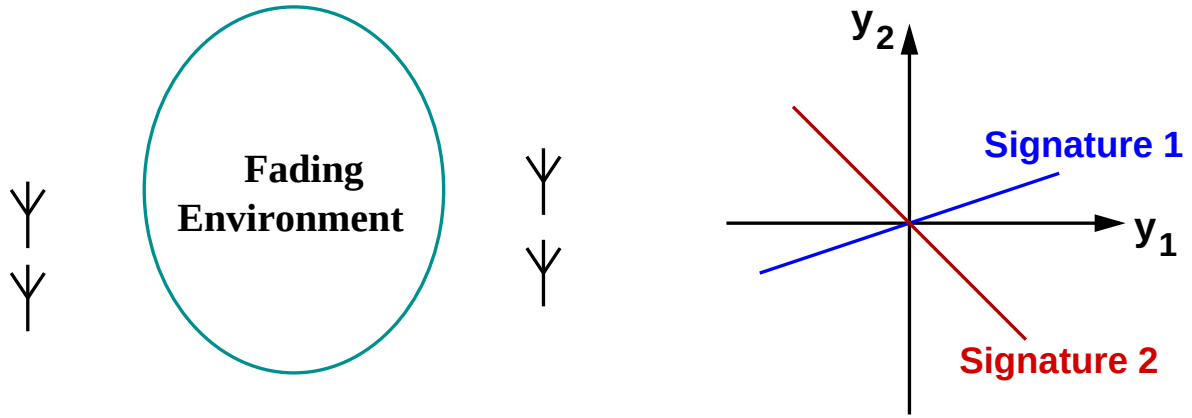
Degrees of Freedom



Signals arrive in multiple directions provide multiple degrees of freedom for communication.

Same effect can be obtained via scattering even when antennas are close together.

Degrees of Freedom



Signals arrive in multiple directions provide multiple degrees of freedom for communication.

Same effect can be obtained via scattering even when antennas are close together.

In a m by n channel with rich scattering, there are $\min\{m, n\}$ degrees of freedom.

Diversity and Freedom

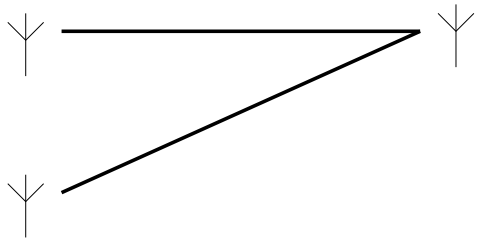
In a MIMO channel with rich scattering:

maximum diversity $= mn$

degrees of freedom $= \min\{m, n\}$

The name of the game in space-time coding is to design schemes which exploit as much of both these resources as possible.

Space-Time Code Examples: 2×1 Channel



Repetition Scheme:

$$X = \begin{bmatrix} x_1 & 0 \\ 0 & x_1 \end{bmatrix}$$

time (→), space (↓)

diversity: 2

data rate: 1/2 sym/s/Hz

Alamouti Scheme:

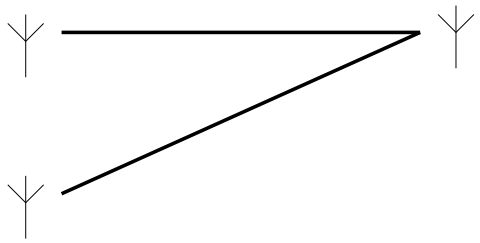
$$X = \begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}$$

time (→), space (↓)

diversity: 2

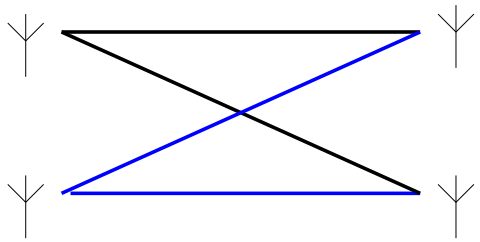
data rate: 1 sym/s/Hz

Performance Summary: 2×1 Channel



	Diversity gain	Degrees of freedom utilized /s/Hz
Repetition	2	$1/2$
Alamouti	2	1
channel itself	2	1

Space-Time Code Examples: 2×2 Channel



Repetition Scheme:

$$\mathbf{X} = \begin{bmatrix} x_1 & 0 \\ 0 & x_1 \end{bmatrix}$$

time

space

diversity gain : 4

data rate: 1/2 sym/s/Hz

Alamouti Scheme:

$$\mathbf{X} = \begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}$$

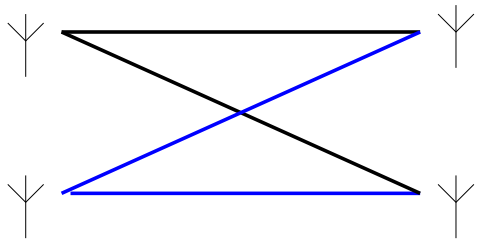
time

space

diversity gain : 4

data rate: 1 sym/s/Hz

Space-Time Code Examples: 2×2 Channel



Repetition Scheme:

$$X = \begin{bmatrix} x_1 & 0 \\ 0 & x_1 \end{bmatrix}$$

time (horizontal) / space (vertical)

diversity: 4

data rate: 1/2 sym/s/Hz

Alamouti Scheme:

$$X = \begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}$$

time (horizontal) / space (vertical)

diversity: 4

data rate: 1 sym/s/Hz

But the 2×2 channel has 2 degrees of freedom!

V-BLAST with Nulling

Send two independent uncoded streams over the two transmit antennas.

Demodulate each stream by nulling out the other stream.

Data rate: 2 sym/s/Hz

Diversity: 1

Winters, Salz and Gitlins 93:

Nulling out k interferers using n receive antennas yields a diversity gain of $n - k$.

Performance Summary: 2×2 Channel

	Diversity gain	d.o.f. utilized /s/Hz
Repetition	4	$1/2$
Alamouti	4	1
V-Blast with nulling	1	2
channel itself	4	2

Questions:

- Alaomuti is clearly better than repetition, but how can it be compared to V-Blast?

Performance Summary: 2×2 Channel

	Diversity gain	d.o.f. utilized /s/Hz
Repetition	4	$1/2$
Alamouti	4	1
V-Blast with nulling	1	2
channel itself	4	2

Questions:

- Alaomuti is clearly better than repetition, but how can it be compared to V-Blast?
- How does one quantify the "optimal" performance achievable by any scheme?

Performance Summary: 2×2 Channel

	Diversity gain	d.o.f. utilized /s/Hz
Repetition	4	$1/2$
Alamouti	4	1
V-Blast with nulling	1	2
channel itself	4	2

Questions:

- Alaomuti is clearly better than repetition, but how can it be compared to V-Blast?
- How does one quantify the "optimal" performance achievable by any scheme?
- We need to make the notions of "fiversity gain" and "d.o.f. utilized" precise and enrich them.

Classical Diversity Gain

Motivation: PAM

$$\mathbf{y} = \mathbf{h}\mathbf{x} + \mathbf{w} \qquad P_e \approx P(\|\mathbf{h}\| \text{ is small }) \propto \text{SNR}^{-1}$$

$$\left.\begin{aligned} \mathbf{y}_1 &= \mathbf{h}_1\mathbf{x} + \mathbf{w}_1 \\ \mathbf{y}_2 &= \mathbf{h}_2\mathbf{x} + \mathbf{w}_2 \end{aligned}\right\} \qquad \begin{aligned} P_e &\approx P(\|\mathbf{h}_1\|, \|\mathbf{h}_2\| \text{ are both small}) \\ &\propto \text{SNR}^{-2} \end{aligned}$$

Classical Diversity Gain

Motivation: PAM

$$\mathbf{y} = \mathbf{h}\mathbf{x} + \mathbf{w} \qquad P_e \approx P(\|\mathbf{h}\| \text{ is small }) \propto \text{SNR}^{-1}$$

$$\left.\begin{aligned} \mathbf{y}_1 &= \mathbf{h}_1\mathbf{x} + \mathbf{w}_1 \\ \mathbf{y}_2 &= \mathbf{h}_2\mathbf{x} + \mathbf{w}_2 \end{aligned}\right\} \qquad \begin{aligned} P_e &\approx P(\|\mathbf{h}_1\|, \|\mathbf{h}_2\| \text{ are both small}) \\ &\propto \text{SNR}^{-2} \end{aligned}$$

General Definition

A space-time coding scheme achieves (classical) diversity gain $d_{\max}$, if

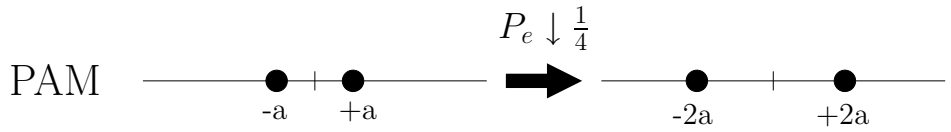
$$P_e(\text{SNR}) \sim \text{SNR}^{-d_{\max}}$$

for a fixed data rate.

i.e. error probability deceases by $2^{-d_{\max}}$ for every 3 dB increase in SNR, by $4^{-d_{\max}}$ for every 6dB increase, etc.

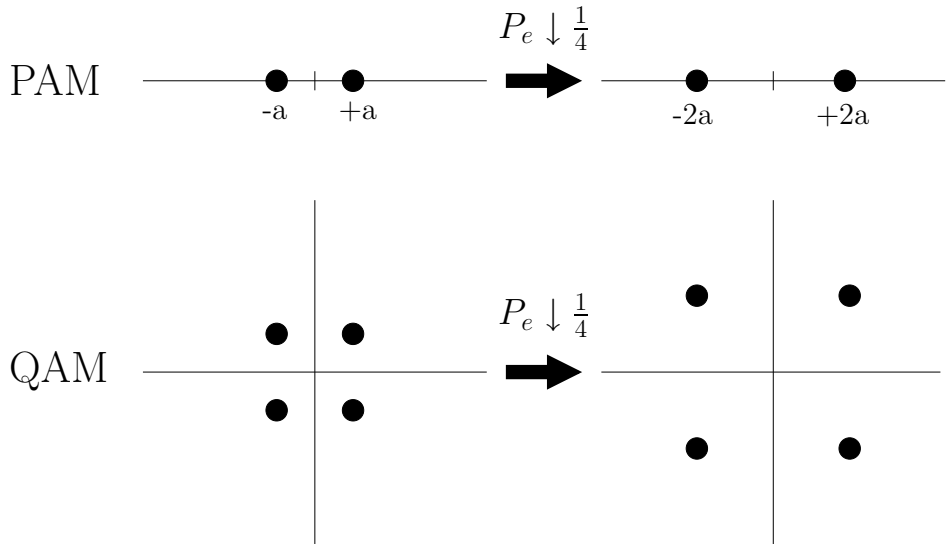
Example: PAM vs QAM in 1 by 1 Channel

Every 6 dB increase in SNR doubles the distance between constellation points for a given rate.



Example: PAM vs QAM in 1 by 1 Channel

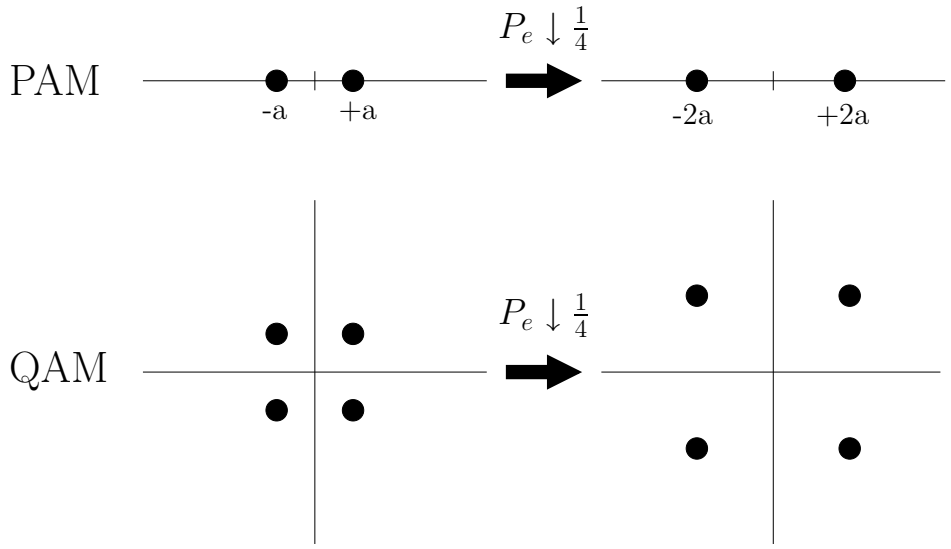
Every 6 dB increase in SNR doubles the distance between constellation points for a given rate.



Both PAM and QAM have the same (classical) diversity gain of 1.

Example: PAM vs QAM in 1 by 1 Channel

Every 6 dB increase in SNR doubles the distance between constellation points for a given rate.

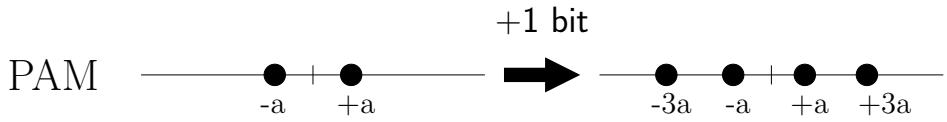


Both PAM and QAM have the same (classical) diversity gain of 1.

(classical) diversity gain does not say anything about the d.o.f. utilized by the scheme.

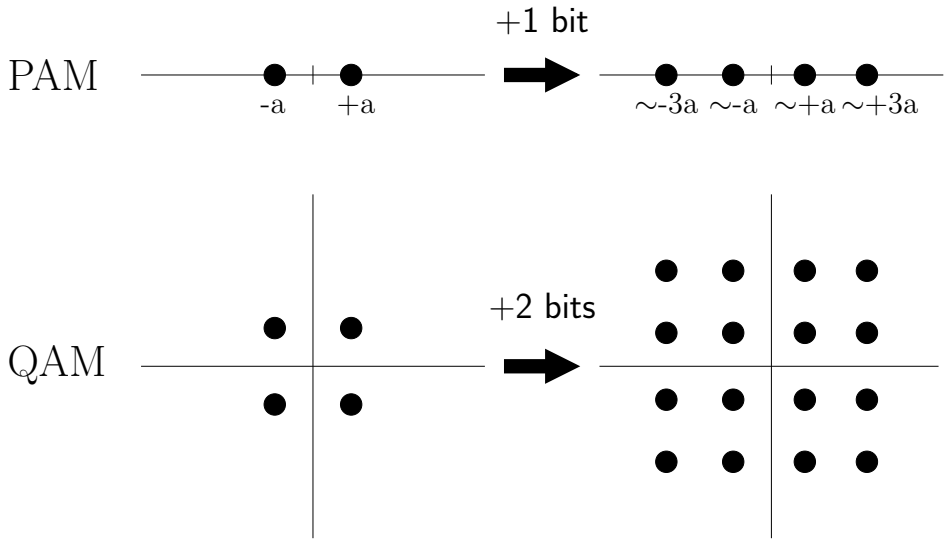
Ask a Dual Question

Every 6 dB doubles the constellation size for a given reliability, for PAM.



Ask a Dual Question

Every 6 dB doubles the constellation size for a given reliability, for PAM



But for QAM, every 6 dB quadruples the constellation size.

Degrees of Freedom Utilized

Definition:

A space-time coding scheme utilizes $r_{\max}$ degrees of freedom/s/Hz if the data rate scales like

$$R(\text{SNR}) \sim r_{\max} \log_2 \text{SNR} \quad \text{bits/s/Hz}$$

for a fixed error probability (reliability)

In a 1×1 channel, $r_{\max} = 1/2$ for PAM, $r_{\max} = 1$ for QAM.

Degrees of Freedom Utilized

Definition:

A space-time coding scheme utilizes $r_{\max}$ degrees of freedom/s/Hz if the data rate scales like

$$R(\text{SNR}) \sim r_{\max} \log_2 \text{SNR} \quad \text{bits/s/Hz}$$

for a fixed error probability (reliability)

In a 1×1 channel, $r_{\max} = 1/2$ for PAM, $r_{\max} = 1$ for QAM.

Note: A space-time coding scheme is a family of codes within a certain structure, with varying symbol alphabet as a function of SNR.

Diversity-Multiplexing Tradeoff

Every 3 dB increase in SNR yields

either

a $2^{-d_{\max}}$ decrease in error probability for a fixed rate;

or

$r_{\max}$ additional bits/s/Hz for a fixed reliability.

Diversity-Multiplexing Tradeoff

Every 3 dB increase in SNR yields

either

a $2^{-d_{\max}}$ decrease in error probability for a fixed rate;

or

$r_{\max}$ additional bits/s/Hz for a fixed reliability.

But these are two extremes of a rate-reliability tradeoff.

Diversity-Multiplexing Tradeoff

Every 3 dB increase in SNR yields

either

a $2^{-d_{\max}}$ decrease in error probability for a fixed rate;

or

$r_{\max}$ additional bits/s/Hz for a fixed reliability.

But these are two extremes of a rate-reliability tradeoff.

More generally, one can increase reliability and the data rate at the same time.

Diversity-Multiplexing Tradeoff of A Scheme

(Zheng and Tse 03)

Definition

A space-time coding scheme achieves a diversity-multiplexing tradeoff curve $d(r)$ if for each multiplexing gain r, simultaneously

$$R(\text{SNR}) \sim r \log_2 \text{SNR} \quad \text{bits/s/Hz}$$

and

$$P_e(\text{SNR}) \sim \text{SNR}^{-d(r)}.$$

Diversity-Multiplexing Tradeoff of A Scheme

(Zheng and Tse 03)

Definition

A space-time coding scheme achieves a diversity-multiplexing tradeoff curve $d(r)$ if for each multiplexing gain r, simultaneously

$$R(\text{SNR}) \sim r \log_2 \text{SNR} \quad \text{bits/s/Hz}$$

and

$$P_e(\text{SNR}) \sim \text{SNR}^{-d(r)}.$$

The largest multiplexing gain is $r_{\max}$, the d.o.f. utilized by the scheme.

The largest diversity gain is $d_{\max} = d(0)$, the classical diversity gain.

Diversity-Multiplexing Tradeoff of the Channel

Definition

The diversity-multiplexing tradeoff $d^*(r)$ of a MIMO channel is the best possible diversity-multiplexing tradeoff achievable by any scheme.

$r^*_{\max}$ is the largest multiplexing gain achievable in the channel.

$d^*_{\max} = d^*(0)$ is the largest diversity gain achievable.

Diversity-Multiplexing Tradeoff of the Channel

Definition

The diversity-multiplexing tradeoff $d^*(r)$ of a MIMO channel is the best possible diversity-multiplexing tradeoff achievable by any scheme.

$r^*_{\max}$ is the largest multiplexing gain achievable in the channel.

$d^*_{\max} = d^*(0)$ is the largest diversity gain achievable.

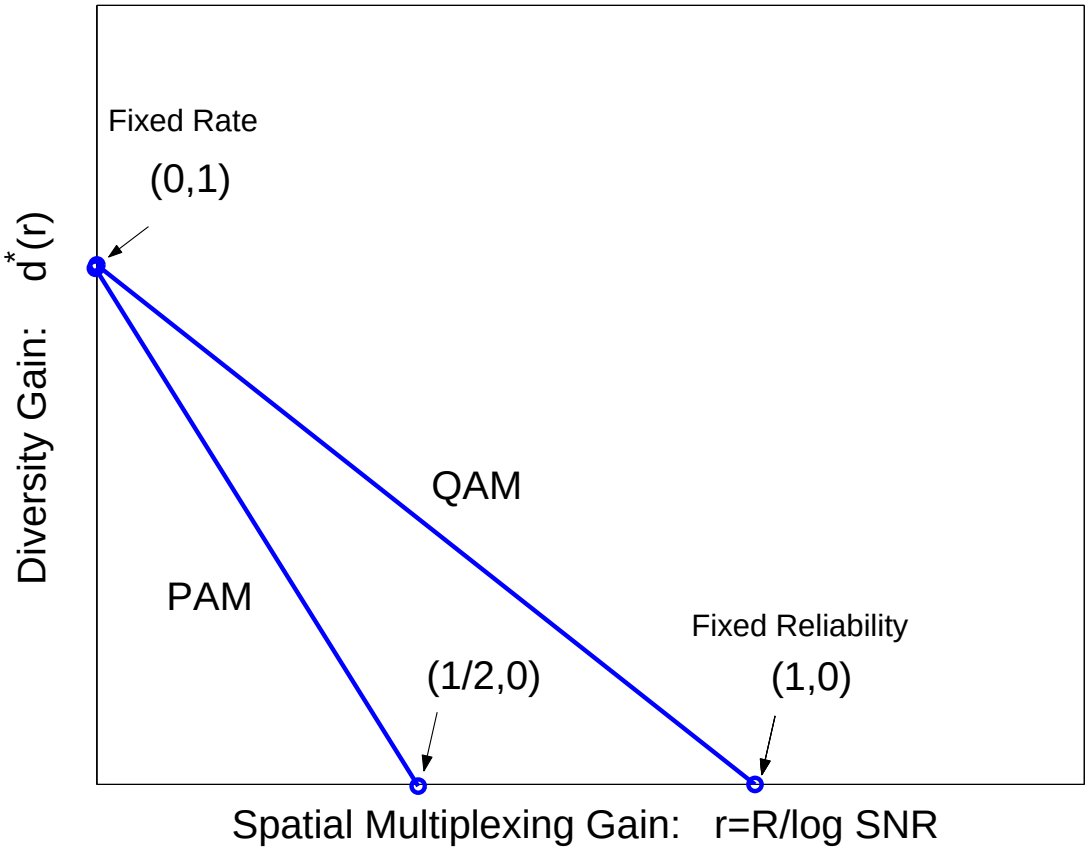
For a $m \times n$ MIMO channel, it is not difficult to show:

$$r^*_{\max} = \min\{m, n\}$$

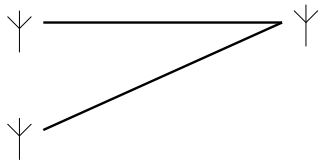
$$d^*_{\max} = mn$$

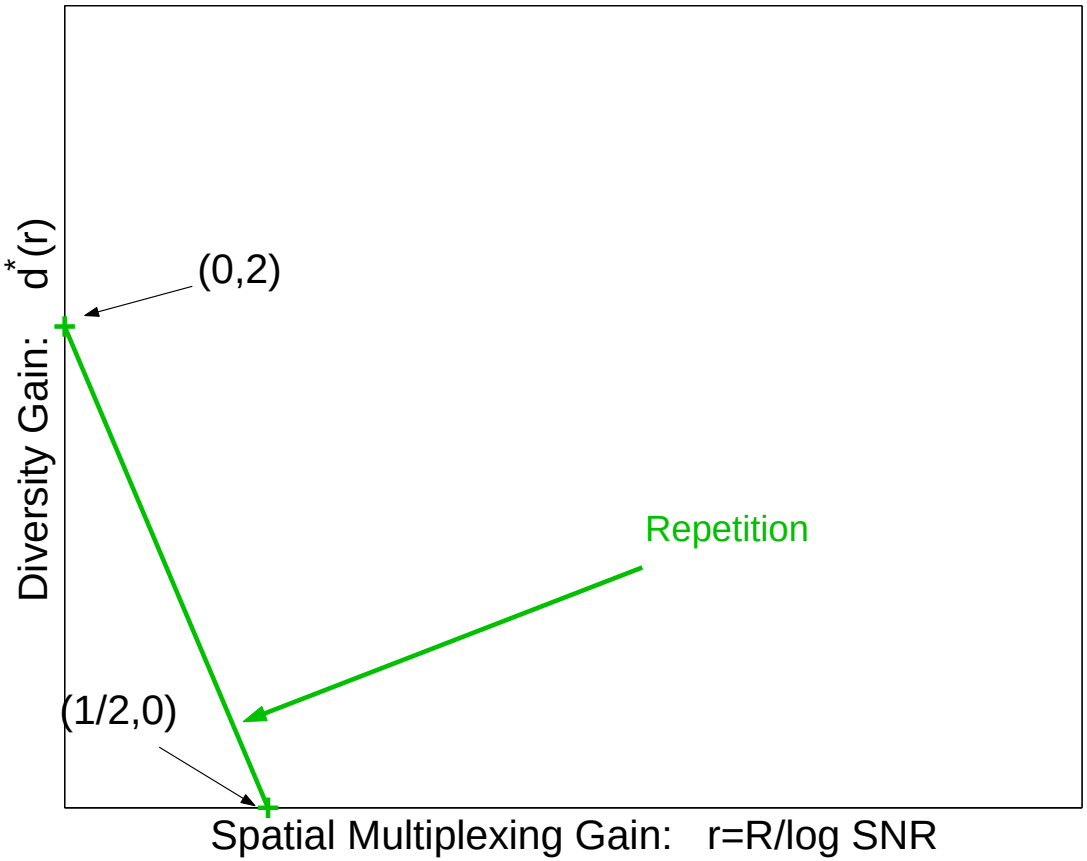
What is more interesting is how the entire curve looks like.

Example: 1×1 Channel

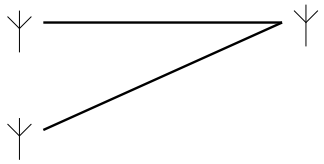


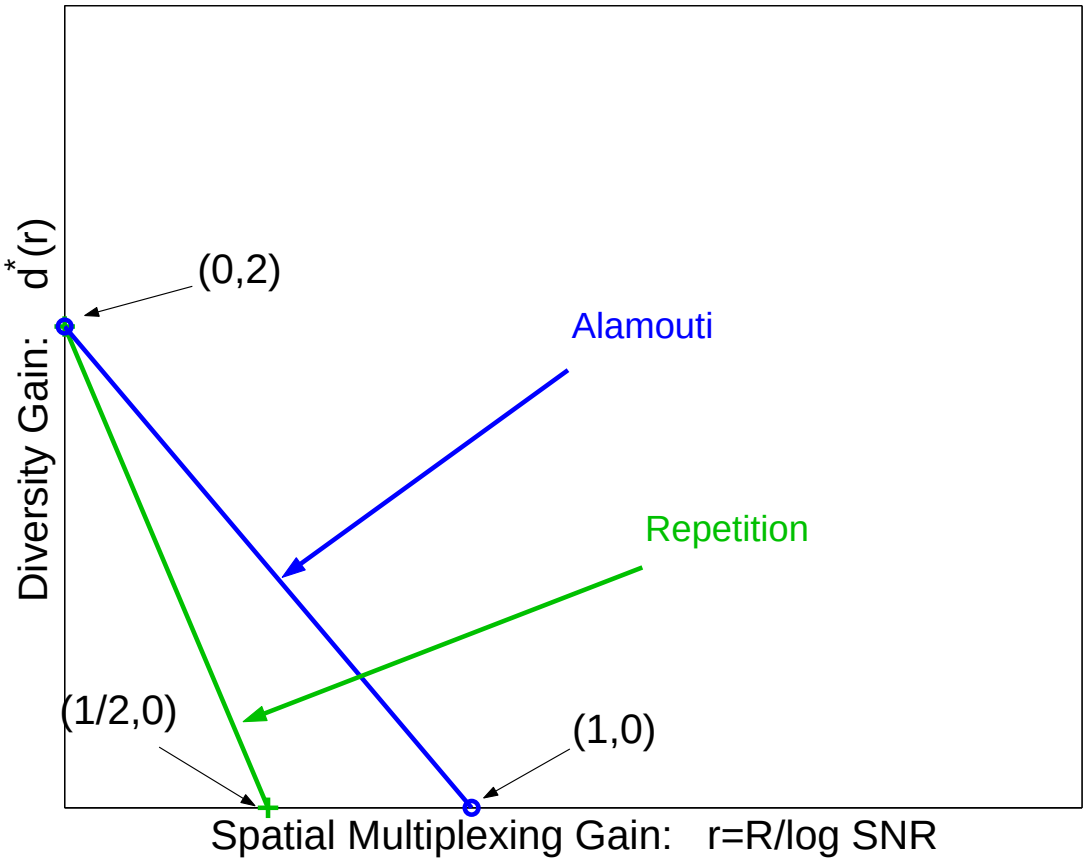
Example: 2×1 Channel



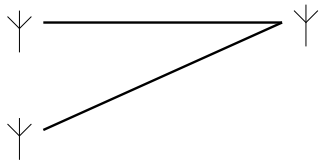


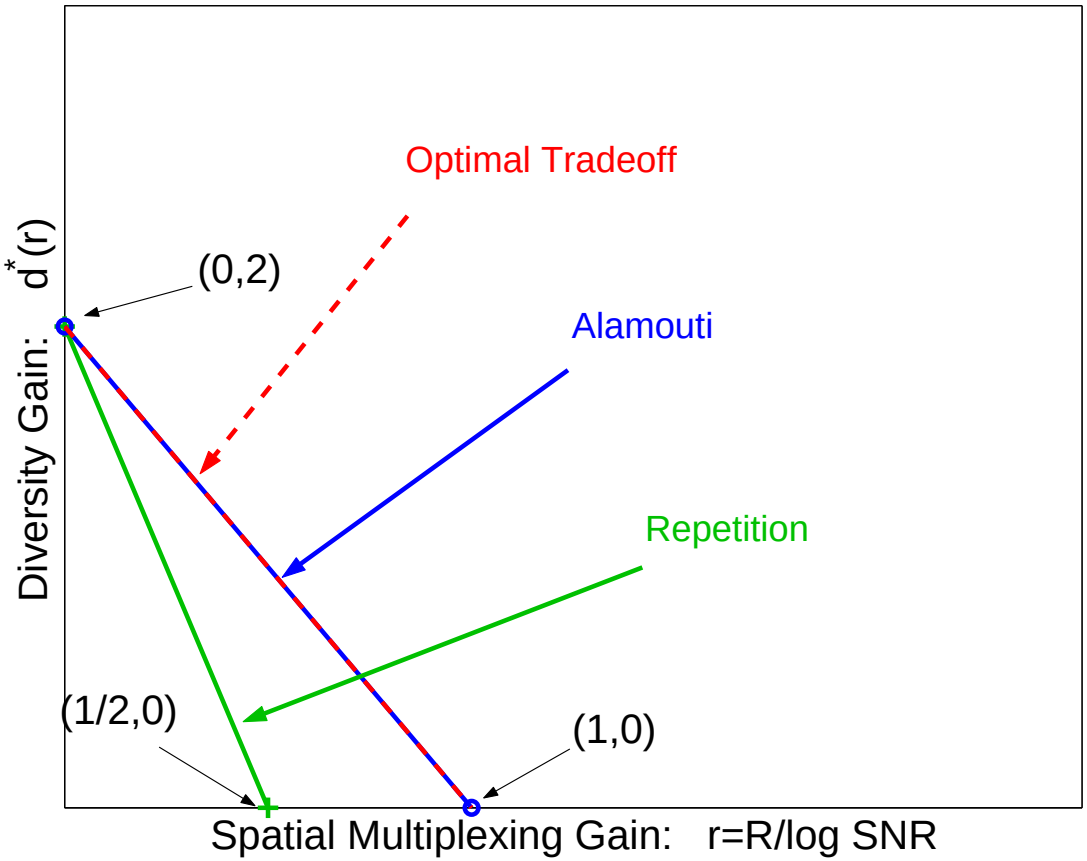
Example: 2×1 Channel



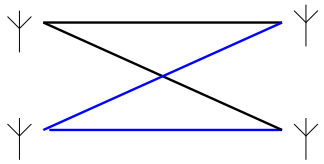


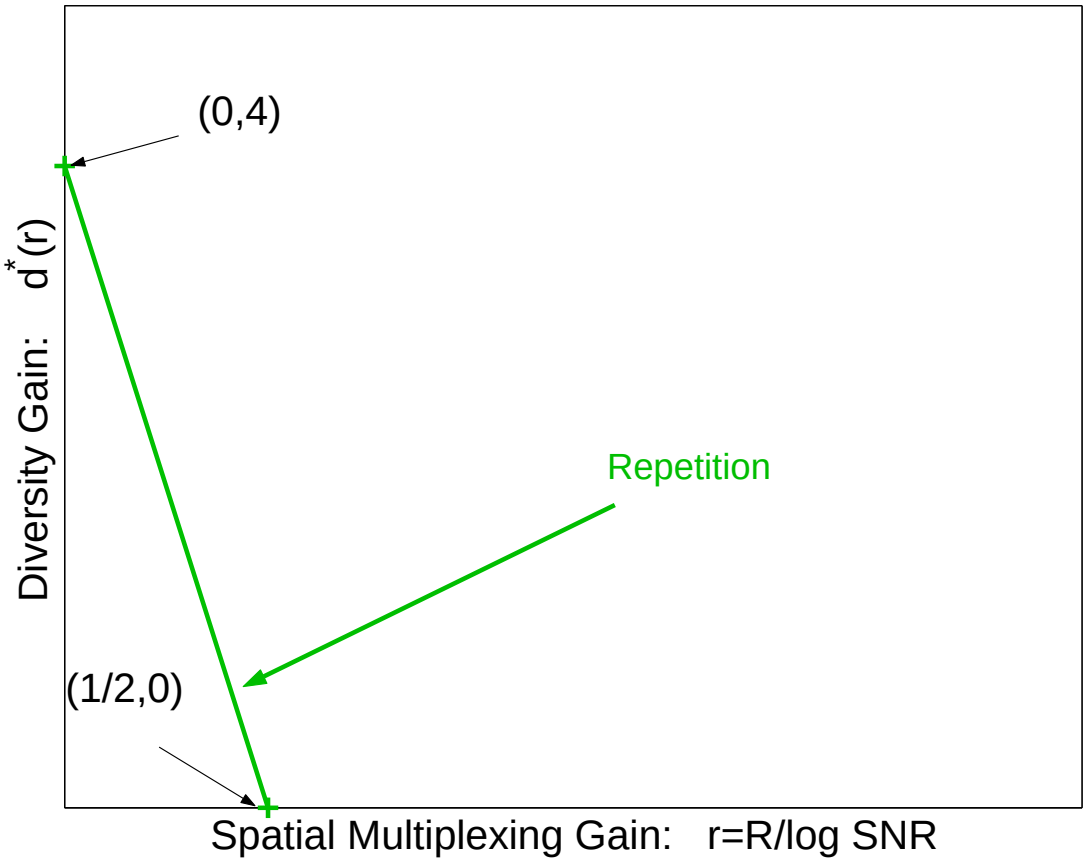
Example: 2×1 Channel



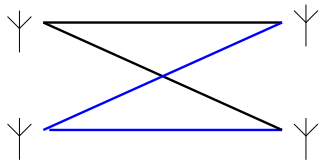


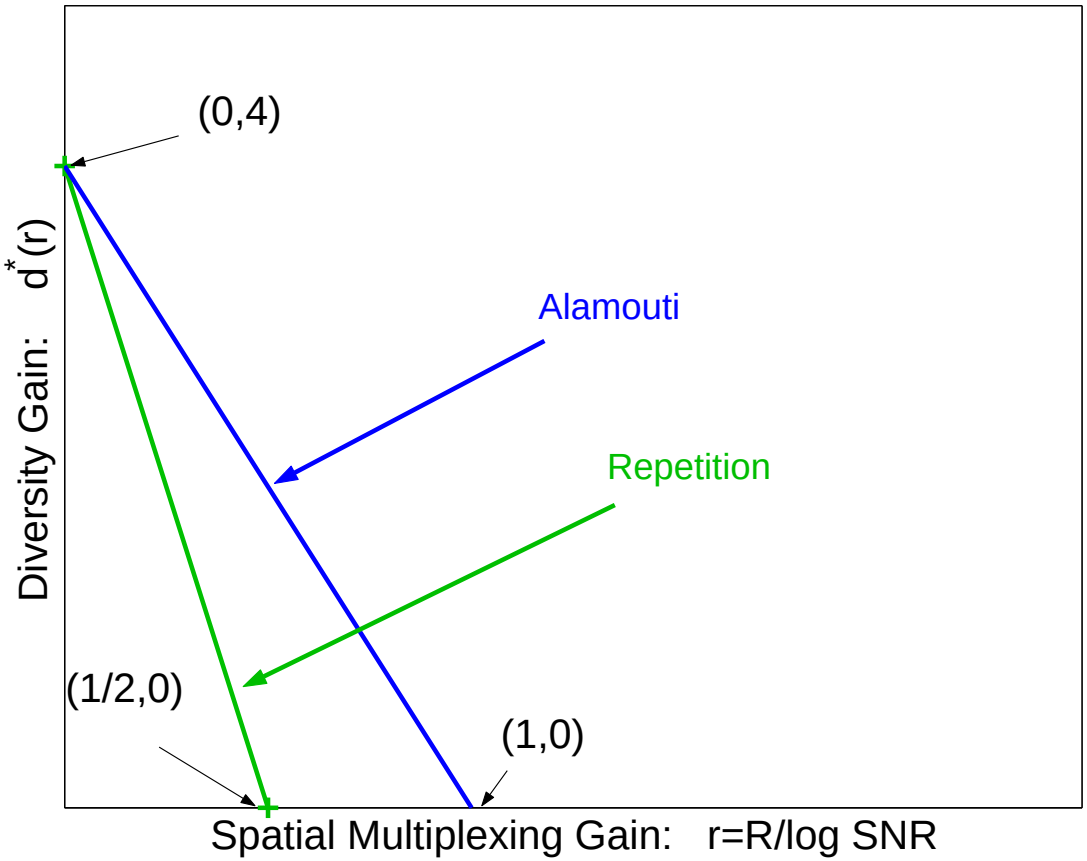
Example: 2×2 Channel



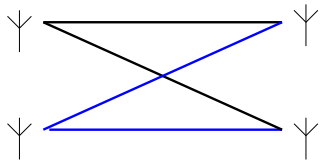


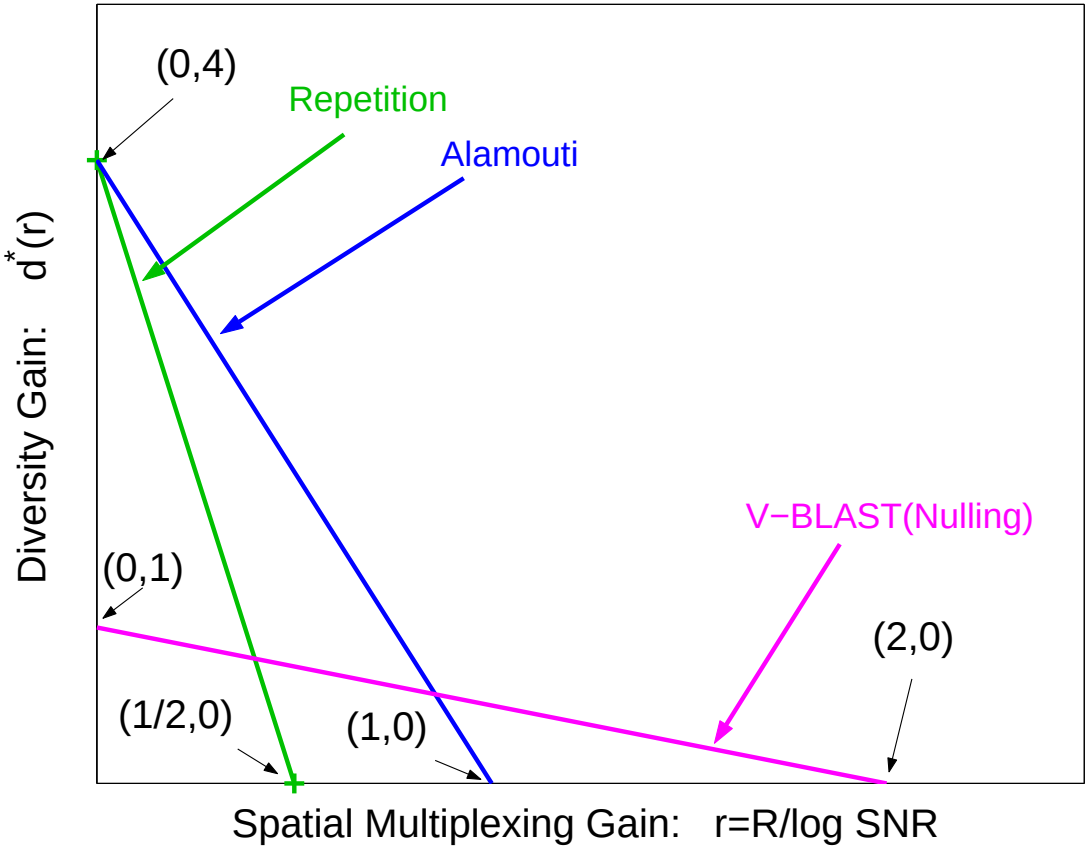
Example: 2×2 Channel



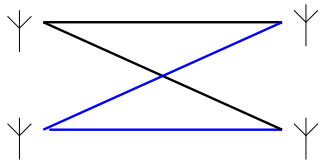


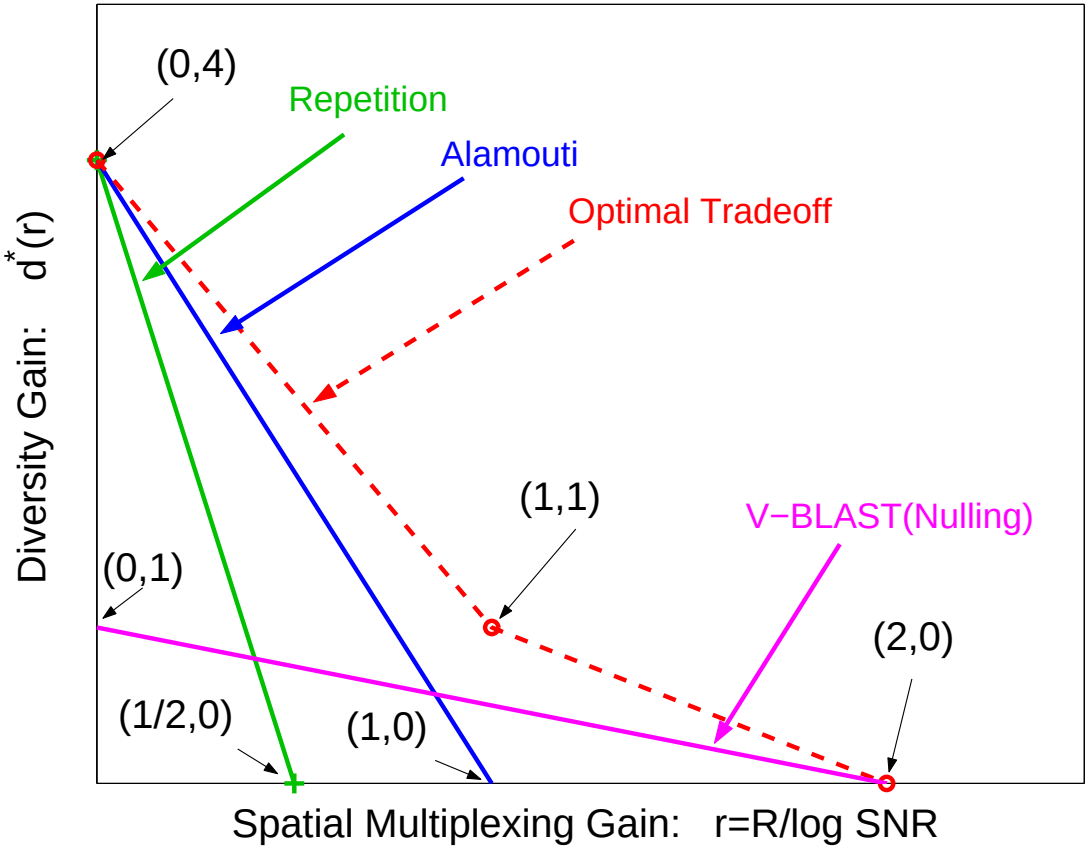
Example: 2×2 Channel



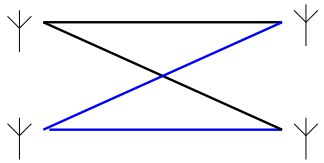


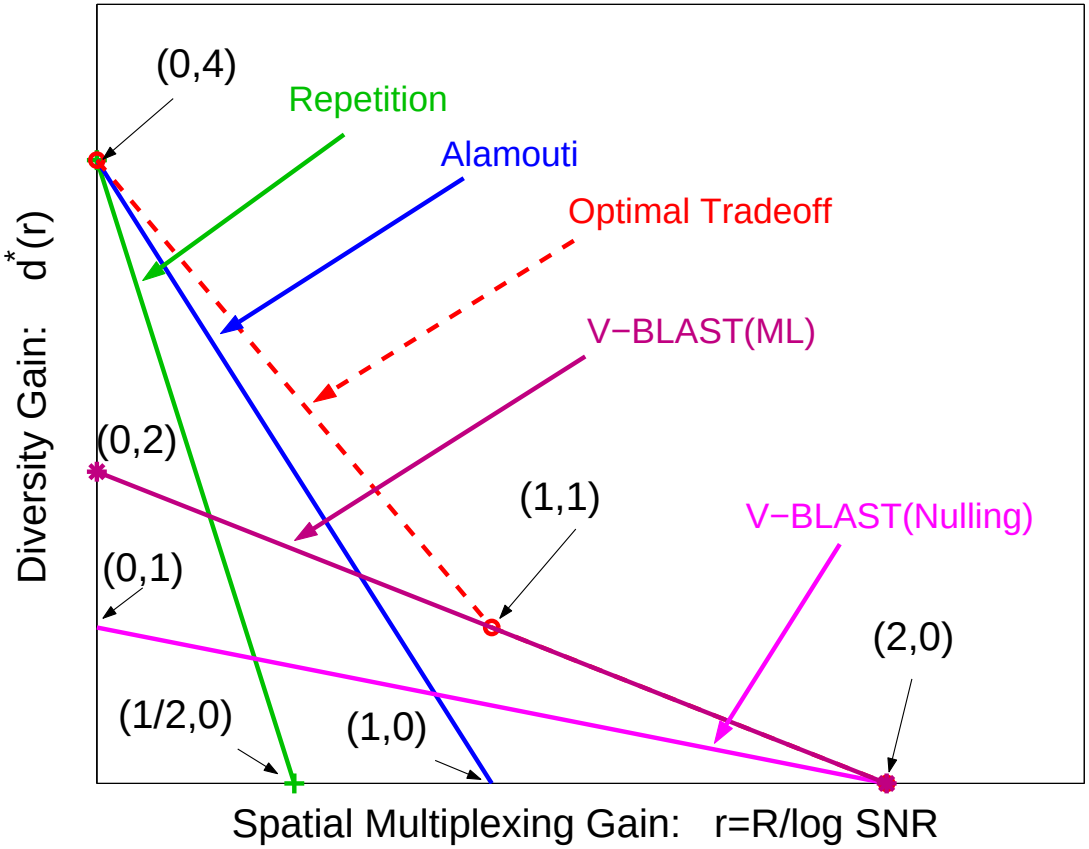
Example: 2×2 Channel



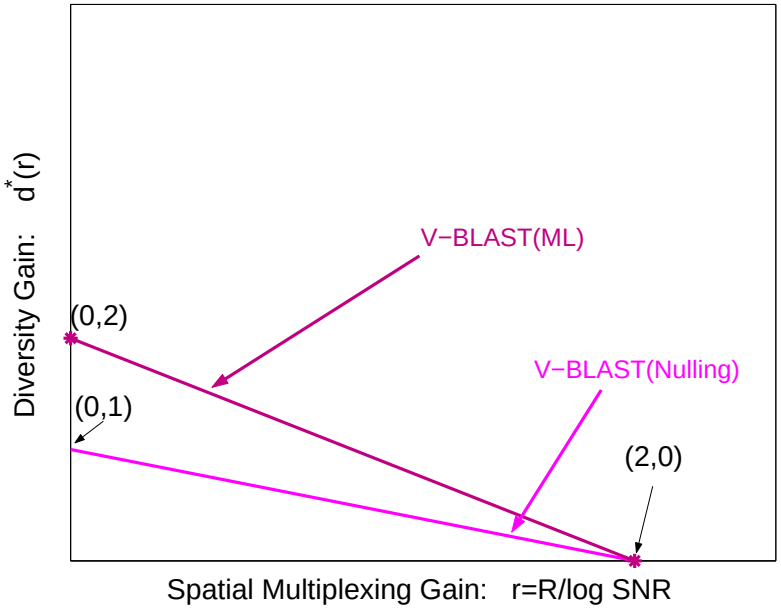


Example: 2×2 Channel





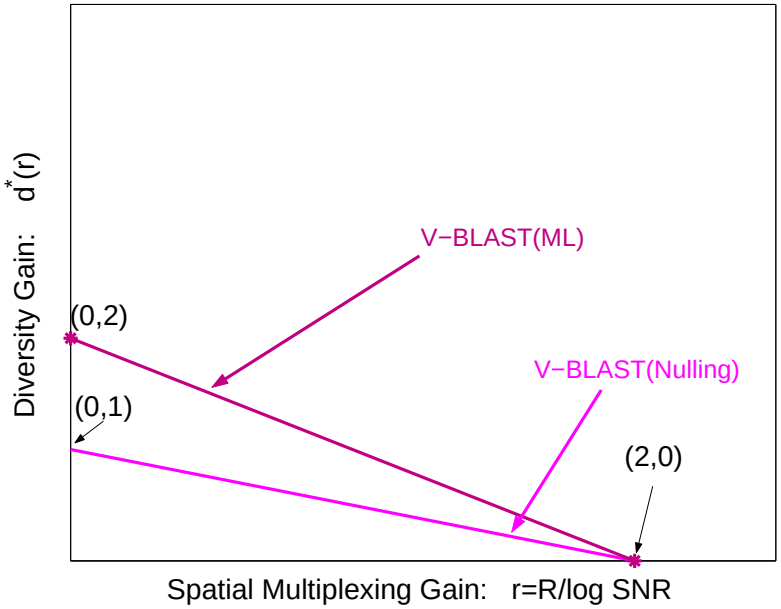
ML vs Nulling in V-Blast



Winters, Salz and Gitlins 93:

Nulling out k interferers using n receive antennas provides a diversity gain of $n-k$.

ML vs Nulling in V-Blast



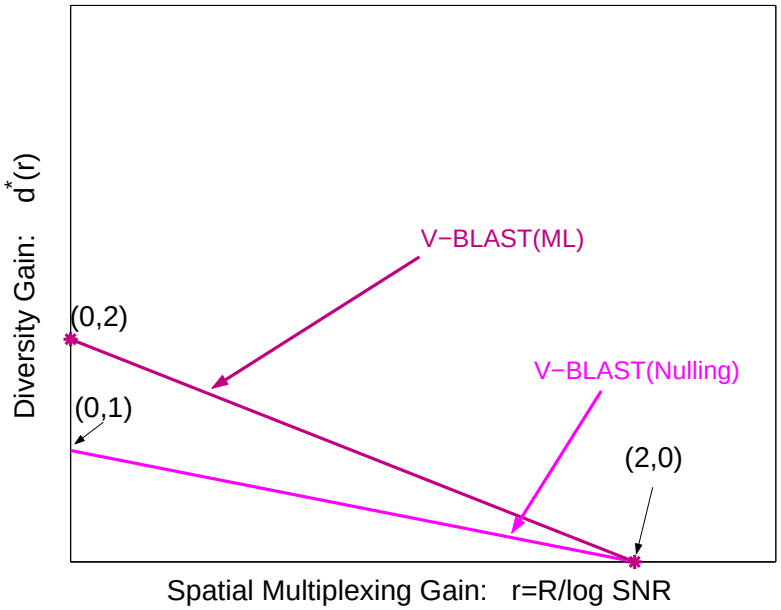
Winters, Salz and Gitlins 93:

Nulling out k interferers using n receive antennas provides a diversity gain of $n-k$.

Tse,Viswanath and Zheng 03:

Jointly detecting all users provides a diversity gain of n to each.

ML vs Nulling in V-Blast



Winters, Salz and Gitlins 93:

Nulling out k interferers using n receive antennas provides a diversity gain of $n-k$.

Tse,Viswanath and Zheng 03:

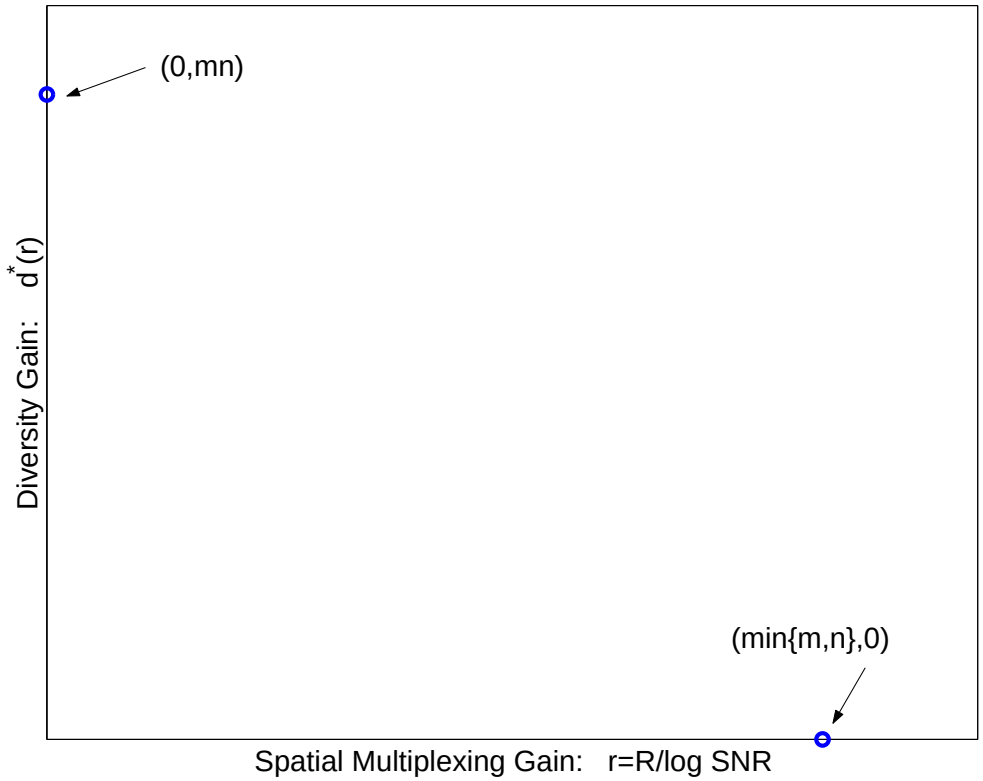
Jointly detecting all users provides a diversity gain of n to each.

There is free lunch. (?)

Optimal D-M Tradeoff for General $m \times n$ Channel

(Zheng and Tse 03)

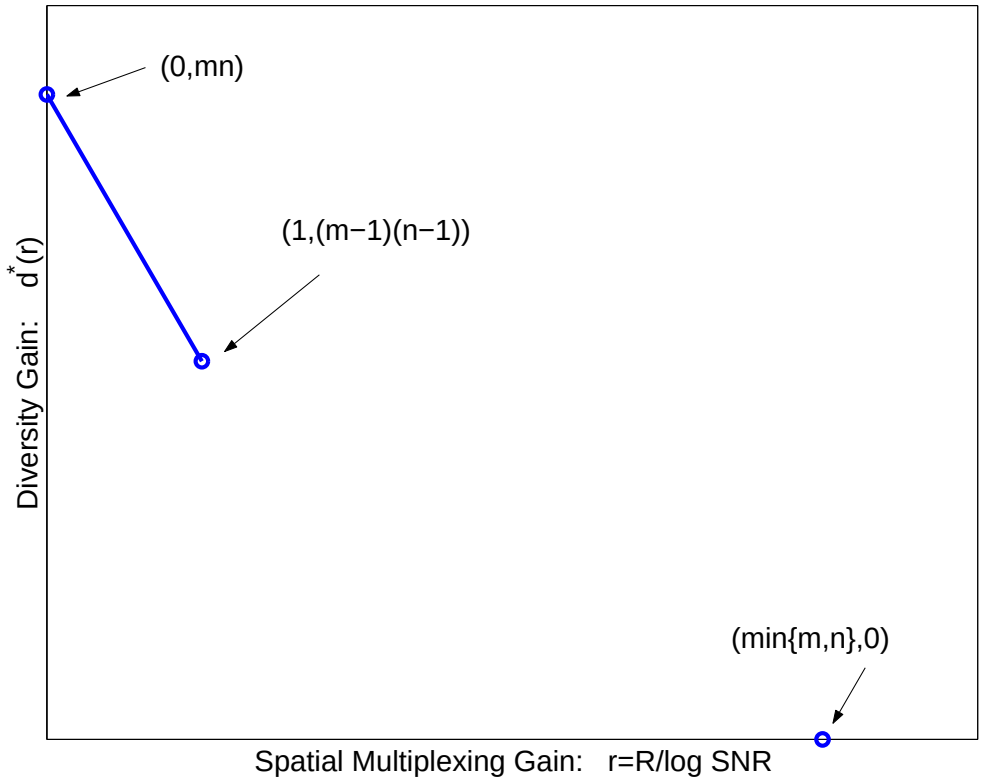
As long as block length $l \geq m + n - 1$:



Optimal D-M Tradeoff for General $m \times n$ Channel

(Zheng and Tse 03)

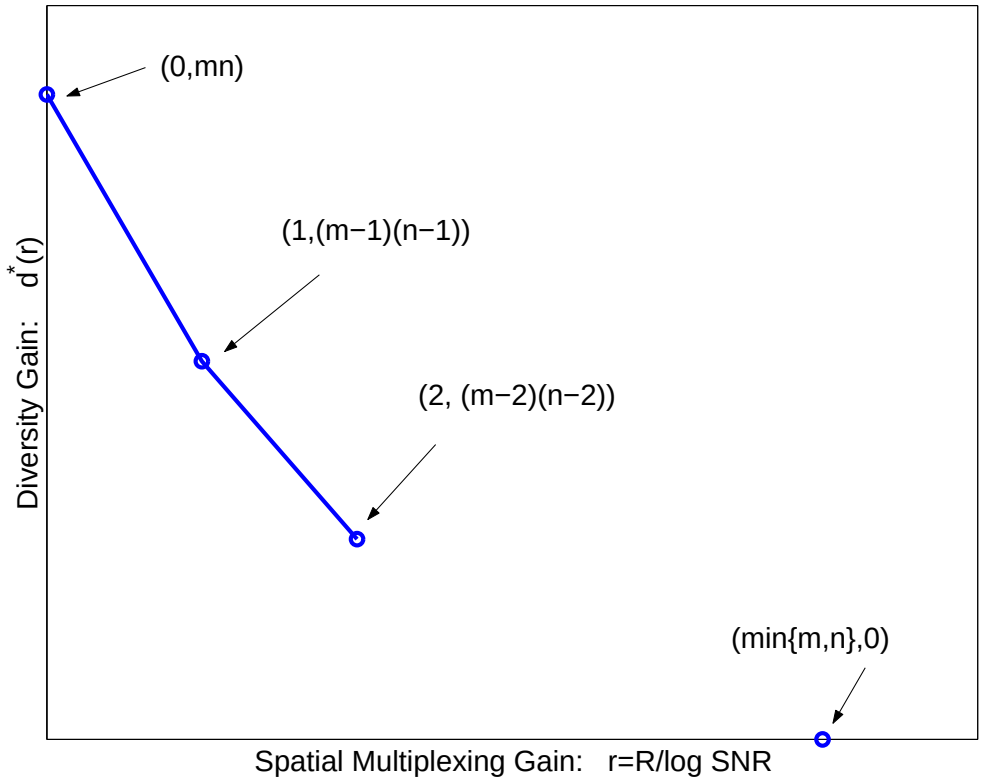
As long as block length $l \geq m + n - 1$:



Optimal D-M Tradeoff for General $m \times n$ Channel

(Zheng and Tse 03)

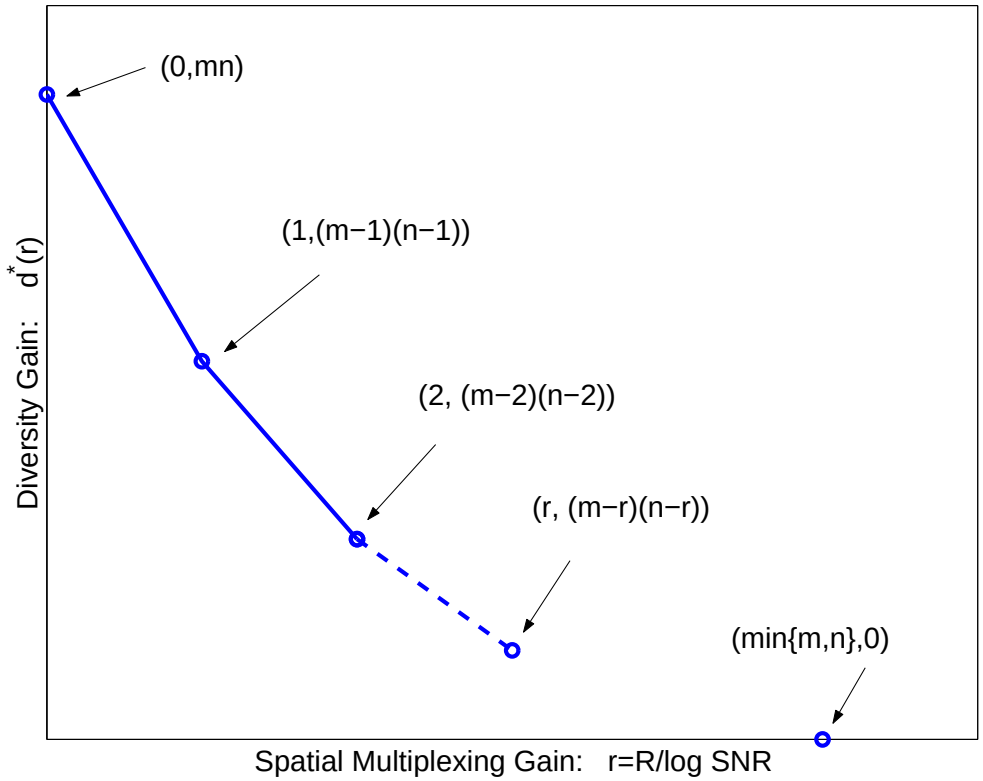
As long as block length $l \geq m + n - 1$:



Optimal D-M Tradeoff for General $m \times n$ Channel

(Zheng and Tse 03)

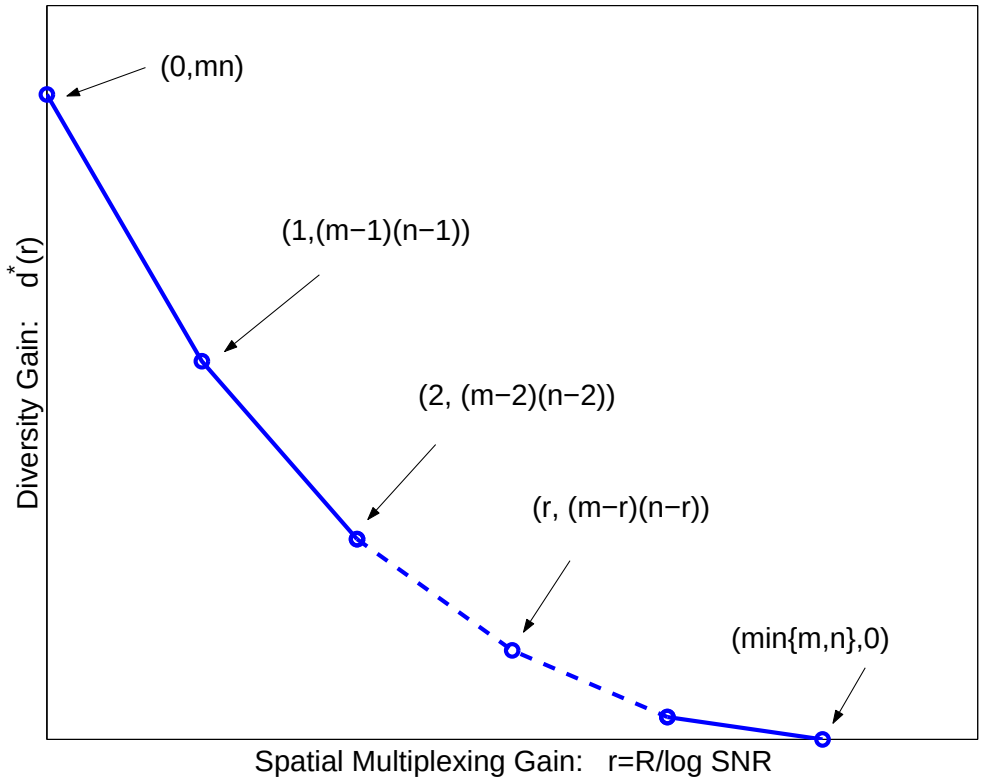
As long as block length $l \geq m + n - 1$:



Optimal D-M Tradeoff for General $m \times n$ Channel

(Zheng and Tse 03)

As long as block length $l \geq m + n - 1$:



For integer r, it is *as though* r transmit and r receive antennas were dedicated for multiplexing and the rest provide diversity.

Achieving Optimal Diversity-Multiplexing Tradeoff

- Hao and Wornell 03: MIMO rotation code (2×2 channel only).
- Tavildar and Viswanath 04: D-Blast plus permutation code.
- El Gamal, Caire and Damen 03: Lattice codes.

Hao and Wornell 03

Alamouti scheme:

$$\begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}$$

Hao and Wornell's scheme:

$$\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$$

where

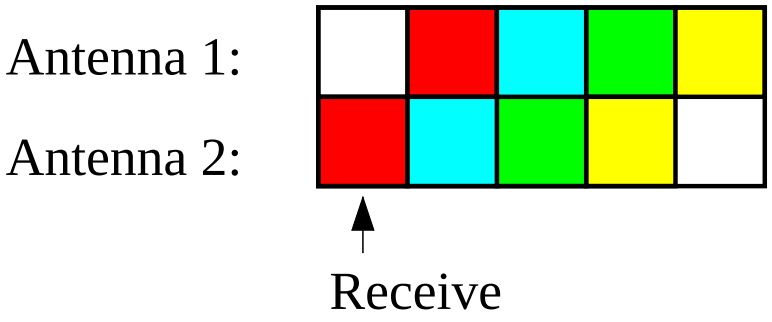
$$\begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \text{Rotate}(\theta_1^*) \begin{bmatrix} u_1 \\ u_4 \end{bmatrix} \qquad \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \text{Rotate}(\theta_2^*) \begin{bmatrix} u_2 \\ u_3 \end{bmatrix}$$

and u_1, u_2, u_3, u_4 are independent QAM symbols.

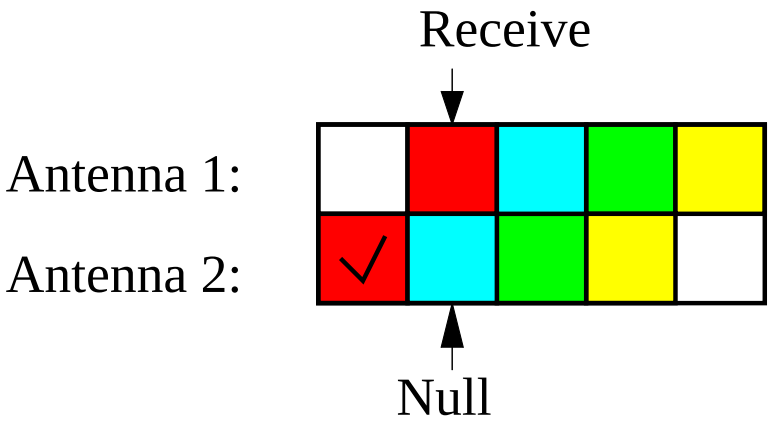
Tavildar and Viswanth 04

- First use D-Blast to convert the MIMO channel into a parallel channel.
- Then design permutation codes to achieve the optimal diversity-multiplexing tradeoff on the parallel channel.

D-BLAST



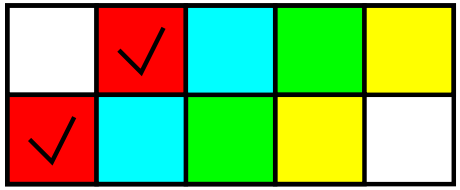
D-BLAST



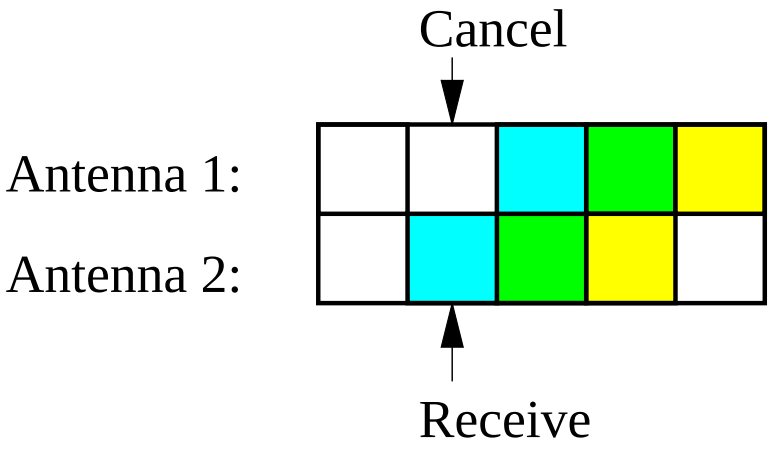
D-BLAST

Antenna 1:

Antenna 2:

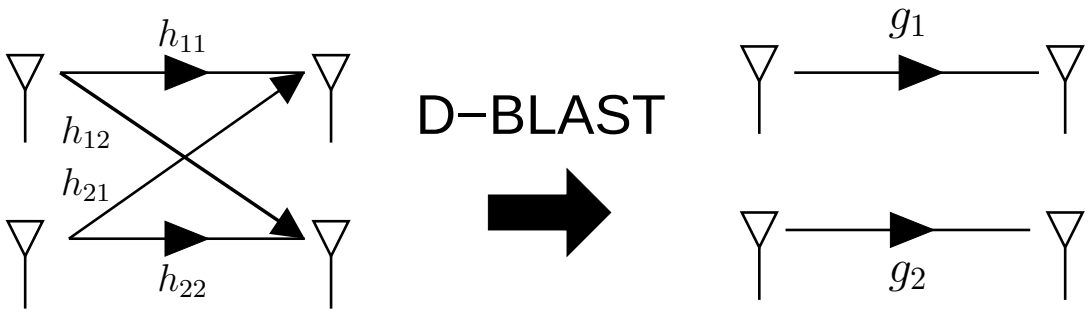


D-BLAST



Original D-Blast is sub-optimal.

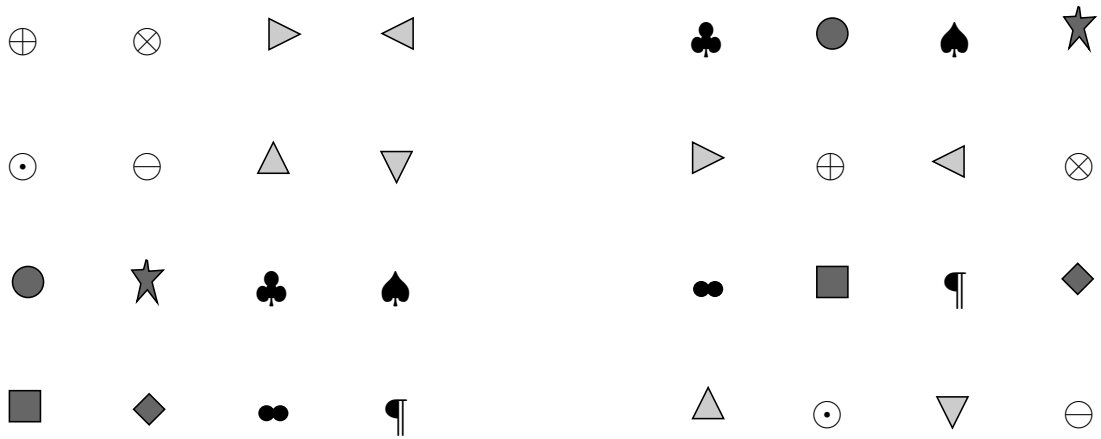
D-Blast with MMSE suppression is information lossless



Permutation Coding for Parallel Channel

The channel is parallel but the fading at the different sub-channels are correlated.

Nevertheless it is shown that the permutation codes can achieve the optimal diversity-multiplexing tradeoff of the parallel channel.



Conclusion

Diversity-multiplexing tradeoff is a unified way to look at space-time code design for MIMO channels.

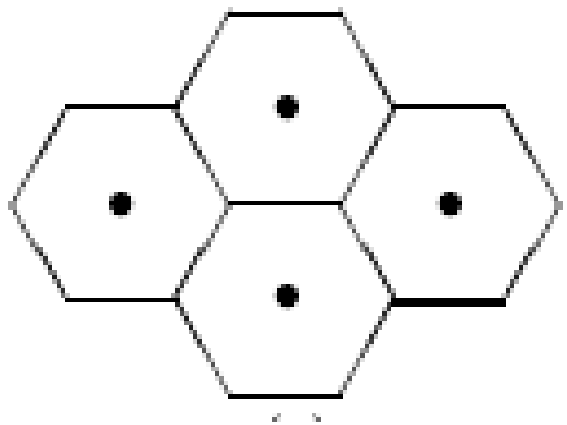
It puts diversity and multiplexing on an equal footing.

It provides a framework to compare existing schemes as well as stimulates the design of new schemes.

7. Cellular Systems: Multiple Access and Interference Management

Cellular Systems

- So far we have focused on point-to-point communication.
- In a cellular system, additional issues come into forefront:
  - Multiple access
  - Inter-cell interference management



Some History

- Cellular concept (Bell Labs, early 70's)
- AMPS (analog, early 80's)
- GSM (digital, narrowband, late 80's)
- IS-95 (digital, wideband, early 90's)
- 3G/4G systems

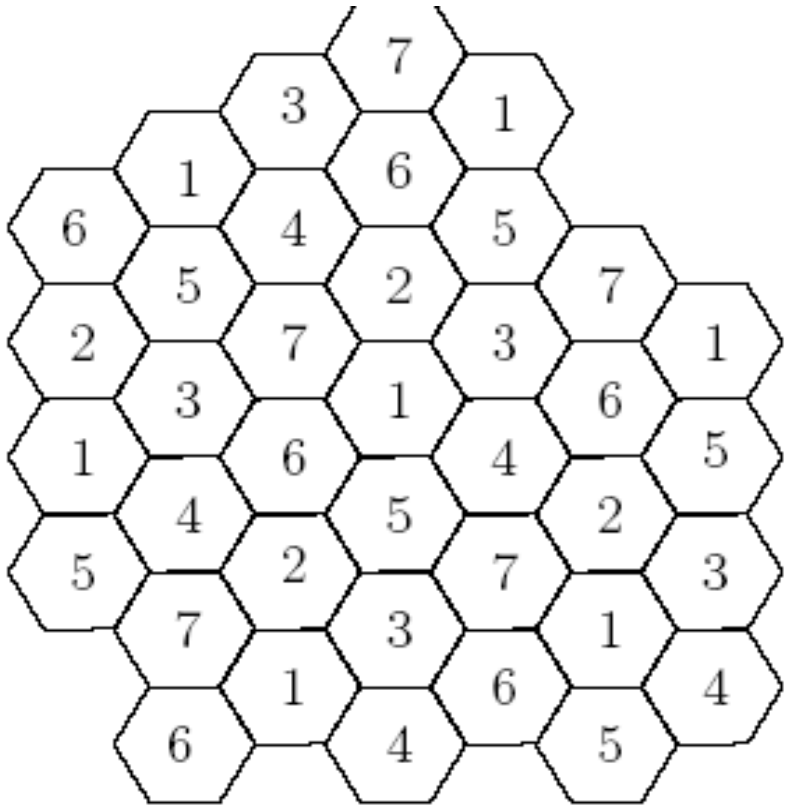
Four Systems

- Narrowband (GSM)
- Wideband CDMA (IS-95)
- Wideband OFDM (Flash OFDM)
- Opportunistic Communication (1x EV-DO)

Narrowband (GSM)

- The total bandwidth is divided into many narrowband channels. (200 kHz in GSM)
- Users are given time slots in a narrowband channel (8 users)
- Multiple access is orthogonal: users within the cell never interfere with each other.
- Interference between users on the same channel in different cells is minimized by reusing the same channel only in cells far apart.
- Users operate at high SINR regime
- The price to pay is in reducing the overall available degrees of freedom.

Frequency Reuse



Frequency reuse is poor in narrowband systems because of lack of interference averaging.

Wideband System: IS-95

- Universal frequency reuse: all the users in all cells share the same bandwidth (1.25 MHz)
- Main advantages:
  - Maximizes the degrees of freedom usage
  - Allows interference averaging across many users.
  - Soft capacity limit
  - Allows soft handoff
  - Simplify frequency planning
- Challenges
  - Very tight power control to solve the near-far problem.
  - More sophisticated coding/signal processing to extract the information of each user in a very low SINR environment.

Design Goals

- 1) make the interference look as much like a white Gaussian noise as possible:
  - Spread each user's signal using a pseudonoise noise sequence
  - Tight power control for managing interference within the cell
  - Averaging interference from outside the cell as well as fluctuating voice activities of users.
- 2) apply point-to-point design for each link
  - Extract all possible diversity in the channel

Point-to-Point Link Design

- Very low SINR per chip: can be less than -15 dB.
- Diversity is very important at such low SINR.
- Use very low-rate convolution codes
- Time diversity is obtained by interleaving across different coherence time.
- Frequency diversity is obtained by Rake combining of the multipaths.
- Transmit diversity in 3G CDMA systems

Power Control

- Maintain equal received power for all users in the cell
- Tough problem since the dynamic range is very wide. Users' attenuation can differ by many 10's of dB
- Consists of both open-loop and closed loop
- Open loop sets a reference point
- Closed loop is needed since IS-95 is FDD (frequency-division duplex)
- Consists of 1-bit up-down feedback at 800 Hz.
- Not cheap: consumes about 10% of capacity for voice.

Interference Averaging

The received signal-to-interference-plus-noise ratio for a user:

$$\text{SINR} = \frac{P}{N_0 + KP + \sum_i I_i}$$

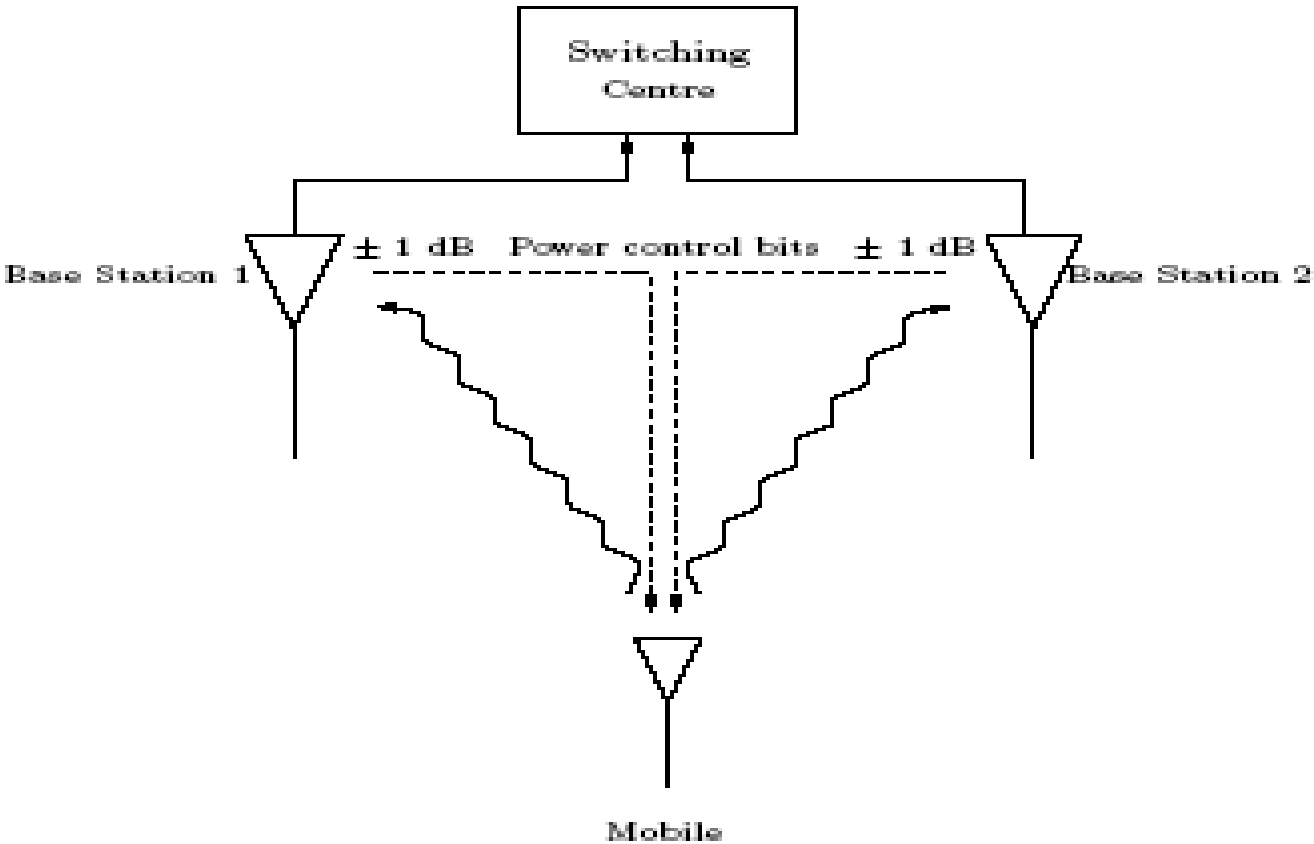
In a large system, each interferer contributes a small fraction of the total out-of-cell interference.

This can be viewed as providing **interference diversity**.

Same interference-averaging principle applies to voice activity and imperfect power control.

Soft Handoff

- Provides another form of diversity: macrodiversity



Uplink vs Downlink

- Can make downlink signals for different users orthogonal at the transmitter. Still because of multipaths, they are not orthogonal at the receiver.
- Less interference averaging: interference come from a few high-power base stations as opposed to many low-power mobiles.

Problem with CDMA

- In-cell interference reduces capacity.
- More importantly, power control is expensive, particularly for data applications where users are very bursty and have low duty cycle.
- In-cell interference is not an inhererent property of systems with universal frequency reuse.
- We can keep users in the cell orthogonal.

Wideband System: OFDM

- We have seen OFDM as a point-to-point modulation scheme, converting the frequency-selective channel intlo a parallel channel.
- It can also be used as a multiple access technique.
- By assigning different time/frequency slots to users, they can be kept orthogonal, no matter what the multipath channels are.
- The key property of sinusoids is that they are eigenfunctions of all linear time-invariant channels.

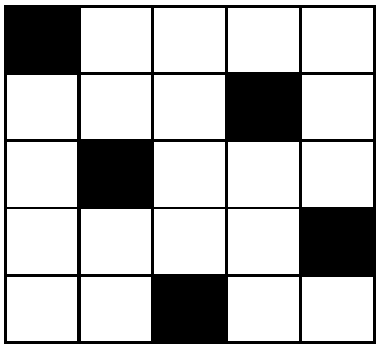
In-cell Orthogonality

- The basic unit of resource is a virtual channel: a hopping sequence.
- Each hopping sequence spans all the sub-carriers to get full frequency-diversity.
- Coding is performed across the symbols in a hopping sequence.
- Hopping sequences of different virtual channels in a cell are orthogonal.
- Each user is assigned a number of virtual channels depending on their data rate requirement.

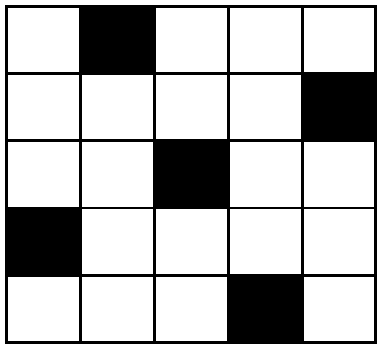
Example

相对GSM，徒增了time diversity

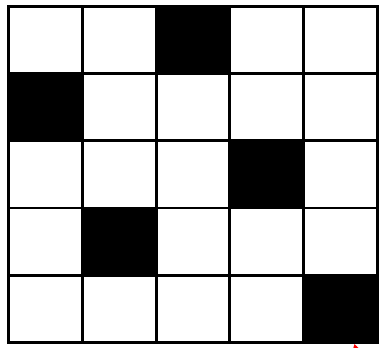
Virtual Channel 0



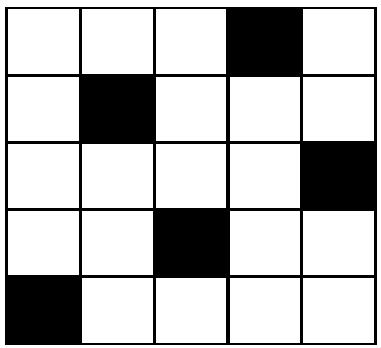
Virtual Channel 1



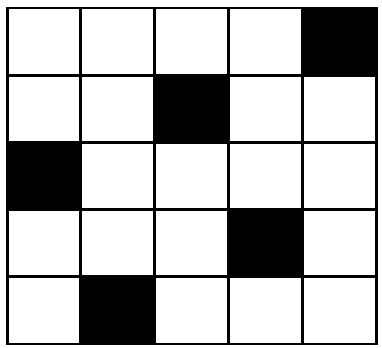
Virtual Channel 2



无需均衡，但需要精确同步



Virtual Channel 3



Virtual Channel 4

1ms

Out-of-Cell Interference Averaging

- The hopping patterns of virtual channels in adjacent cells are designed such that any pair has minimal overlap.

  the prob of failing can be very low

- This ensures that a virtual channel sees interference from many users instead of a single strong user.

- This is a form of interference diversity.

"fixed overhead"

Example: Flash OFDM

- Bandwidth = 1.25 Mz
- # of data sub-carriers = 113
- OFDM symbol = 128 samples = 100 μ s
- Cyclic prefix = 16 samples = 11 μ s delay spread

States of Users

- Users are divided into 3 states:
  - Active: users that are currently assigned virtual channels (<30)
  - Hold: users that are not sending data but maintain synchronization (<130)
  - Inactive (<1000)
- Users in hold state can be moved into active states very quickly.
- Because of the orthogonality property, tight power control is not crucial and this enables quick access for these users
- Important for certain applications (requests for http transfers, acknowledgements, etc.)

可以立即激活开始传输，无需功率控制。

8. Opportunistic Communication and Multiuser Diversity

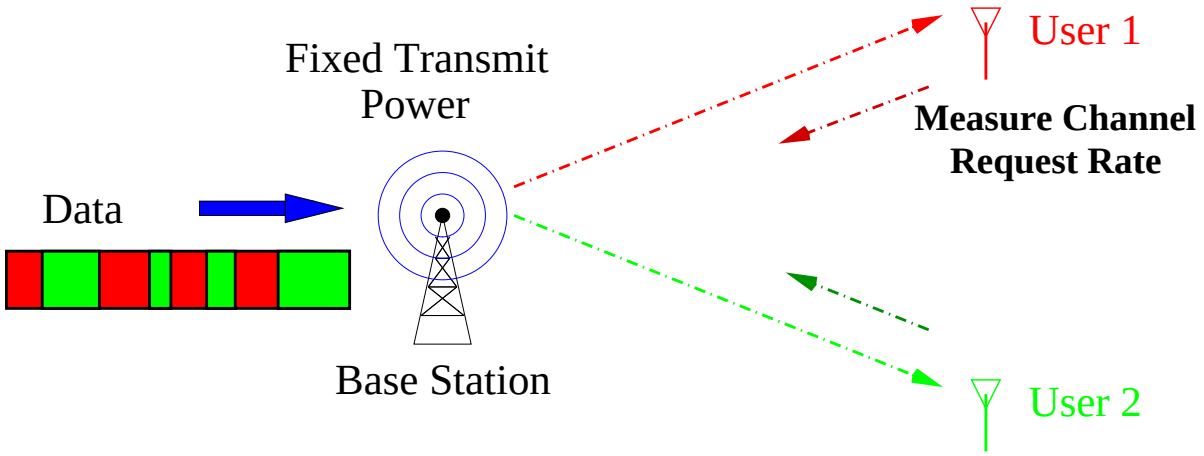
Opportunistic Communication

One line summary:

Transmit when and where the channel is good.

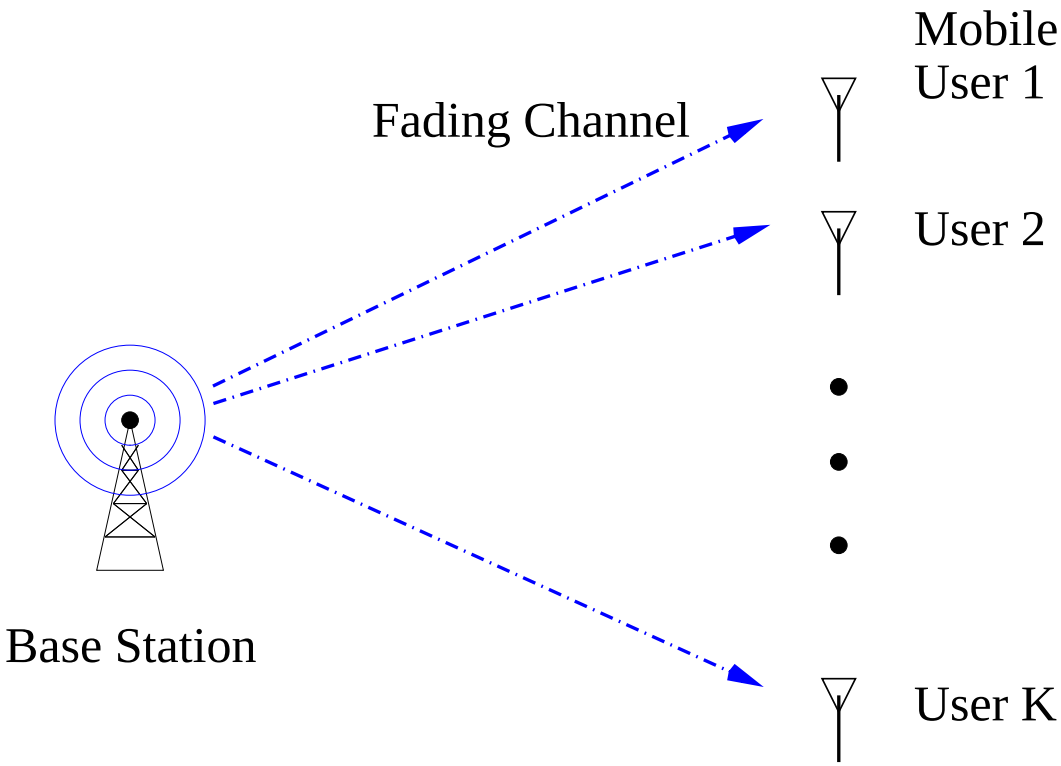
Qualcomm HDR's DownLink

HDR (1xEV-DO): a wireless data system operating on IS-95 band (1.25 MHz)



- HDR downlink operates on a time-division basis.
- Scheduler decides which user to serve in each time-slot.

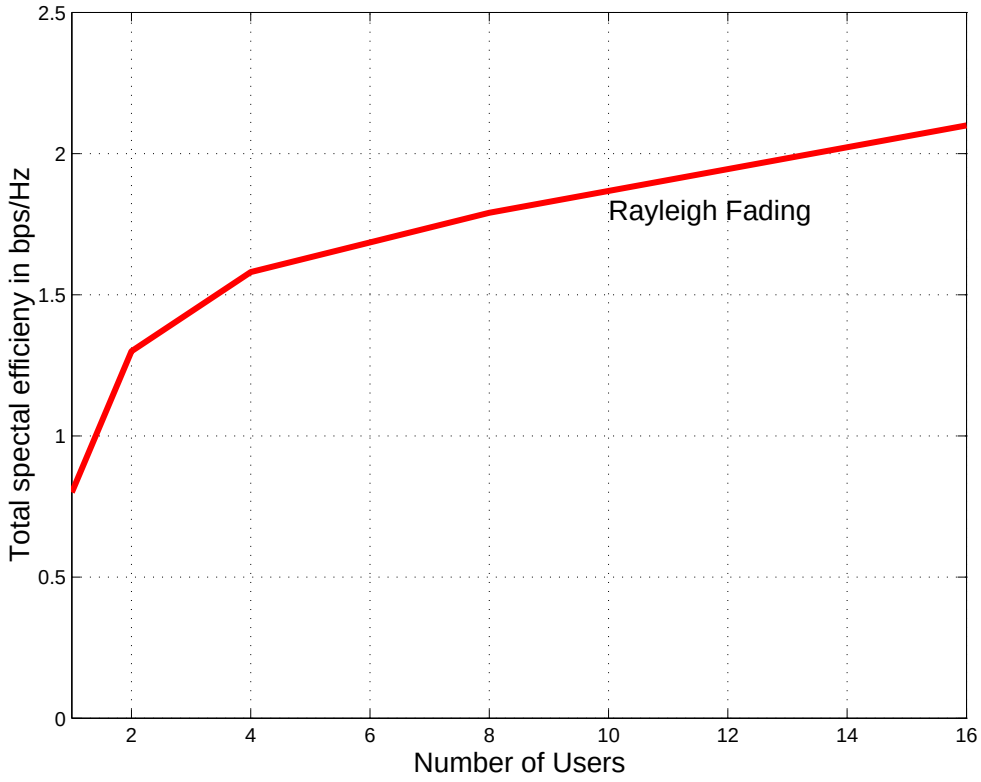
Downlink Multiuser Fading Channel



What is the sum capacity with channel state feedback?

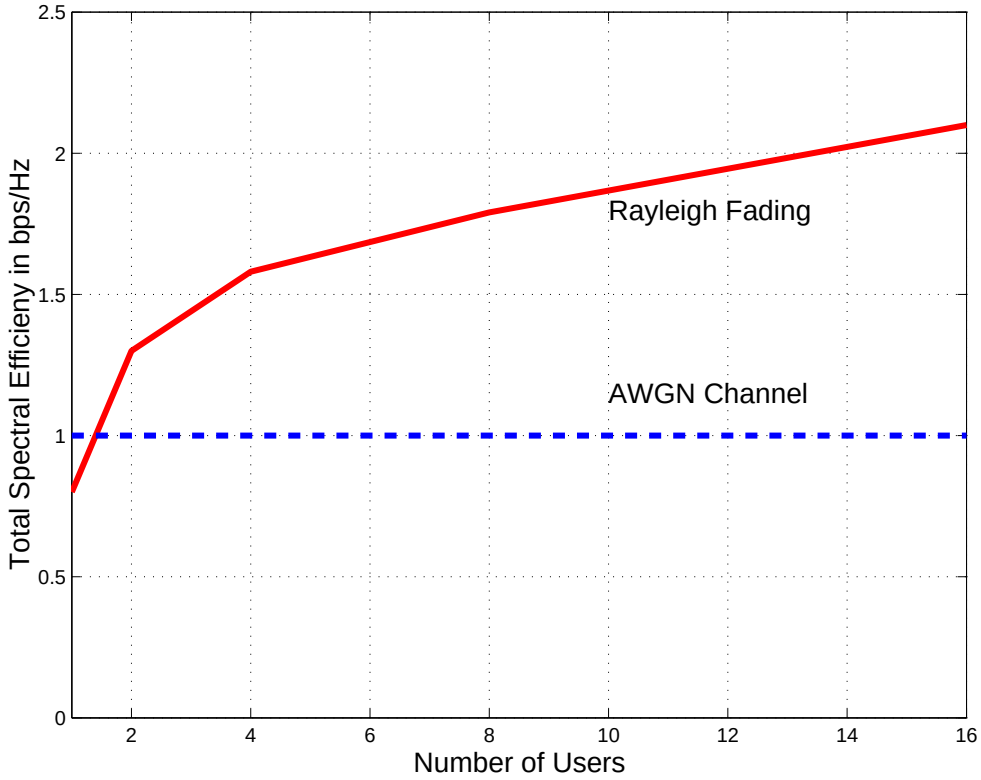
Information Theoretic Capacity of Downlink

(Tse 97)



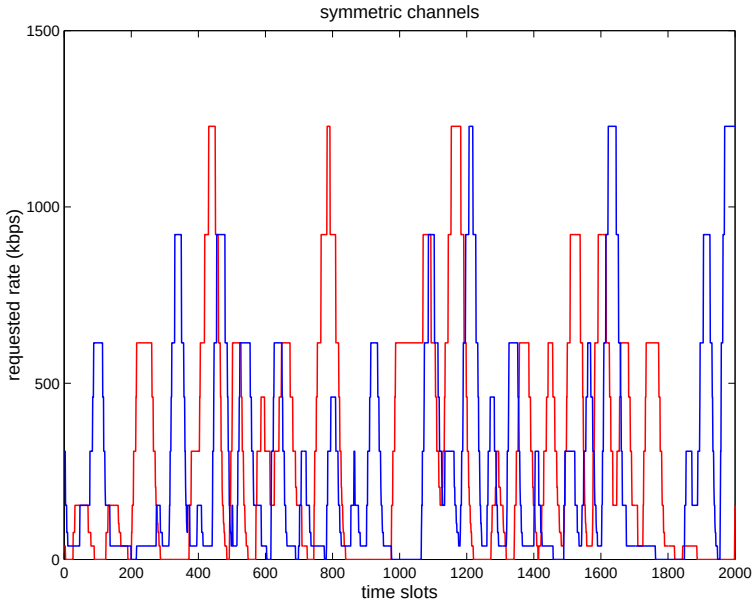
Each user undergoes independent Rayleigh fading with average received signal-to-noise ratio SNR $= 0$dB.

To Fade or Not to Fade?



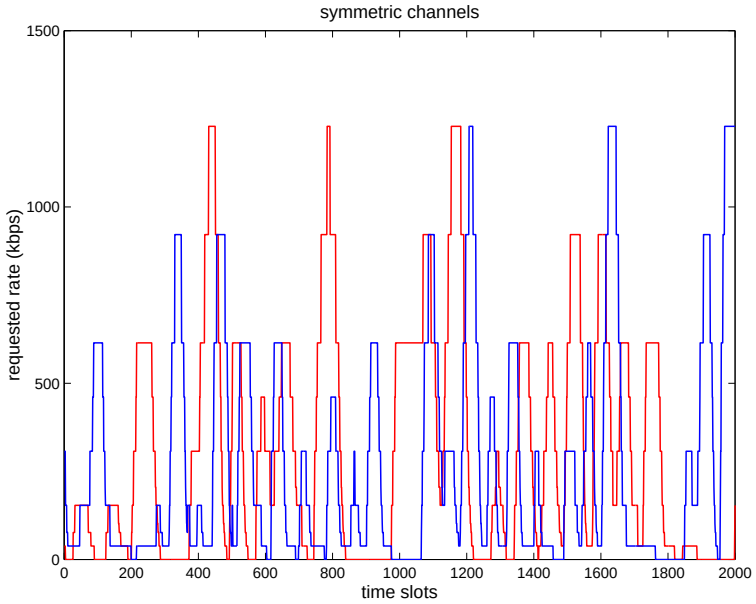
Sum Capacity of fading channel much larger than non-faded channel!

Multiuser Diversity



- In a large system with users fading independently, there is likely to be a user with a very good channel at any time.
- Long term total throughput can be maximized by always serving the user with the strongest channel.

Multiuser Diversity



- In a large system with users fading independently, there is likely to be a user with a very good channel at any time.
- Long term total throughput can be maximized by always serving the user with the strongest channel.

$$\text{effective SNR at time } t = \max_{1 \le k \le K} |h_k(t)|^2.$$

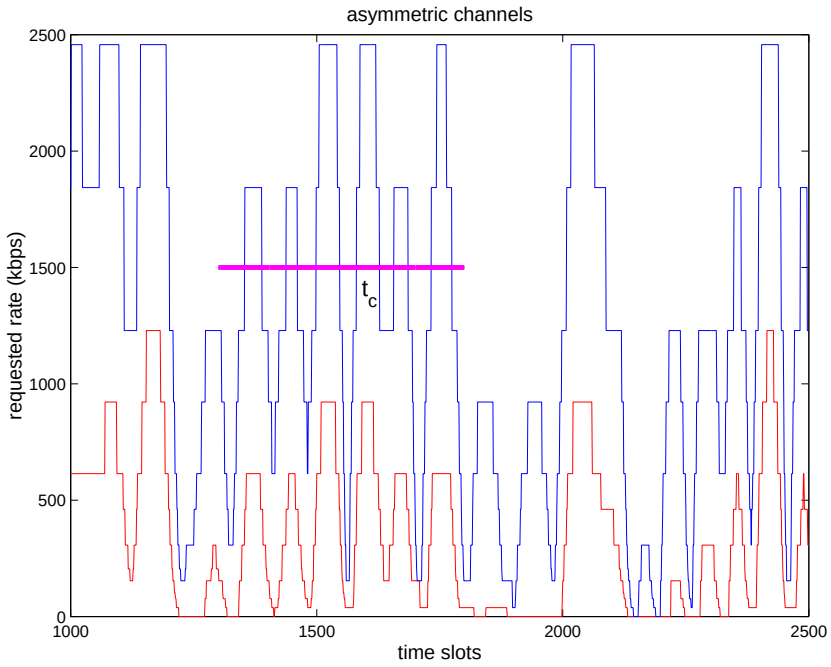
Multiuser Diversity

- Diversity in wireless systems arises from independent signal paths.
- Traditional forms of diversity includes time, frequency and antennas.
- Multiuser diversity arises from independent fading channels across different users.

Multiuser Diversity

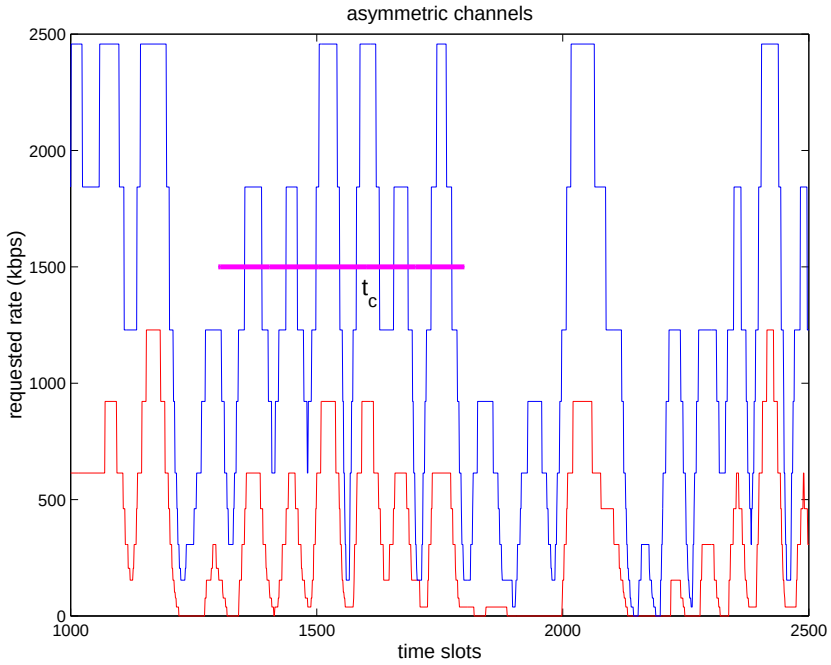
- Diversity in wireless systems arises from independent signal paths.
- Traditional forms of diversity includes time, frequency and antennas.
- Multiuser diversity arises from independent fading channels across different users.
- Fundamental difference: Traditional diversity modes pertain to point-to-point links, while multiuser diversity provides network-wide benefit.

Fairness and Delay



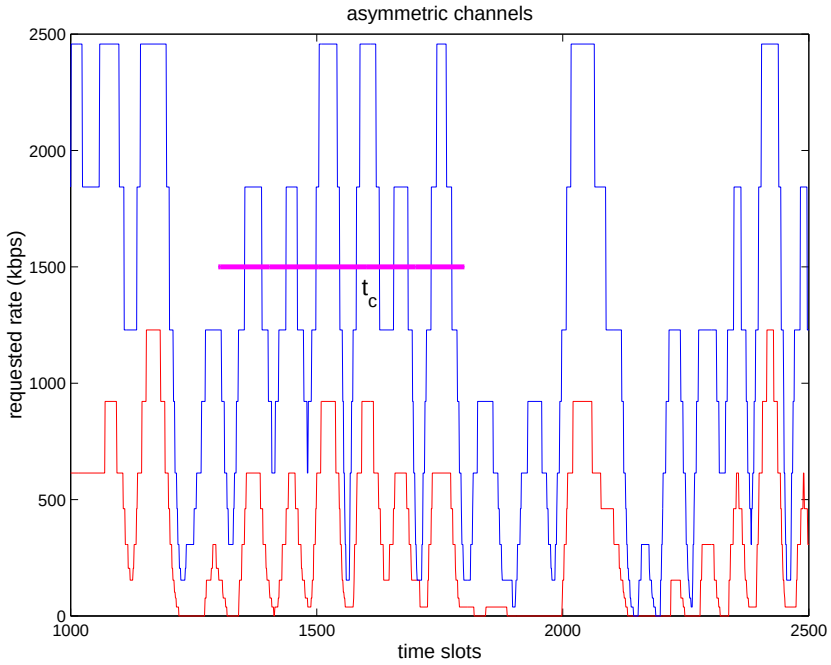
Challenge is to exploit multiuser diversity while sharing the benefits fairly and timely to users with asymmetric channel statistics.

Hitting the Peaks



- Want to serve each user when it is near its peak within a latency time-scale t_c.

Hitting the Peaks



- Want to serve each user when it is near its peak within a latency time-scale t_c.
- In a large system, at any time there is likely to be a user whose channel is near its peak.

Proportional Fair Scheduler

At time slot t, given

1) users' average throughputs $T_1(t), T_2(t), \ldots, T_K(t)$ in a past window.

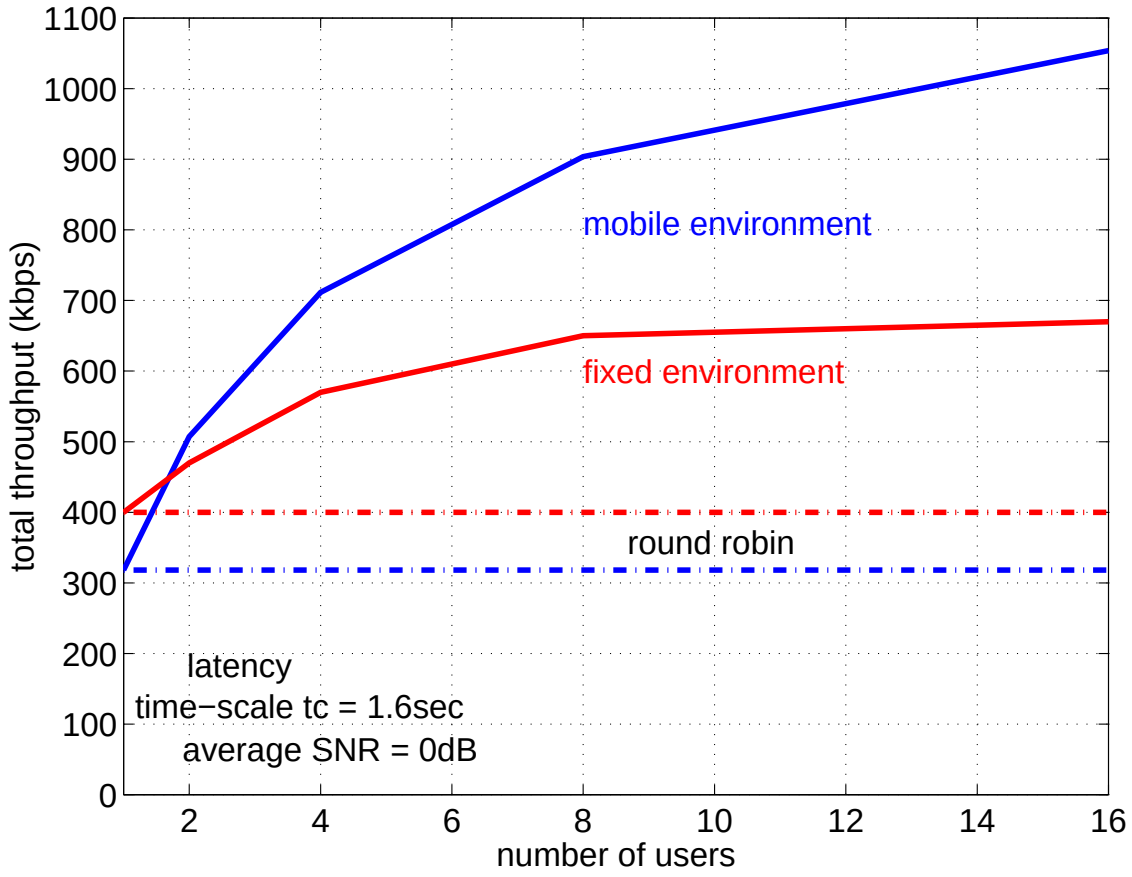
2) current requested rates $R_1(t), R_2(t), \ldots, R_K(t)$

transmit to the user k^* with the largest

$$\frac{R_k(t)}{T_k(t)}.$$

Average throughputs $T_k(t)$ can be updated by an exponential filter with time constant t_c.

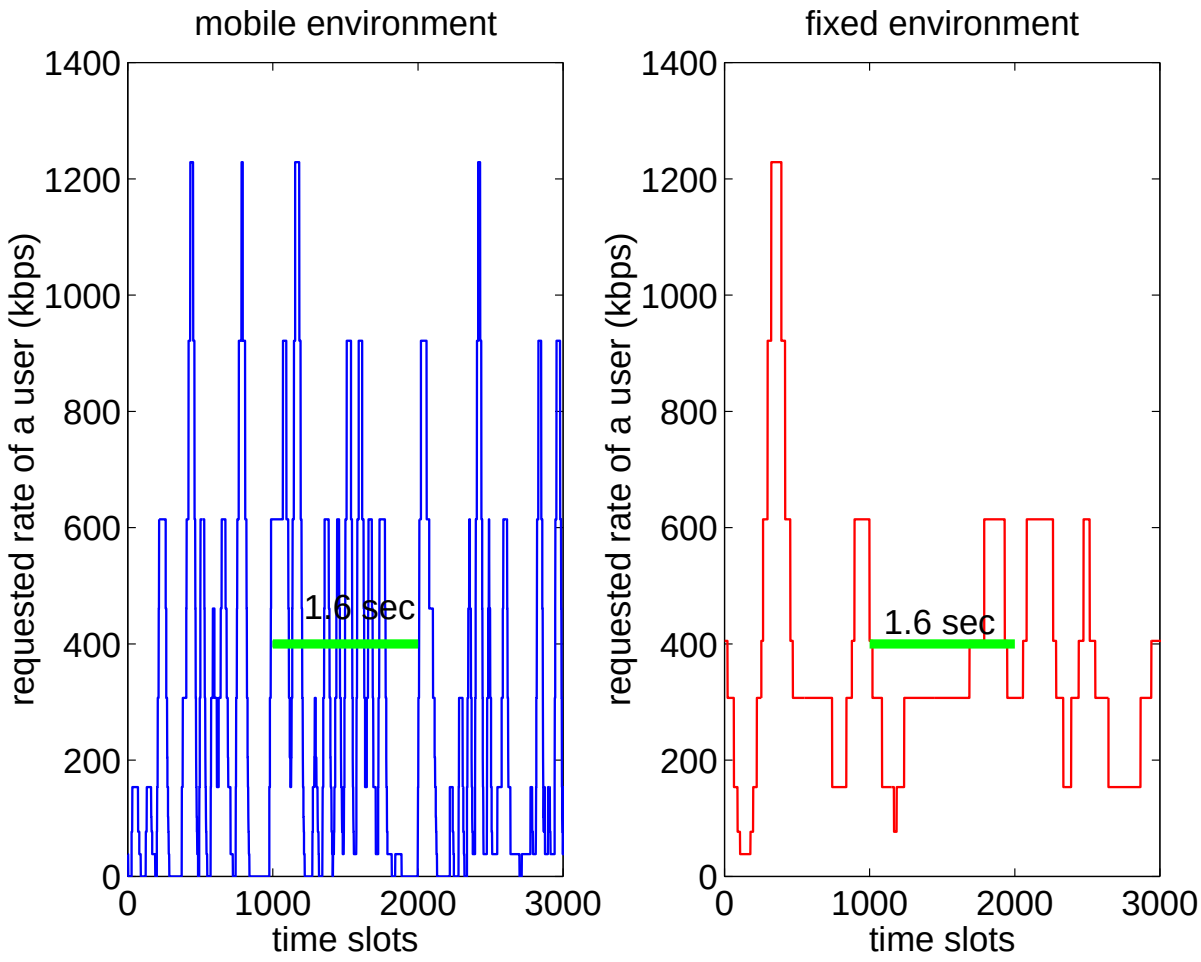
Throughput of HDR Scheduler: Symmetric Users



Mobile environment: 3 km/hr, Rayleigh fading

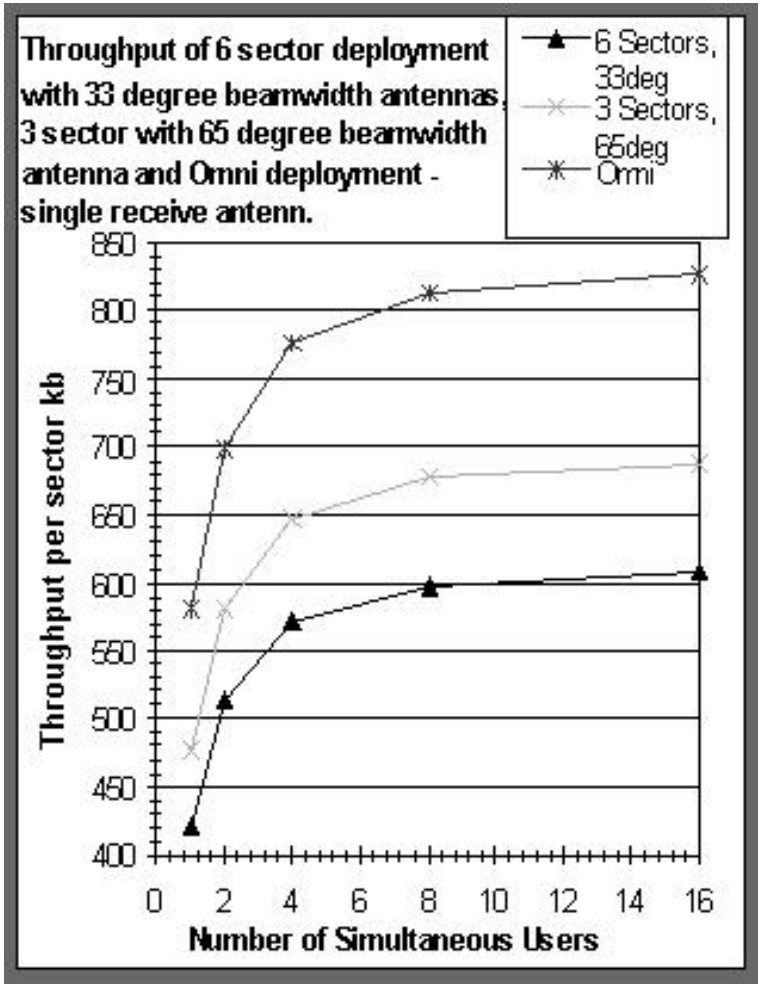
Fixed environment: 2Hz Rician fading with $E_{\text{fixed}}/E_{\text{scattered}} = 5$.

Channel Dynamics



Channel varies faster and has more dynamic range in mobile environments.

Throughput of Scheduler: Asymmetric Users

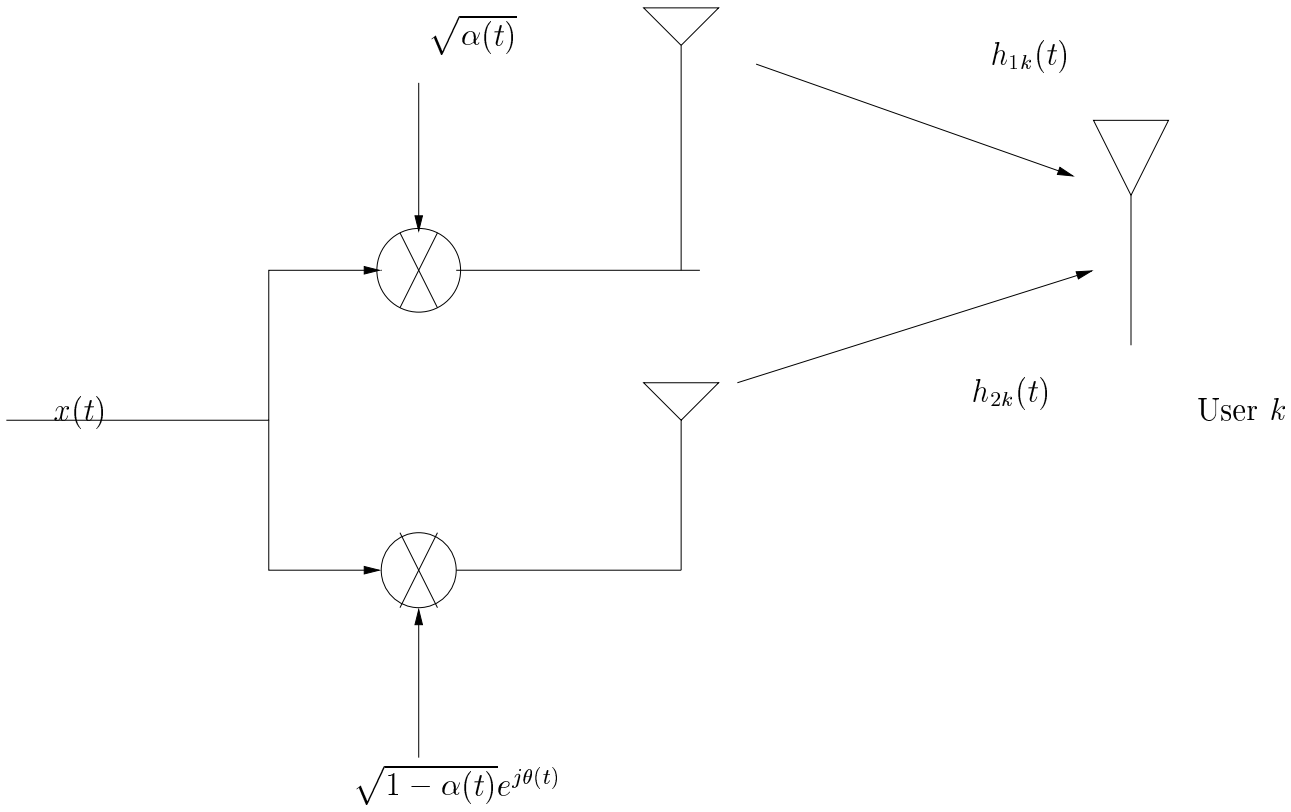


(Jalali, Padovani and Pankaj 2000)

Inducing Randomness

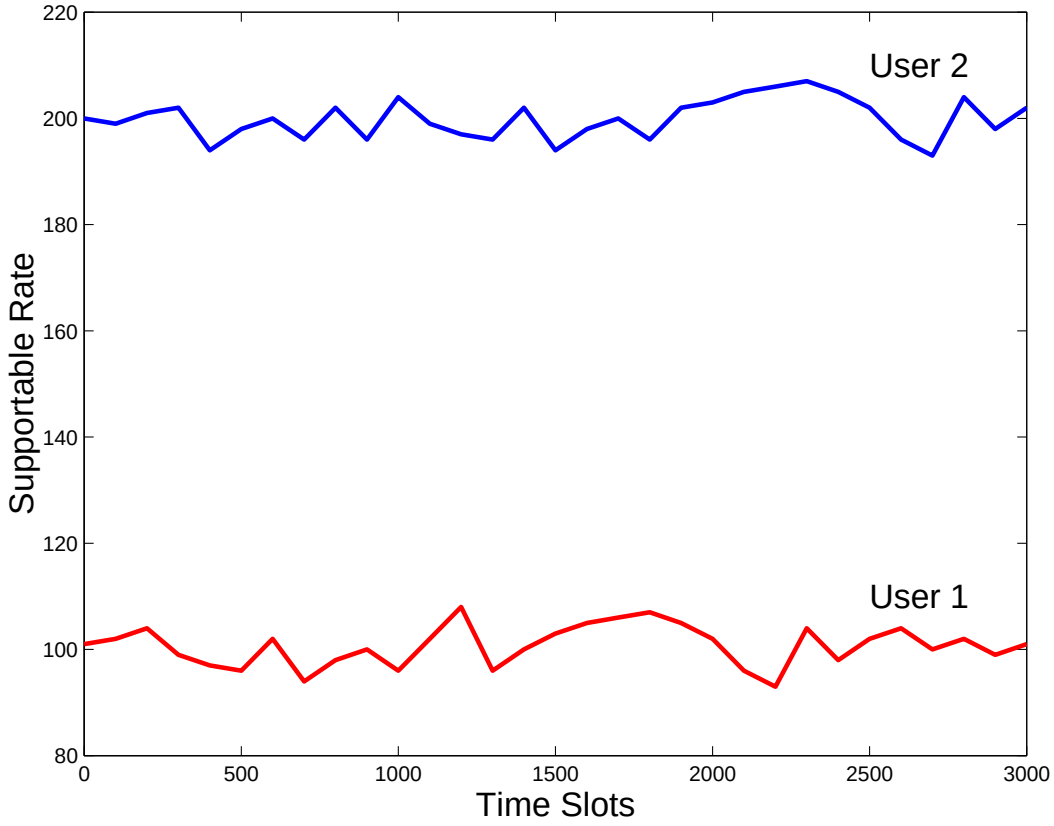
- Scheduling algorithm exploits the nature-given channel fluctuations by hitting the peaks.
- If there are not enough fluctuations, why not purposely induce them?

Dumb Antennas

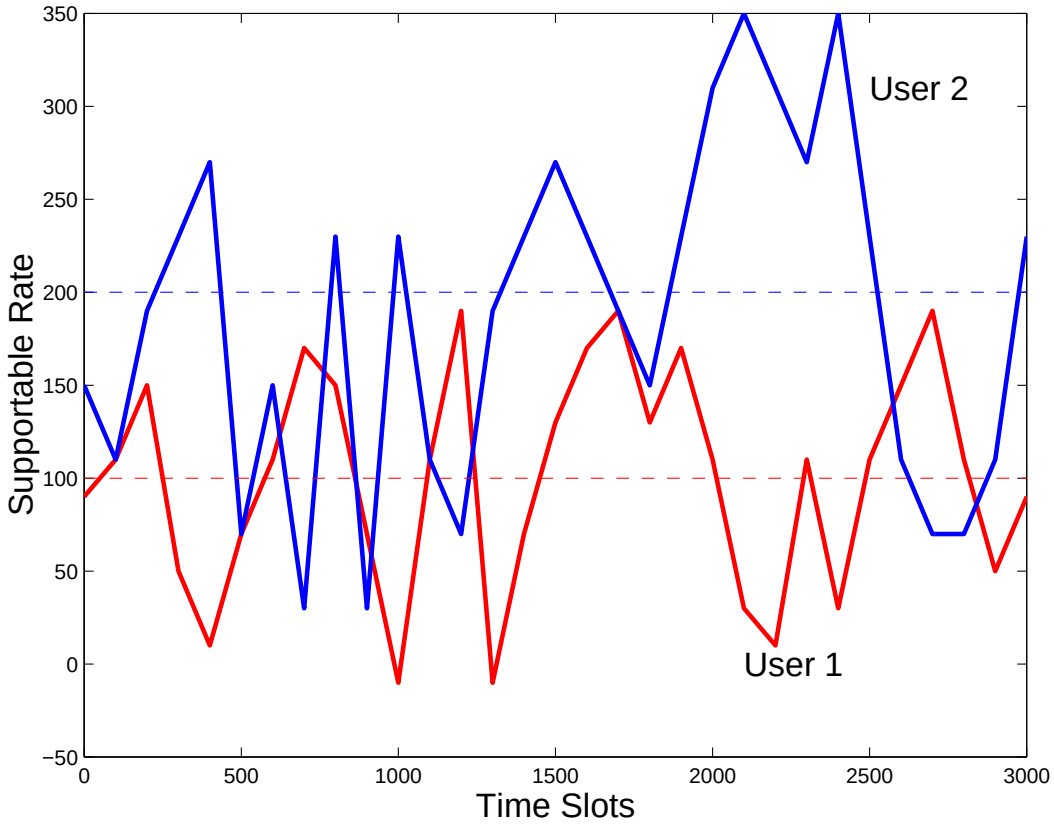


Received signal at user k: $\left[\sqrt{\alpha(t)}h_{1k}(t) + \sqrt{1-\alpha(t)}\exp(j\theta(t))h_{2k}(t)\right]x(t).$

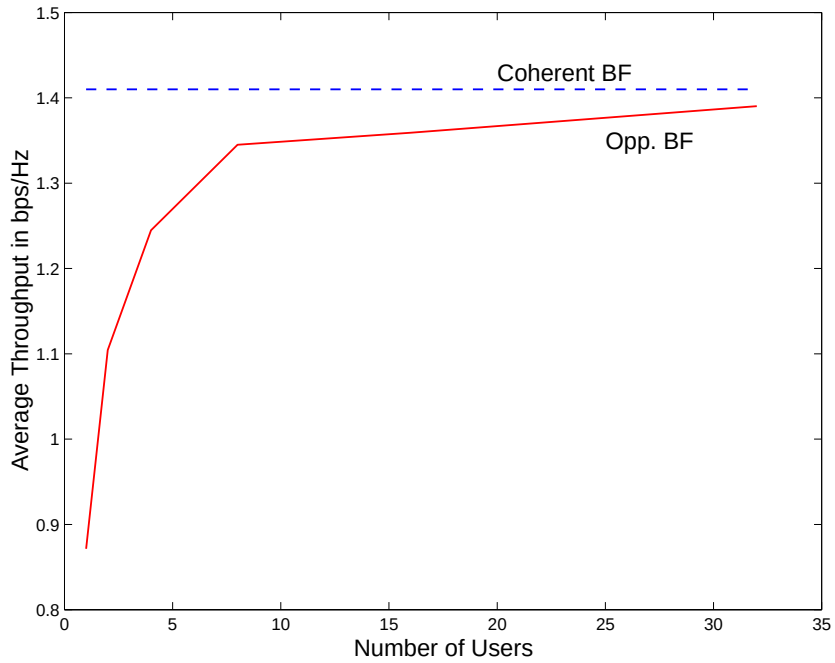
Slow Fading Environment: Before



After


Supportable Rate
Time Slots
User 2
User 1
350
300
250
200
150
100
50
0
−50
0
500
1000
1500
2000
2500
3000

Opportunistic Beamforming: Slow Fading



- Consider first a slow fading environment when channels of the users are fixed (but random).
- Dumb antennas can approach the performance of true beamforming when there are many users in the systems.

Opportunistic versus True Beamforming

- If the gains h_{1k} and h_{2k} are known at the transmitter, then true beamforming can be performed:

$$\begin{aligned} \alpha &= \frac{|\, h_{1k} \,|^2}{|\, h_{1k} \,|^2 + |\, h_{2k} \,|^2} \\ \theta &= \angle h_{1k} - \angle h_{2k} \end{aligned}$$

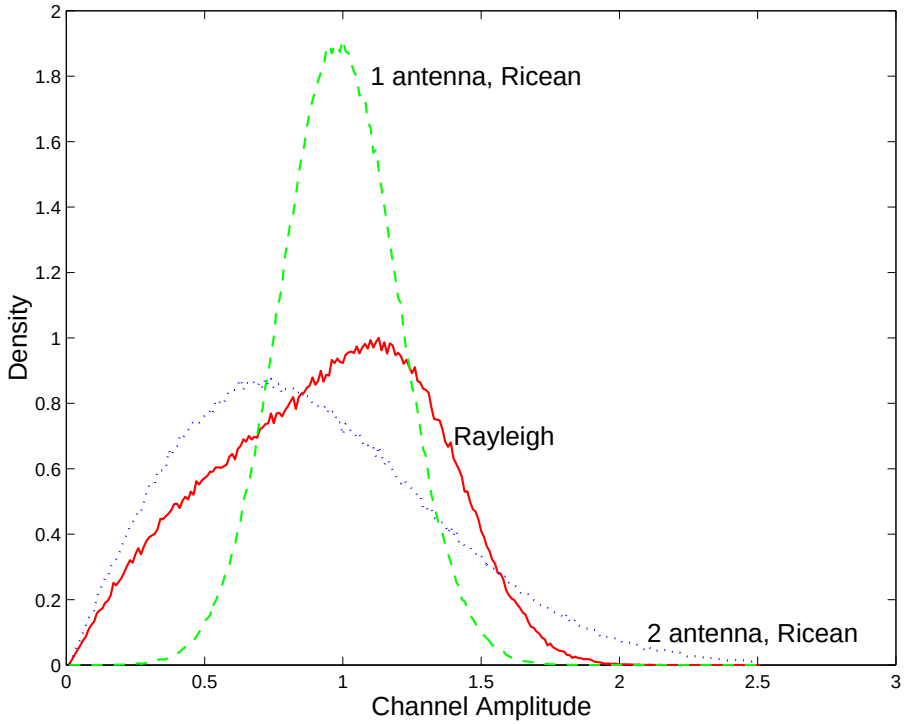
Opportunistic versus True Beamforming

- If the gains h_{1k} and h_{2k} are known at the transmitter, then true beamforming can be performed:

$$\begin{aligned} \alpha &= \frac{|\, h_{1k} \,|^2}{|\, h_{1k} \,|^2 + |\, h_{2k} \,|^2} \\ \theta &= \angle h_{1k} - \angle h_{2k} \end{aligned}$$

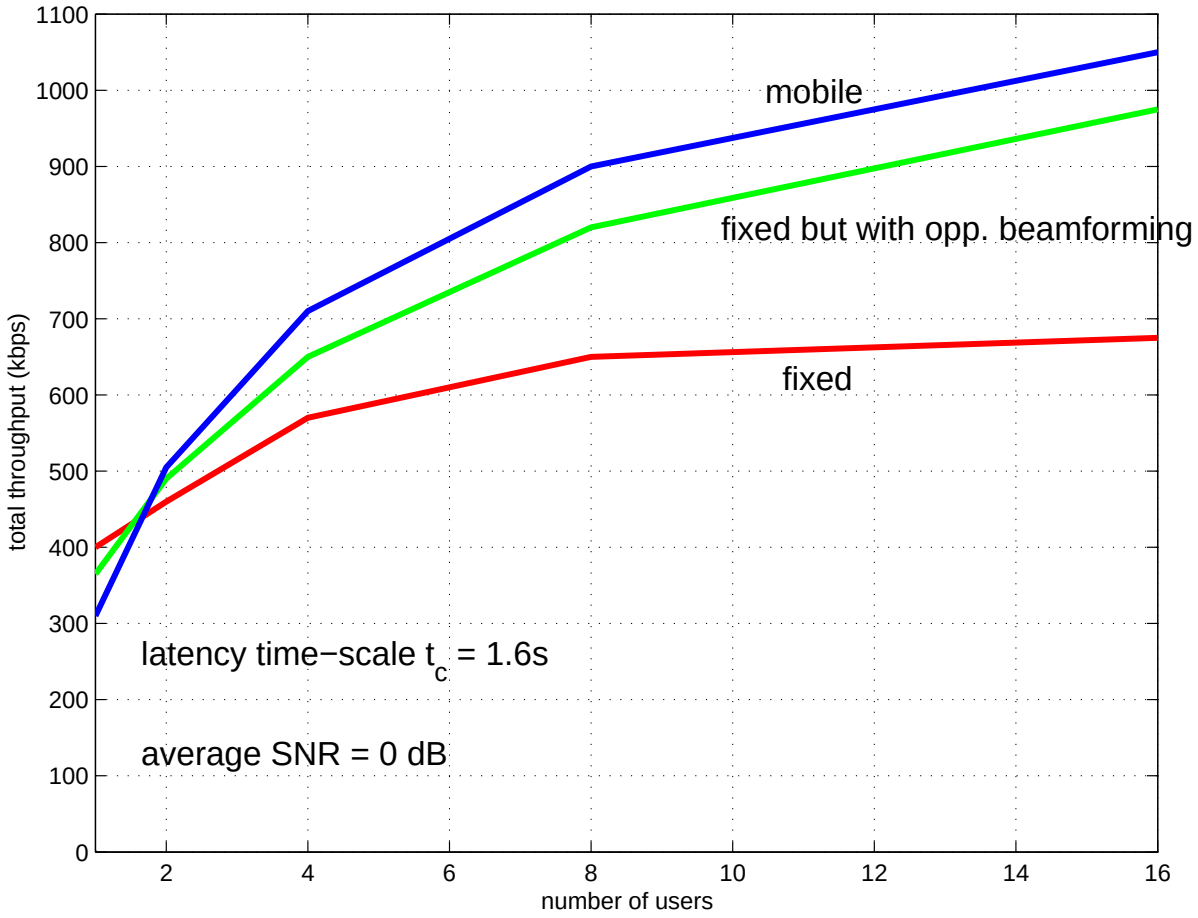
- Dumb antennas randomly sweep out a beam and opportunistically sends data to the user closest to the beam.
- Opportunistic beamforming can approach the performance of true beamforming when there are many users in the systems, but with much less feedback and channel measurements.

Opportunistic Beamforming: Fast Fading



Improves performance in fast fading Rician environments by spreading the fading distribution.

Overall Performance Improvement



Mobile environment: 3 km/hr, Rayleigh fading

Fixed environment: 2Hz Rician fading with $E_{\text{fixed}}/E_{\text{scattered}} = 5$.

Comparison to Space Time Codes

- Space time codes: intelligent use of transmit diversity to improve reliability of point-to-point links.
- In contrast, opportunistic beamforming requires no special multi-antenna encoder or decoder nor MIMO channel estimation.
- In fact the mobiles are completely oblivious to the existence of multiple transmit antennas.
- Antennas are truly dumb, but yet can surpass performance of space time codes.

Cellular System: Opportunistic Nulling

- In a cellular systems, users are scheduled when their channel is strong and the interference from adjacent base-stations is weak.

Cellular System: Opportunistic Nulling

- In a cellular systems, users are scheduled when their channel is strong and the interference from adjacent base-stations is weak.
- Multiuser diversity allows interference avoidance.

Cellular System: Opportunistic Nulling

- In a cellular systems, users are scheduled when their channel is strong and the interference from adjacent base-stations is weak.
- Multiuser diversity allows interference avoidance.
- Dumb antennas provides opportunistic nulling for users in other cells.
- Particularly important in interference-limited systems with no soft handoff.

Traditional CDMA Downlink Design

- orthogonalize users (via spreading codes)

Traditional CDMA Downlink Design

- orthogonalize users (via spreading codes)
- Makes individual point-to-point links reliable by averaging:
  - interleaving
  - multipath combining,
  - soft handoff
  - transmit/receive antenna diversity

Traditional CDMA Downlink Design

- orthogonalize users (via spreading codes)
- Makes individual point-to-point links reliable by averaging:
  - interleaving
  - multipath combining,
  - soft handoff
  - transmit/receive antenna diversity
- Important for voice with very tight latency requirements.

Downlink Design: Modern View

- Shifts from the point-to-point view to a multiuser network view.

Downlink Design: Modern View

- Shifts from the point-to-point view to a multiuser network view.
- Wants large and fast fluctuations of both channel and interference so that we can ride the peaks.

Downlink Design: Modern View

- Shifts from the point-to-point view to a multiuser network view.
- Wants large and fast fluctuations of both channel and interference so that we can ride the peaks.
- Exploits more relaxed latency requirements of data as well as MAC layer packet scheduling mechanisms.