DAYANANDA SAGAR COLLEGE OF ENGINEERING

An Autonomous Institute Affiliated to Visvesvaraya Technological University, Belagavi, Approved by AICTE & ISO 9001 -2008 Certified

Shavige Malleshwara Hills, Kumaraswamy Layout, Bengaluru-560078

Accredited by National Assessment and Accreditation Council (NAAC) with 'A' Grade

DEPARTMENT OF TELECOMMUNICATION ENGINEERING

Accredited by National Board of Accreditation (NBA)

NOTES: MODULE 2

COURSE NAME: MIMO Communication

COURSE CODE: TE814

Faculty Incharge: Dr Vinod B Durdi, Dr Sayed Abdulhayan

CONTENTS:

Capacity and information rates of MIMO channels: Capacity and Information rates of noisy, AWGN and fading channels – Capacity of MIMO channels – Capacity of non-coherent MIMO channels – Constrained signaling for MIMO communications

Capacity and Information Rates of MIMO Channels

One of the primary concerns in the study of noisy channels is the computation of channel capacity, which is defined as the maximum rate at which we can transmit information with an arbitrarily low probability of error. These stablishes a fundamental limit on reliable communications.

Therefore, the channel capacity is widely used for evaluating the performanceof communication systems. Our main objective in this chapter is to study channel capacityissues pertaining to MIMO wireless channels to demonstrate that the use of MIMO systemswill increase the transmission rates significantly without requiring additional poweror bandwidth. This will establish our motivation for the rest of the book, i.e., detailedcoverage of various coding techniques for MIMO communication systems.

Parallel to the capacity of MIMO channels, we will also consider achievable information rates for various communication scenarios. By achievable information rates, we refer to the transmission rates that can be supported with an arbitrarily low probability of error for a given type of input, such as inputs drawn independently from a certain signal constellation (e.g., BPSK, PAM, QAM). These results will be very useful in evaluating practical communication system performance, since in practice we are limited to these types of input signals anyway.

We begin the module with a brief discussion of capacity and achievable information rates, and then specialize the results to the case of AWGN channels. We then extend the discussion to fading channels. After establishing the channel capacity for single-input single-output systems, we move on to the study of wireless MIMO channels and demonstrate the tremendous capacity improvements promised by the use of multiple antennas.

Focus on Capacity of Noisy Channel.
Data rate governs the speed of data transmission. A very important consideration in data communication is how fast we can send data, in bits per second, over a channel. Data rate depends upon 3 factors:

- The bandwidth available
- Number of levels in digital signal
- The quality of the channel – level of noise

Two theoretical formulas were developed to calculate the data rate: one by Nyquist for a noiseless channel, another by Shannon for a noisy channel.

In reality, we cannot have a noiseless channel; the channel is always noisy. Shannon capacity is used, to determine the theoretical highest data rate for a noisy channel:

Capacity = bandwidth * $\log_2(1 + SNR)$

In the above equation, bandwidth is the bandwidth of the channel, SNR is the signal-to-noise ratio, and capacity is the capacity of the channel in bits per second.
Bandwidth is a fixed quantity, so it cannot be changed. Hence, the channel capacity is directly

proportional to the power of the signal, as SNR = (Power of signal) / (power of noise).
The signal-to-noise ratio (S/N) is usually expressed in decibels (dB) given by the formula:

$10 * \log_{10}(S/N)$

so for example a signal-to-noise ratio of 1000 is commonly expressed as:

$10 * \log_{10}(1000) = 30$ dB.

Examples:
Input1 : A telephone line normally has a bandwidth of 3000 Hz (300 to 3300 Hz) assigned for data communication. The SNR is usually 3162. What will be the capacity for this channel?
Output1 : $C = 3000 * \log_2(1 + SNR) = 3000 * 11.62 = 34860$ bps
Input2 : The SNR is often given in decibels. Assume that SNR(dB) is 36 and the channel bandwidth is 2 MHz. Calculate the theoretical channel capacity.
Output2: $SNR(dB) = 10 * \log_{10}(SNR)$
$SNR = 10^{(SNR(dB)/10)}$
$SNR = 10^{3.6} = 3981$
Hence, $C = 2 * 10^6 * \log_2(3982) = 24$ MHz

Capacity and Information Rates of Noisy Channels

Consider the generic communication channel model depicted in Figure 1. The message*W* is encoded to a sequence of *n* channel inputs*X*1*,X*2*, . . .,Xn*, and the sequence*Y*1*, Y*2*, . . . , Yn*is observed at the channel output. The received sequence is then used to recover the transmitted message, producing the estimate ˆ*W* . The channel input is characterized by a joint p.d.f. *p(x*1*, x*2*, . . . ,xn)*, the channel output has p.d.f. *p(y*1*, y*2*, . . . , yn)*, and the channel statistics are described by the conditional p.d.f. *p(y*1*, y*2*, . . . , yn|x*1*, x*2*, . . . , xn)*.

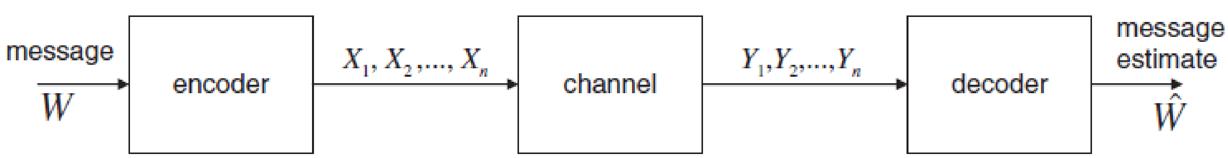


Figure 1 Generic block diagram for a channel coded communication system

Let us now characterize the limits of reliable communications over this generic channel model. In his pioneering work, Shannon (1948) has proved that over a noisy channel, the channel capacity which is defined as the highest transmission rate per channel use that can
be supported with an arbitrarily low probability of error *P(ˆW _= W)* as the block length *n*goes to infinity, is given by the maximum mutual information between the channel input and the channel output (denoted by *I (;   )*,where the maximum is taken over the joint distribution of the input sequence, i.e., *p(x*1*, x*2*, . . . , xn)*).

To prove this result, Shannon employed a random coding argument. The idea is to pick codes with long block lengths at random, and to upper bound the resulting error probability averaged over all possible codes. He showed that if the transmission rate is below the channel capacity, then the average error probability goes to zero as the block length *n* goes to infinity.

Mathematically, the channel capacity C is given by

$$C = \lim_{n\to\infty} \frac{1}{n} \max_{p(x_1,x_2,\ldots,x_n)} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n).$$

For memoryless channels, i.e., if the input−output relationship can be written as

$$p(y_1, y_2, \ldots, y_n | x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p(y_i | x_i),$$

then the channel capacity can be simplified to a single letter representation given by

$$C = \max_{p(x)} I(X; Y),$$

where the mutual information $I(X; Y)$ is defined as

$$I(X; Y) = E\left[\log \frac{p(X, Y)}{p(X)p(Y)}\right],$$

Definition of information rate

Capacity and Information Rates of AWGN and Fading Channels

We now consider widely encountered AWGN and fading channel models in more detail, and characterize their capacity and achievable information rates with constrained signaling. As a basic but important channel model, let us consider the capacity of a widely encountered communication channel, namely, the discrete time AWGN channel. The input–output relationship is given by

$$y = \sqrt{\rho}x + n,$$

where is the channel input, y is the channel output, and n represents the additive noise term. To keep the terminology consistent with the rest of the chapter (and the book), we assume that the input, output and noise terms are complex (two-dimensional). The noise is modeled as a complex Gaussian random variable with zero mean and variance 1⁄2 per dimension. The noise terms at different uses of the channel are assumed to be independent.

The channel input is assumed to satisfy the power constraint $E[|x|2] \leq 1$, thus the constant represents the signal-to-noise ratio at the receiver. It can be shown that for this channel model the input distribution that maximizes the mutual information is zero mean complex Gaussian with variance 1⁄2 per dimension(independent for different uses of the channel), and the resulting channel capacity is given by

$$C = \log(1 + \rho)$$

In practice, using Gaussian channel inputs is not feasible. Therefore although the capacity expression given above is an upper limit for reliable transmission, it may not be possible to achieve.

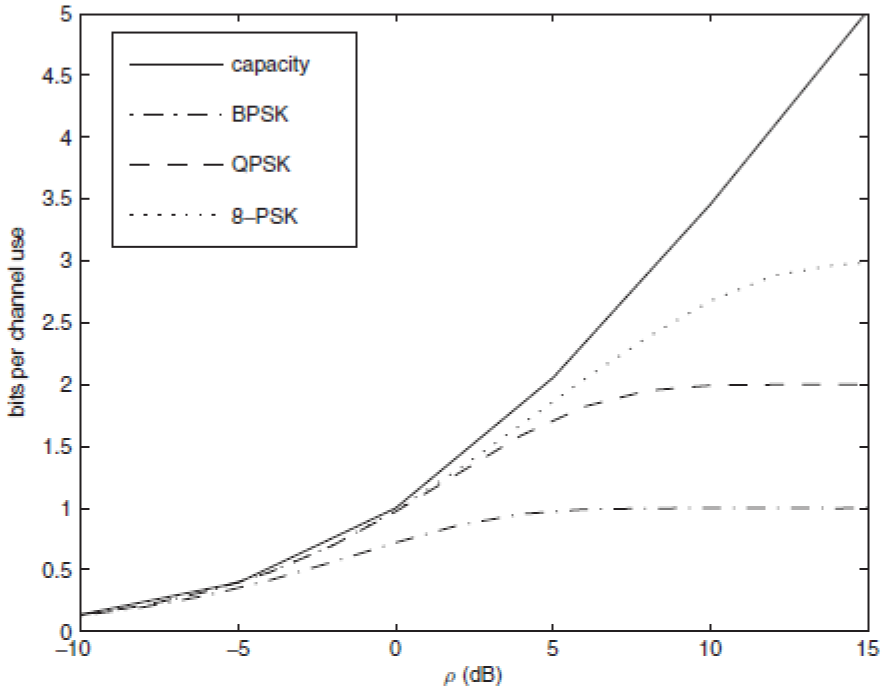


Figure 2 Capacity and information rates for several modulation schemes over AWGN channels.

In Figure 2, we illustrate the capacity of an AWGN channel as well as the achievable information rates for several modulation schemes (BPSK, QPSK and 8-PSK). It is clear that for an *M*-ary modulation scheme, the maximum achievable rate is limited by log2*M*bits per channel use.It is also observed that for low signal-to-noise ratios, the constrained information rate results are very close to the channel capacity obtained by using Gaussian inputs.

Fading Channels

Let us now consider flat fading channels as extensions of AWGN channels studied in the previous section. Consider the flat fading channel model, where the input–output relationship is given by

$$y = \sqrt{\rho} h x + n,$$

where is the complex channel gain. The input x has the same power constraint as in the AWGN case, i.e., $E[|X|2] \leq 1$, and the noise is complex Gaussian with zero mean and variance 1/2 per dimension. For the specific case of Rayleigh fading, the channel gain is zero mean complex Gaussian, or equivalently, $|h|$ is a Rayleigh random variable and $\varphi = \angle h$ is a uniform random variable on $(0, 2\pi)$. We normalize the channel gain such that $E[|h|2] = 1$, thus ρ is the received signal-to-noise ratio. Regarding the availability of the channel state information, we consider only one case, that is, the channel fade coefficient is known precisely at the receiver, and its distribution is known both at the receiver and the transmitter. For fading channels, there are two capacity definitions that are important: ergodic (Shannon type) capacity and outage capacity for non-ergodic channels

Ergodic Capacity and Information Rates:

Assume that the fading channel is ergodic. In this case, the channel coefficients vary overtime, and it is possible to average over their statistics by coding over large blocks of data.For this scenario, the channel capacity is obtained by taking the expected value of capacityof an AWGN channel with the signal-to-noise ratio $|h|2\rho$
That is, defining $z = |h|$, we have

$$C_{fading} = \int_0^\infty \log\left(1 + z^2\rho\right) p(z)dz,$$

or, equivalently by averaging over the p.d.f. of the instantaneous signal-to-noise ratio $\rho\_$ ofthe channel as

$$C_{fading} = \int_0^\infty \log\left(1 + \rho'\right) p(\rho')d\rho',$$

whereρ
_ = $|h|_2\rho$. Clearly, for some cases, this integral can be evaluated analytically.

For instance, we can use **Monte Carlo techniques** for evaluation, i.e., we can generate a large number of realizations for the instantaneous signal to-noise ratio ρ' using the channel statistics, compute $\log(1 + \rho')$ corresponding to each, and average these quantities to estimate the capacity. This is guaranteed to converge using the law of large numbers.

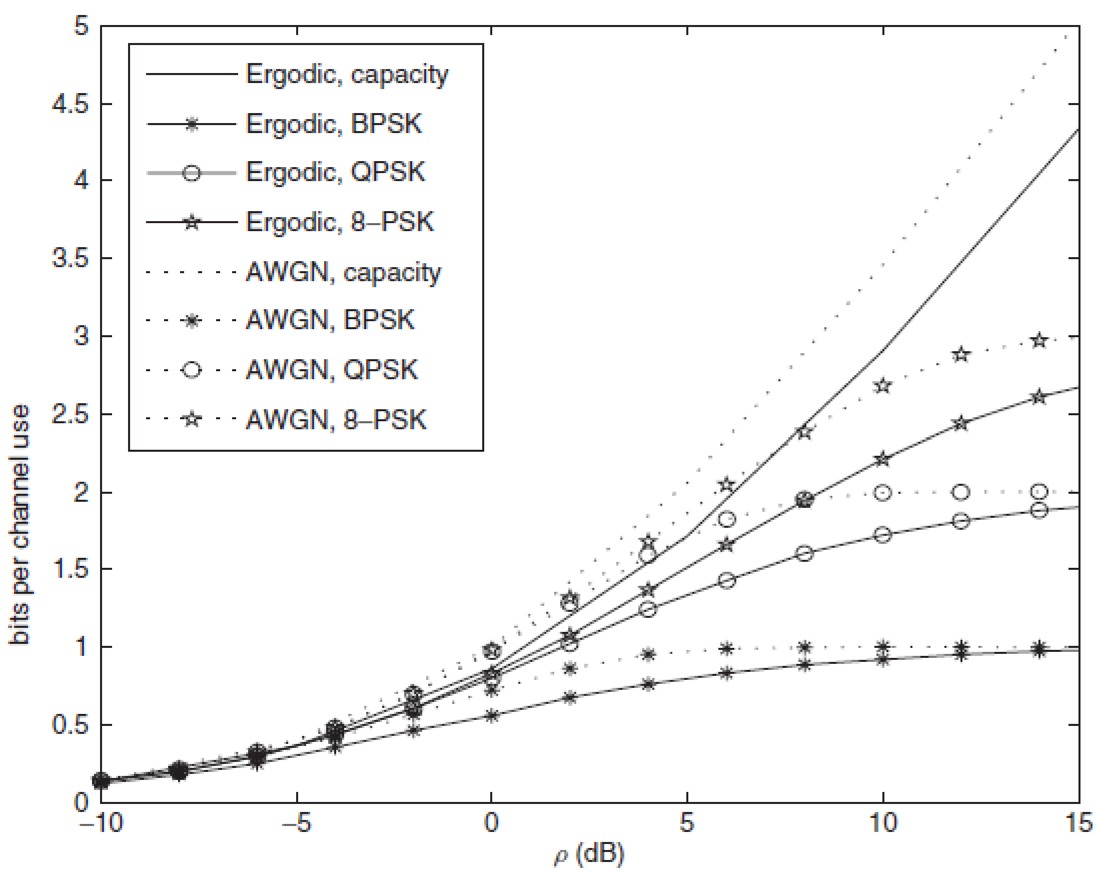


Figure 3 Capacity and information rates for several modulation schemes over ergodic

Rayleigh fading channels

As an example, Figure 3 shows the capacity and achievable information rates with BPSK, QPSK and 8-PSK modulation schemes over ergodic Rayleigh flat fading channels as a function of the average signal-to-noise ratio. We observe that the channel fading deteriorates the capacity and information rates considerably compared with the case ofAWGN channels, as expected. Also, the information rates with BPSK and QPSK inputs are very close to the channel capacity for low signal-to-noise ratios. However, they level off at one bit/channel use and two bits/channel use for high signal-to-noise ratios. We also observe from the figure a large difference between the information rates and capacity results for large signal-to-noise ratios.

Outage Capacity and Outage Information Rates

Let us now assume that fading is non-ergodic, that is, it is not possible to code across different states of the channel, and we have to live with a limited number of realizations for an entire frame of data.

For example quasi static model

Consider a quasi-static fading channel, for a given channel realization, or instantaneous signal-to-noise ratio, the channel capacity is a random variable. For instance, for aquatic-static Rayleigh flat fading channel, the instantaneous channel capacity as a function of the signal-to-noise ratio is given by $C(\rho) = \log(1 + \rho)$, where $\rho\_$ is an exponential random variable. Therefore, for any given rate R, there is a probability that any coding scheme will not be supported reliably over this channel. That is, the outage probability is given by

$$P_{out} = P(C(\rho') < R)$$

To interpret this in another way, for any given outage probability level, there is an outage capacity associated with it, with the interpretation that when the system is not in outage(which occurs with probability $1 - Pout$), this particular transmission rate can be supported.

A similar definition holds for the outage information rates. That is, when a specific channel input is employed, theoutage probability is given by

$$P_{out} = P(I(\rho') < R)$$

for a given rate R where the mutual information is computed for the given signal-to-noise ratio level using specific input constraints.

As a simple example, in Figure 4, we provide the $Pout$ = 10% outage capacity and outage information rates with BPSK, QPSK and 8-PSK signals over a Rayleigh fading channel. Again for low signal-to-noise ratios, the capacity and information rates are similar, but for large signal-to-noise ratios, they differ significantly.

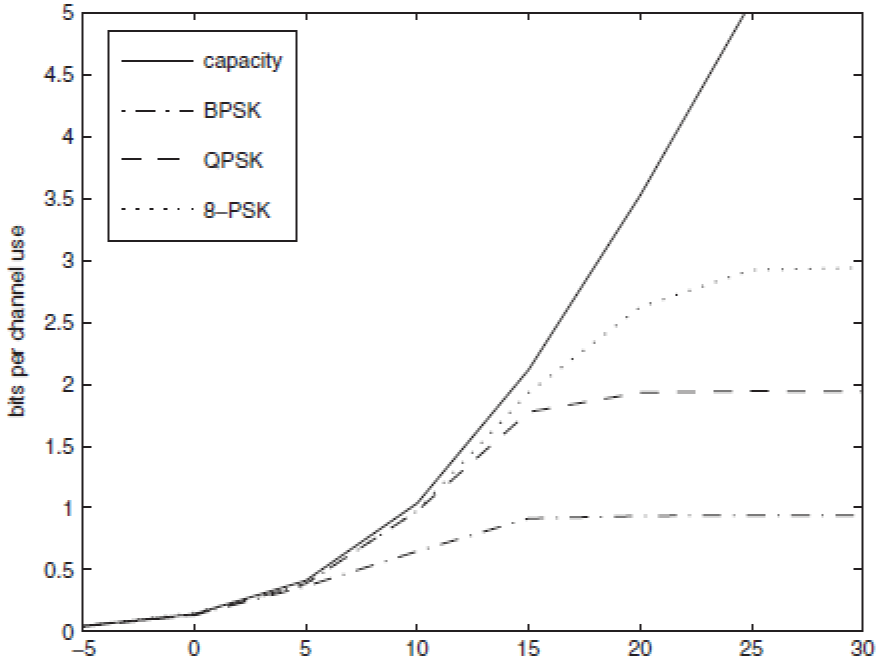


Figure 4 Outage capacity and information rates for quasi-static Rayleigh fading channels.

Other Possible Extensions

Admittedly, the above coverage of fading channel capacity is limited to the case of known channel state information at the receiver, and known input distribution both at the transmitter and the receiver. Other assumptions are certainly possible.

An important but difficult problem is the calculation of channel capacity when neither the transmitter nor the receiver knows the channel state information, but they have access to its distribution.

Another possible extension is the knowledge of the channel state information both at the transmitter and the receiver. In this case, since the transmitter knows when the channel fading is severe, it is possible to allocate the overall power in an optimal manner by performing water filling. The main idea is to allocate more power to more reliable transmissions, and less to the unreliable ones (deep fades) while keeping a total power constraint over a long block of symbols.

Other Points: Examine Capacity and Information Rates of Noisy Channels

In **flat fading**, the coherence bandwidth of the **channel** is larger than the bandwidth of the signal. Therefore, all frequency components of the signal will experience the same magnitude of **fading**. In frequency-selective **fading**, the coherence bandwidth of the **channel** is smaller than the bandwidth of the signal.

1. Capacity of Flat-Fading Channels:

• Depends on what is known about the channel.

• Three cases: 1) Fading statistics known; 2) Fade value known at receiver; 3) Fade value known at transmitter and receiver.

• When only fading statistics are known, capacity is difficult to compute. Only known results are for Finite State Markov channels, Rayleigh fading channels, and block fading.

2. Fading Known at the Receiver:

• Capacity given by $C = \int_0^\infty B \log_2(1+\gamma)p(\gamma)d\gamma$ bps, where p(γ) is the distribution of the fading SNR γ.
• By Jensen's inequality this capacity always less than that of an AWGN channel.
• "Average" capacity formula, but transmission rate is fixed.

3. Capacity with Fading Known at Transmitter and Receiver

• For fixed transmit power, same capacity as when only receiver knows fading.
• By Jensen's inequality, fading reduces capacity w.r.t. AWGN for fixed transmit power.
• Transmit power as well as rate can be adapted.
• Under variable rate and power
$C = \max_{P(\gamma): \int P(\gamma)p(\gamma)d\gamma = S} \int_0^\infty B \log_2\left(1 + \frac{P(\gamma)\gamma}{S}\right) p(\gamma)d\gamma,$, where P(γ) is power adaptation

4. Optimal Power and Rate Adaptation

• Optimal adaptation found via Lagrangian differentiation.
• Optimal power adaptation is a "water-filling" in time: power $P(\gamma) = \gamma_0^{-1} - \gamma^{-1}$ increases with channel quality γ above an optimized cutoff value γ0.
• Rate adaptation relative to $\gamma \geq \gamma_0$ is $B \log_2(\gamma/\gamma_0)$: also increases with γ above cutoff.
• Resulting capacity is $C = \int_{\gamma_0}^\infty B \log_2(\gamma/\gamma_0)p(\gamma)d\gamma.$.
• Capacity with power and rate adaptation not much larger than when just receiver knows channel, but has lower complexity and yields more insight into practical schemes.
• Capacity in flat-fading can exceed the capacity in AWGN, typically at low SNRs.

5. Channel Inversion

• Suboptimal transmission strategy where fading is inverted to maintain constant received SNR. • Simplifies system design and is used in CDMA systems for power control.
• Capacity with channel inversion greatly reduced over that with optimal adaptation (capacity equals zero in Rayleigh fading).
• Truncated inversion: performance greatly improved by inverting above a cutoff γ0.

6.Capacity of Frequency Selective Fading Channels

• Capacity for time-invariant frequency-selective fading channels is hard: result is a "water-filling" of power over frequency.
• For time-varying ISI channels, capacity is unknown.
• Approximate by dividing up the bandwidth subbands of width equal to the coherence bandwidth (same premise as multicarrier modulation).
• We assume independent fading in each subband.
• Capacity in each subband obtained from flat-fading analysis. Power is optimized over both frequencies and time.

Main Points

• Capacity of flat-fading channels depends on what is known about the fading at receiver and transmitter.
• Capacity when only the receiver knows the fading is the same as when the transmitter also

knows but does not adapt power.
• Capacity-achieving transmission scheme uses variable-rate variable-power transmission with power water-filling in time.
• Power and rate adaptation does not significantly increase capacity, and rate adaptation alone yields no increase. These results may not carry over to practical schemes.
• Channel inversion practical but has poor performance. Performance improved by truncating.
• Capacity of frequency-selective fading channels obtained by breaking up wideband channel into sub bands (similar to multicarrier).

Capacity of MIMO Channels

Let us now consider capacity and information rate issues for wireless MIMO systems. Our main objective is to illustrate that the channel capacity improves significantly when MIMOsystems are used without requiring additional bandwidth or power.

We assume that there are *Nt* transmit and *Nr* receive antennas as illustrated in Figure 5,and that there is no inter symbol interference (i.e., the sub-channels are flat fading). The input–output relationship of the MIMO channel is given by

$$y = \sqrt{\rho} x H + n$$

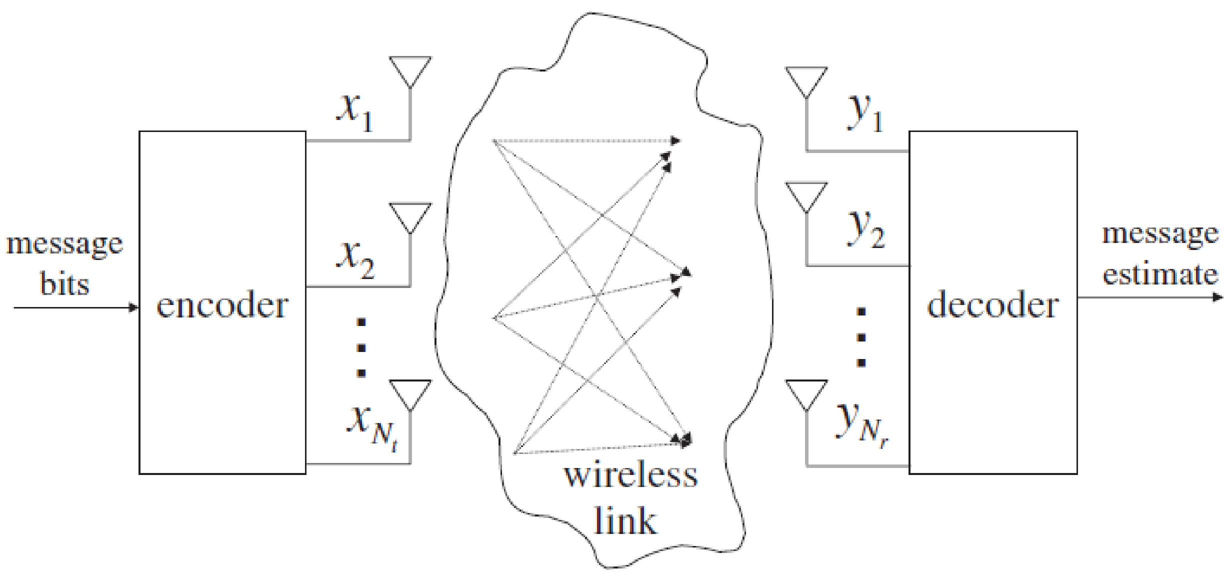


Figure 5. Generic block diagram for a channel coded MIMO communication system

where $\boldsymbol{x}$ is the 1 × *Nt*vector of transmitted signals, $\boldsymbol{H}$ is the *Nt*× *Nr*matrix denoting the channel gains for each transmit and receive antenna pair, $\boldsymbol{n}$ is the 1 × *Nr*vector of independent complex Gaussian noise terms. The signal vector satisfies the power constraint $E[\boldsymbol{xx}H]\leq 1$. Each component of the noise vector has a variance of 1/2 per dimension, and the noise terms at different receive antennas and for different uses of the channel are independent.

In the following, we will consider deterministic MIMO channels and random (fading)MIMO channels. For the case of fading channels, we will study the ergodic and non-ergodic cases separately.

Deterministic MIMO Channels

For deterministic MIMO channels, the assumption is that the channel gain matrix ***H*** is fixed. This may be the case for instance for fixed wireless links when the variations in the environment are negligible. Equivalent channel model

$$\tilde{\boldsymbol{y}} = \sqrt{\rho}\tilde{\boldsymbol{x}}\boldsymbol{\Sigma} + \tilde{\boldsymbol{n}},$$

Therefore, equivalent to the original channel model, we have the following set of parallel channels

$$\tilde{y}_1 = \sqrt{\rho}\sigma_1\tilde{x}_1 + \tilde{n}_1,$$

$$\tilde{y}_2 = \sqrt{\rho}\sigma_2\tilde{x}_2 + \tilde{n}_2,$$

$$\vdots$$

$$\tilde{y}_\nu = \sqrt{\rho}\sigma_\nu\tilde{x}_\nu + \tilde{n}_\nu,$$

$$\tilde{y}_{\nu+1} = \tilde{n}_{\nu+1},$$

$$\vdots$$

$$\tilde{y}_{N_r} = \tilde{n}_{N_r},$$

where the power constraint on the input is *Nti*=1 *E*[| ˜ *xi* |2] ≤ 1 and the noise terms are independent complex Gaussian with variance 1/2 per dimension. Clearly, this is nothing but a set of usual parallel Gaussian channels with independent noise terms having identical variances.

Equal Transmit Power Allocation

Let us first assume that although the MIMO channel is deterministic (fixed ***H***), the transmitter does not have access to the channel matrix (but the receiver does). Therefore the transmitter cannot optimize its power allocation among its antennas. In this case, the capacity of the MIMO channel is given by

$$C = \max_{p(\boldsymbol{x})} I(X;Y),$$

$$= \max_{p(\boldsymbol{x})} H(Y) - H(Y|X).$$

When the channel input is given, the remaining uncertainty in the output is simply theentropy of the noise. Thus we can write

$$C = \max_{p(\boldsymbol{x})} H(Y) - H(N).$$

Since we assume that the transmitter does not have the channel state information, withthe given trace constraint, it selects an input with covariance $\boldsymbol{R}x = 1Nt\boldsymbol{I}Nt$.
Namely thecapacity-achieving input vector is independent complex Gaussian with equal power oneach of the antennas. Therefore, the channel capacity is given bySince the product $\boldsymbol{HHH}$ is positive semidefinite with positive eigenvalues ($\lambda 1, \lambda 2, \ldots, \lambda v$) as the squares of the non-zero singular values of $\boldsymbol{H}$ (i.e., $\lambda 1 = \sigma 21, \lambda 2 = \sigma 22, \ldots, \lambda v = \sigma 2v$), diagonal matrix containing the eigenvalues $\lambda 1, \lambda 2, \ldots, \lambda v$we obtain

$$C = \sum_{i=1}^{\nu} \log\left(1 + \frac{\rho}{N_t}\lambda_i\right)$$

as the MIMO channel capacity.
Therefore, for capacity-achieving signaling, independent Gaussian inputs over the set of Nr parallel channels are used (each with power 1 /Nt), and the overall capacity is simply the sum of the capacities of these parallel channels'

Beamforming

In the previous subsection, we have assumed that the transmitter does not have access to the channel state information, hence it cannot optimize its power allocation. Let us now assume that the transmitter knows the channel matrix $\boldsymbol{H}$.With the ordering of the equivalent parallel channels from good to bad (i.e., with $\lambda 1 \geq \lambda 2 \geq \quad \geq \lambda v$), the transmitter may choose to allocate all its power to the first channel and no power to the remaining links, resulting ina capacity of

$$C = \log(1 + \rho\lambda_1)$$

This is referred to as the beamforming scheme. This is because,in order to accomplish the use of only the best channel, the transmitter emits the samesignal from each of the antennas scaled by a certain coefficient.

Other points on Beamforming

- Beamforming sends the same symbol over each transmit antenna with a different scale factor.
- At the receiver, all received signals are coherently combined using a different scale factor.
- This produces a transmit/receiver diversity system, whose SNR can be maximized by optimizing the scale factors (MRC).
- Beamforming leads to a much higher SNR than on the individual channels in the parallel channel decomposition.
- Thus, there is a design tradeoff in MIMO systems between capacity and diversity.

Waterfilling

When the transmitter has access to the channel state information, the use of equal powerallocation or beamforming (transmitting the same signal from each of the antennas) issuboptimal in general. Instead, the optimal solution is obtained by applying what is knownas the water filling principle as described next.

Denoting the power of the symbol transmitted on the ith parallel channel by Pi (withthe nonzero gain), the capacity is given by

$$C = \max_{\boldsymbol{P}:\sum_{j=1}^{\nu} P_j \leq 1} \sum_{i=1}^{\nu} \log\left(1 + \rho\lambda_i P_i\right)$$

with$\boldsymbol{P} = [P1, P2, \ldots, Pv]$.

Clearly, this capacity is achieved if the inputs are independentcomplex Gaussian.the resulting capacity is

$$C = \sum_{i=1}^{\nu} \left(\log(\mu\rho\lambda_i)\right)^{+}$$

This solution shows that the optimal scheme only uses some of the equivalent parallel channels depending on the signal-to-noise ratio. If the signal-to-noise ratio is very low, only the channel with the best gain will be excited making beamforming optimal. As the signal-to-noise ratio increases, more and more of the parallel channels will be used for transmission.

The capacity of this channel obtained by equal power allocation, beamforming and optimal water filling is shown in Figure 6

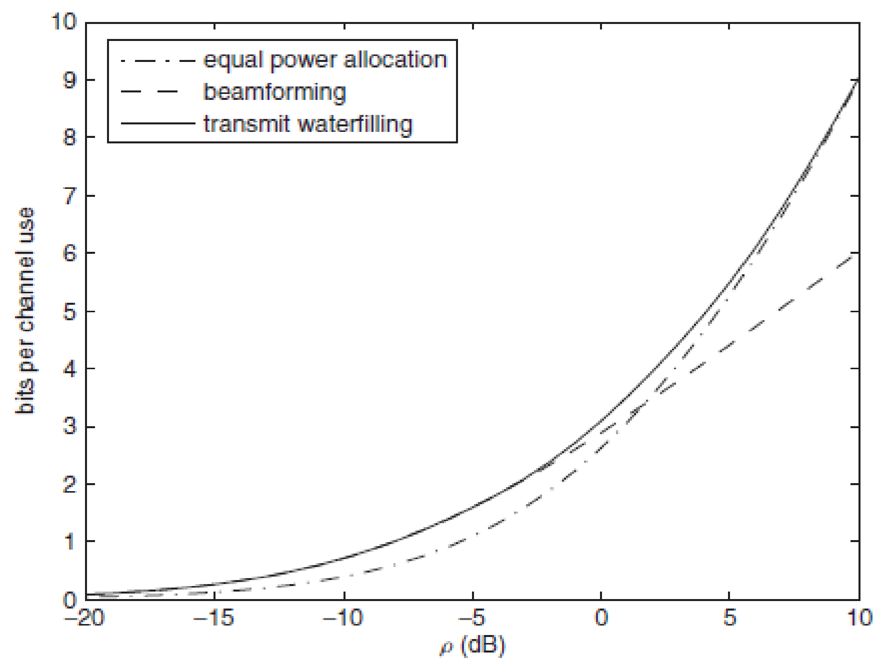


Figure 6: Capacity of channel H1

beamforming is significantly suboptimal, whereas equal power allocation is almost optimalsince excitation of all three parallel channels is the right choice, and the power levels shouldbe almost the same.

Other main on Points on MIMO channel capacity

1. MIMO Systems

- MIMO systems have multiple antennas at the transmitter and receiver.
- The antennas can be used for capacity gain and/or diversity gain.

• MIMO system design and analysis can be complex since it requires vector signal processing.
• The performance and complexity of MIMO systems depends on what is known about the channel at both the transmitter and receiver
2. MIMO Channel Decomposition

• With perfect channel estimates at the transmitter and receiver, the MIMO channel decomposes into RH independent parallel channels, where RH is the rank of the channel matrix (min(Mt , Mr) for Mt transmit and Mr receive antennas under rich scattering).
• With this decomposition there is no need for vector signal processing.
• Decomposition is obtained by transmit precoding and receiver shaping.
• MIMO systems exploit multiple antennas at both TX and RX for capacity and/or diversity gain.
• With both TX and RX CSI, multiple antennas at both transmitter and receiver lead to independent parallel channels.
• With TX and RX CSI, static channel capacity is the sum of capacity on each spatial dimension.
• Without TX CSI, use outage as capacity metric.
• For large arrays, random gains become static, and capacity increases linearly with the number of TX/RX antennas.
• With TX and RX CSI, capacity of MIMO fading channel uses waterfilling in space or space/time - leads to min(Mt , Mr) capacity gain.
• Beamforming transforms MIMO system into a SISO system with TX and RX diversity. Beamforming along direction of maximum singular value

Ergodic MIMO Channels:

Let us now consider MIMO channels where the channel matrix ***H*** is random, but ergodic. Assume that the channel matrix is known by the receiver precisely, but it is unknown tothe transmitter. From the discussion in the previous section, we know that for this case theoptimal signaling uses (spatially and temporally) independent complex Gaussian inputs withequal power. Therefore, taking an approach similar to the case of single-input single-outputsystems, we can write the channel capacity as

$$C = E\left[\log\det\left(\boldsymbol{I}_{N_r} + \frac{\rho}{N_t}\boldsymbol{H}^H\boldsymbol{H}\right)\right]$$

where the expectation is taken over the statistics of the random matrix ***H***

Reciprocity

Using the notation above, ρ denotes the signal-to-noise ratio for each receive antenna. If wedefine $\rho total$as the total signal-to-noise ratio ($\rho total = Nr\rho$), we can rewrite the capacityexpression as

$$C = E\left[\log\det\left(\boldsymbol{I}_{N_r} + \frac{\rho_{total}}{N_t N_r}\boldsymbol{H}^H\boldsymbol{H}\right)\right]$$

Observing that the nonzero eigenvalues of $H_H H$ and HH_H are identical, we can also Write

$$C = E\left[\log\det\left(I_{N_t} + \frac{\rho_{total}}{N_t N_r} H H^H\right)\right]$$

Therefore, we observe that, for Rayleigh fading, if the numbers of transmit and receiveantennas are interchanged, the capacity remains the same (for a fixed total signal-to-noiseratio, *ρtotal*).

Single Transmit Antenna

For the case of a single transmit antenna, i.e., Nt= 1, we simply have receive antennadiversity. The channel gain matrix is a row vector of size 1 × Nr, denoted by ***h***. Thecapacity is given by

$$C = E[\log(1 + \rho\|\boldsymbol{h}\|^2)]$$

which, for the case of independent Rayleigh fading, can be evaluated to be

$$C = \frac{1}{(N_r - 1)!}\int_0^\infty \log(1 + \rho\lambda)\lambda^{N_r - 1} e^{-\lambda} d\lambda.$$

The capacity approaches log(1 + ρNr) as the number of receive antennas increases whichshows that capacity increases only logarithmically with Nr

Single Receive Antenna

For the case of a single receive antenna, the channel matrix is a column vector of sizeNt× 1, denoted by ***h***. The channel capacity is given by

$$C = E\left[\log\left(1 + \frac{\rho}{N_t}\|\boldsymbol{h}\|^2\right)\right]$$

and for independent Rayleigh fading, we have

$$C = \frac{1}{(N_t - 1)!}\int_0^\infty \log\left(1 + \frac{\rho\lambda}{N_t}\right)\lambda^{N_t - 1} e^{-\lambda} d\lambda.$$

which approaches a constant, log(1 + ρ), as the number of transmit antennas is increased.This seems to be contradictory to the reciprocity result

Equal Number of Transmit and Receive Antennas

Let us now consider the case of equal numbers of transmit and receive antennas, i.e.,Nr= Nt= n. For independent Rayleigh fading, the capacity is given by

$$C = n\int_0^{\infty} \log(1+\rho u)\sum_{k=0}^{n-1} L_k(nu)^2 e^{-nu}\,du$$

which, as n is increased, can be well approximated as

$$C \approx n\int_0^{4} \log(1+\rho u)\frac{1}{\pi}\sqrt{\frac{1}{u}-\frac{1}{4}}\,du.$$

This result is very important as it shows that, for a given transmit power level, the MIMO channel capacity scales linearly with the number of receive and transmit antennas used. This is a tremendous increase that motivates the search for good coding techniques for practical wireless MIMO communications

Example

Let us present several examples for the case of independent Rayleigh fading channels. In Figure 7, we show the capacity of independent Rayleigh fading channel for a single transmit antenna as a function of the number of receive antennas for several signal-tonoiseratios. Clearly, the capacity increases logarithmically with *Nr* as expected from our previous discussion.

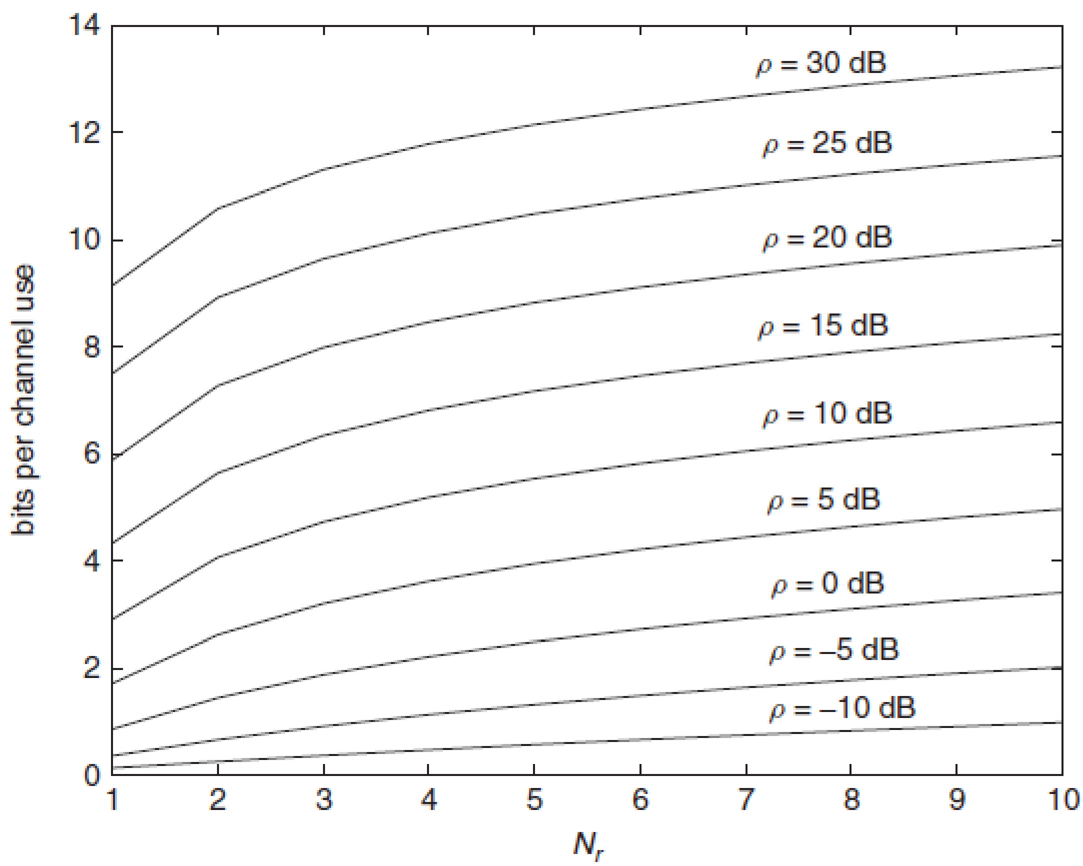


Figure 7 Ergodic capacity of MIMO Rayleigh fading channels with Nt= 1.

For independent Rayleigh fading, the ergodic channel capacity for the case of a single receive antenna as a function of the number of transmit antennas is illustrated in Figure 8. In this case, the capacity approaches a constant as Nt is increased this difference in behavior compared with the Nt= 1 case is only due to the signal-to-noise ratio definition, otherwise the two cases behave the same way

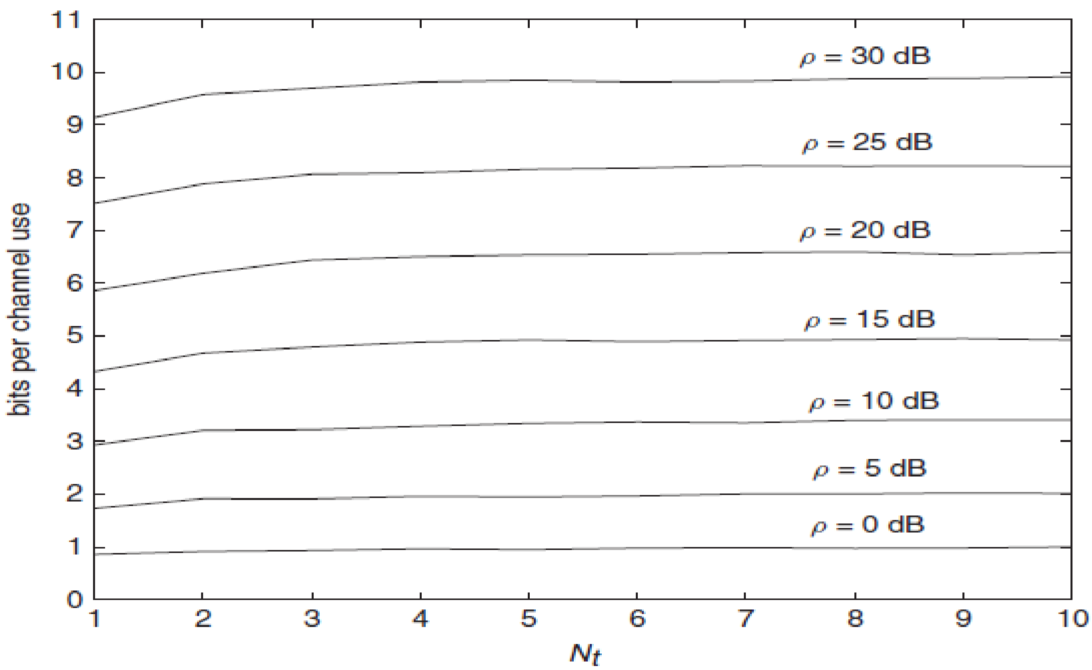


Figure 8 Ergodic capacity of MIMO Rayleigh fading channels with N_r= 1.

Another example is shown in Figure 9 where the channel is again independent Rayleigh fading, but the numbers of transmit and receive antennas are equal. We observe that the channel capacity scales linearly with Nr= Nt= n. clearly there is no increase in the total transmit power or the bandwidth with increasing n. From our earlier discussion, we already expected this result for large n, however, the example shows that it is valid for a lower number of antennas as well. This is a very encouraging result as it shows that the "spatial" dimension can be exploited as another source to improve the capabilities of wireless communication systems considerably.

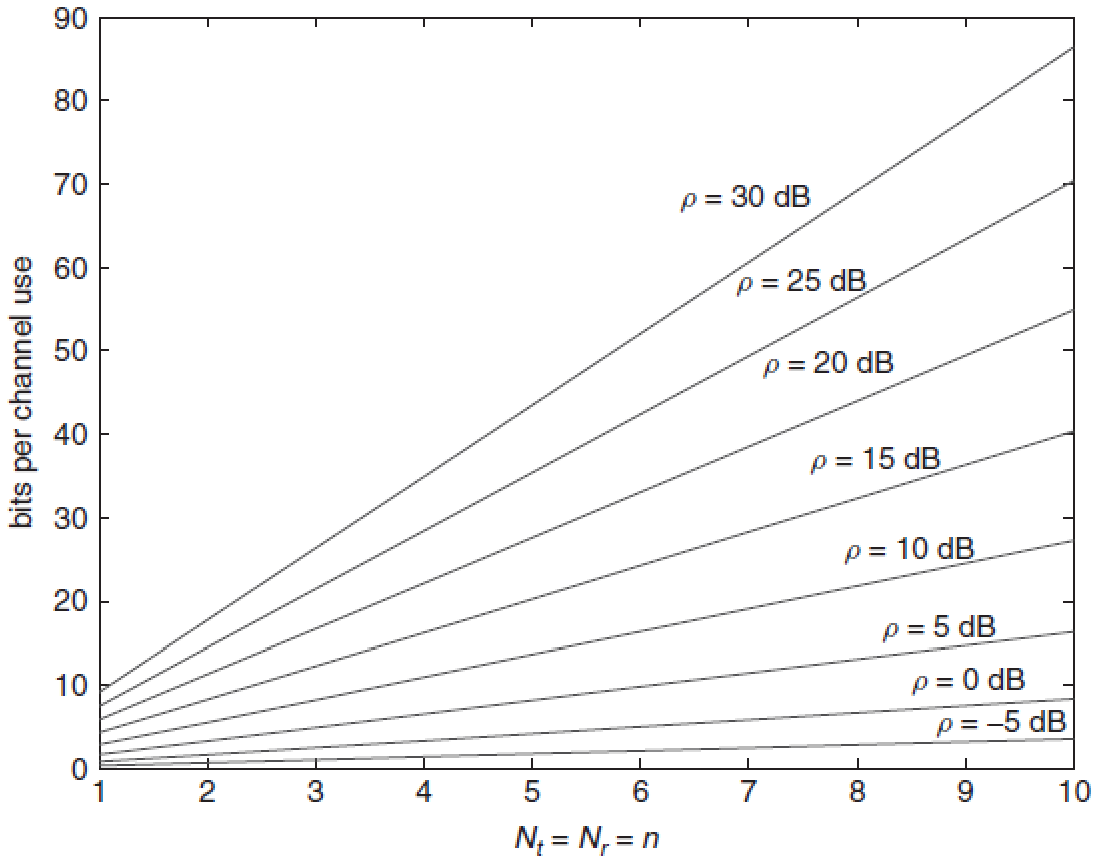


Figure 9 Ergodic capacity of MIMO Rayleigh fading channels with equal number of Transmit and receive antennas

Non-Ergodic MIMO Channels and Outage Capacity:

Let us now turn our attention to the case of non-ergodic fading channels. Specifically, let us assume that the channel is quasi-static, i.e., fading coefficients remain constant over an entire codeword. As discussed earlier in the case of single-input single-output systems, the Shannon type capacity is zero. This is because, regardless of the length of the code chosen, there is a certain probability that the channel cannot support any given fixed rate of transmission. Therefore, we talk about outage capacity. In the discussion of single-input single-output systems, it was clear that the optimal scheme was to use independent Gaussian inputs with variances given by the power constraint.

In the case of multiple antenna systems, for the outage capacity, it is not obvious what the input should be. Telatar (1999) conjectures that the optimal scheme still uses independent Gaussian inputs, but potentially employs a subset of all the available transmit antennas. That is, k antennas, with $k \leq Nt$, are used, and the variance of the Gaussian input for each of them is $1/k$ (no transmit channel state information is assumed). The intuition behind this argument is that, in particular, for very low signal-to-noise ratios, splitting the (already very scarce) power among all the antennas may result in not being able to distinguish the signal from the noise at the receiver end. Instead, using a smaller subset of them with higher power for each transmit antenna could be a better option. For a given signal-to-noise ratio when k of the available antennas are used in transmission, the capacity is a random variable given by

$$C(\rho) = \log\det\left(\boldsymbol{I}_{N_r} + \rho \boldsymbol{H}^H \boldsymbol{R}_x(k)\boldsymbol{H}\right).$$

The outage probability for a given rate of transmission *R* is then given by

$$P_{out} = \min_{k=1,2,\ldots,N_t} P\left(\log\det\left(\boldsymbol{I}_{N_r} + \rho \boldsymbol{H}^H \boldsymbol{R}_x(k)\boldsymbol{H}\right) \le R\right)$$

Single Transmit Antenna: Receive Diversity

For some simple cases, it is possible to analytically evaluate *Pout* . For instance for the case of a single transmit antenna, we obtain

$$P_{out} = \frac{1}{(N_r - 1)!}\,\gamma\left(N_r, \frac{2^R - 1}{\rho}\right)$$

Single Receive Antenna: Transmit Diversity

$$P_{out} = \min_{k=1,2,\ldots,N_t} \frac{1}{(k-1)!}\,\gamma\left(k, \frac{k\;2^R - 1}{\rho}\right).$$

Determine the Capacity of Non Coherent MIMO Channels.

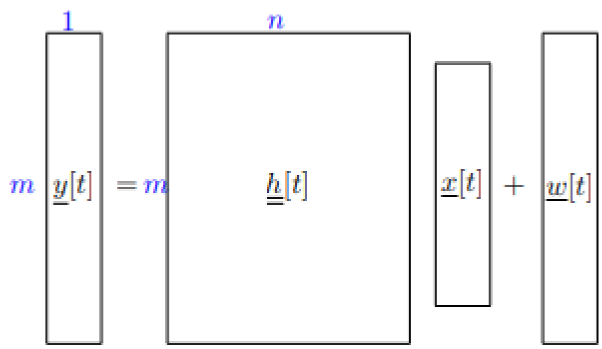
In non-coherent wireless communications, the receiver (neither the transmitter) are not aware of the actual realization of the channel coefficients. The encoding and decoding should be performed by using only the statistics of the coefficients.

Flat Fading Flat fading is referred to the model in which the changes in channel are so slow that the channel coefficients are assumed to be constant over each block of communications.

i.e., h[t] = h for t = 1, 2, · · · , T.

In MIMO communications there are n transmitting antennas and m receive antennas each receiving a linear combination of the transmitted signals from all TX antennas. The channel shown in figure 10 is modeled as following:

$$
\begin{aligned}
y_j[t] &= \sum_{i=1}^{n} h_{i,j}^{*}[t]x_i[t] + w[t], \qquad \text{for } t = 1, \cdots, T \\
\underline{y}[t] &= \underline{\underline{h}}[t]\underline{x}[t] + \underline{w}[t]
\end{aligned}
$$



In *flat fading* MIMO communications, the matrix $\underline{\underline{h}}[t]$ is constant over a block of communication i.e., $\underline{\underline{h}}[t] = \underline{\underline{H}}$ for $t = 1, 2, \cdots, T$.

$$\underline{y}[t] = \underline{\underline{H}}\ \underline{x}[t] + \underline{w}[t] \qquad \text{for } t = 1, \cdots, T$$

$$\underline{\underline{Y}}_{m\times T} = \underline{\underline{H}}_{m\times n}.\underline{\underline{X}}_{n\times T} + \underline{\underline{W}}_{m\times T}$$

Figure 10 channel model

Degrees of Freedom (DOF=d) The received signal in a block of length T in all antennas shown in figure 11, Y ∈Cm×T lives in a space of complex dimension of mT. But not all these dimensions are affected by the input X. The number of dimensions in received signal that are only affected by the input X (i.e.,

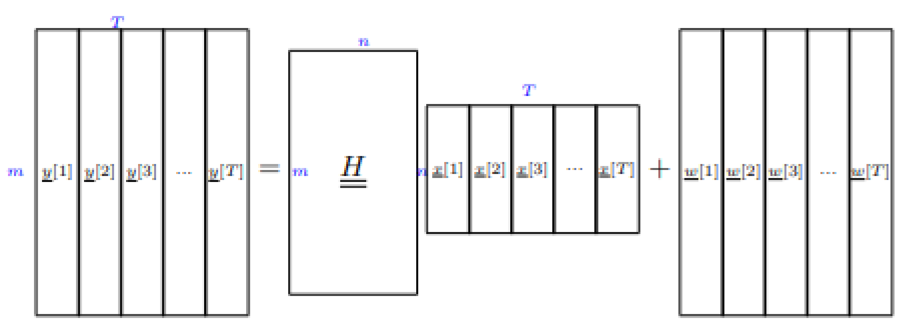


Figure 11. The received signal

the dimension of subspace that Y lives in, fixing all parameters and changing only X) is called the degrees of freedom of the system. The importance of this parameter is in the fact that in high SNR regime, degrees of freedom of the system is the pre-log factor of capacity of the system
i.e., C(SNR) ≈ d log SNR Intuitively, in high SNR regime, the constraint over the input power is not

the limiting factor.
The limiting factor in communications is the number of dimensions of the received signal that we have control over by changing and designing the input signal.

Constrained Signaling for MIMO Communications

Since the capacity of a MIMO channel (in the presence of Gaussian noise) is achieved with Gaussian inputs, using such constellations will prevent us from achieving the channel capacity in general. For instance, if we use BPSK signaling, regardless of the signal-to-noise ratio (and the fading statistics), we can transmit at most Nt bits per channel use. However, it is important to find the limits of reliable transmission when constrained signaling is employed as well. This was rather easy to compute for the case of AWGN channels via a single integration, however for the case of fading channels, in particular for MIMO systems, it is not straightforward to do the computations analytically. In order to compute the information rates with constrained signaling, we go back to the original capacity expression, that is,

$$C = \max_{p(x)} H(Y) - H(N).$$

We note that the maximization should be done using the specific input constraints. This calculation is in general difficult, however we note that

$H(Y) = E[-\log p(Y)],$ where the expectation is over the statistics of the random vector Y. This expectation can be easily calculated using Monte Carlo techniques, i.e., by generating a large number of realizations of the channel output (of course, using the input constraints), computing the log of the joint probability of the realization, and averaging them.

We observe that for low signal-to-noise ratios, the information rates with constrained i.u.d. (independent uniformly distributed, or i.u.d.) inputs are very close to the channel capacity. This is be- cause, when the noise variance is very high, the channel outputs with the constrained inputs are approximately Gaussian distributed, thus the resulting mutual information is almost optimal. However, when the signal-to-noise ratio is increased, the channel capacity grows with the signal-to-noise ratio while the information rates are limited by their maximum values determined by the constellation size and the number of transmit antennas, hence the large gap between them. It is interesting to observe that, there is an almost linear increase of the achievable information rates (or, constrained capacity) with the number of antennas n.

List out the properties of following:

i) Noisy Channels
ii) AWGN Channels
iii) Fading Channels

Noisy Channel:

Noisy coding theorem

The scenario for the noisy coding theorem is the same as that for the noiseless coding theorem, with the addition of a noisy channel between the compressor (renamed as the encoder) and the decompressor (renamed as the decoder). Formany years, it was believed that as the transmission rate through a noisy channel is increased, the error rate also increased, no matter how slow the rate was. However, what Shannon showed, in this second theorem of his, is that as long as the transmission rate is below a certain maximum value, known as the capacity of the channel *C*, messages could be transmitted with *zero* error, asymptotically in the message length. This capacity isC=maxp(X)I(X;Y)

$$\frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}.$$

The noisy coding theorem states that the rate *C* is *optimal*, i.e. that rates larger than *C* are not possible without inducing unrecoverable errors as $n \to \infty$, and *achievable*, meaning a scheme exists which can transmit at any rate less than or equal to *C*, in the limit as $n \to \infty$.

FANO'S INEQUALITY AND THE DATA PROCESSING INEQUALITY

Fano's inequality is used to find a lower bound on the error probability of any decoder; it relates the average information lost in a noisy channel to the probability of the categorization error. We first introduce the concept of a Markov chain, then state Fano's inequality, and finally prove the data processing inequality, a result with many applications in classical information theory.
Random variables, X, Y, and Z, form a Markov chain denoted as X → Y → Z, if the conditional probability distribution of Z depends only on Y (it is independent of X):

The random variables, X1,X2,X3,…,*Xn*, could represent the states of a system and, as we know, the Markov property means that the past and the future state are conditionally independent given the present state; this implies that all terms in this sum except the first are equal to zero.
Consider now a random variable, *X*, with the probability density function, *px(x*), and let |χ| denote the number of elements in the range of *X*. Let *Y* be another random variable related to *X*, with the conditional probability, *pY|X(y|x*). Intuitively, we expect that the error when we use *Y* to estimate *X* will be small when the conditional entropy, *H(X|Y*), is small. Our intuition is quantified by Fano's inequality.

AWGN Channels

Additive white Gaussian noise (**AWGN**) is a basic noise model used in Information theory to mimic the effect of many random processes that occur in nature. The modifiers denote specific characteristics:

- ***Additive*** because it is added to any noise that might be intrinsic to the information system.
- ***White*** refers to the idea that it has uniform to power across the frequency band for the information system. It is an analogy to the color white which has uniform emissions at all frequencies in the visible spectrum.

- ***Gaussian*** because it has a normal distribution in the time domain with an average time domain value of zero.

Wideband noise comes from many natural noise, such as the thermal vibrations of atoms in conductors (referred to as thermal noise or Johnson–Nyquist noise), shot noise, black-body radiation from the earth and other warm objects, and from celestial sources such as the Sun. The central limit theorem of probability theory indicates that the summation of many random processes will tend to have distribution called Gaussian or Normal.

AWGN is often used as a channel model in which the only impairment to communication is a linear addition of wideband or white noise with a constant spectral density (expressed as watts per hertz of bandwidth) and a Gaussian distribution of amplitude. The model does not account for fading, frequency selectivity, interference, nonlinearity or dispersion.

The AWGN channel is a good model for many satellite and deep space communication links. It is not a good model for most terrestrial links because of multipath, terrain blocking, interference, etc.

Fading Channels

Rayleigh and Rician fading channels are useful models of real-world phenomena in wireless communications. These phenomena include multipath scattering effects, time dispersion, and Doppler shifts that arise from relative motion between the transmitter and receiver. This section gives a brief overview of fading channels and describes how to implement them using the toolbox.

This figure 12 depicts direct and major reflected paths between a stationary radio transmitter and a moving receiver. The shaded shapes represent reflectors such as buildings.

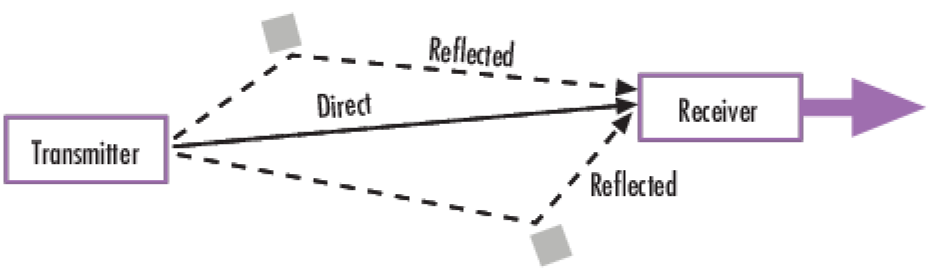


Figure 12 depicts direct and major reflected paths between a stationary radio transmitter and a moving receiver

The major paths result in the arrival of delayed versions of the signal at the receiver. In addition, the radio signal undergoes scattering on a *local* scale for each major path. Such local scattering typically results from reflections off objects near the mobile. These irresolvable components combine at the receiver and cause a phenomenon known as *multipath fading*. Due to this phenomenon, each major path behaves as a discrete fading path. Typically, the fading process is characterized by a Rayleigh distribution for a non line-of-sight path and a Rician distribution for a line-of-sight path.

The relative motion between the transmitter and receiver causes Doppler shifts. Local scattering typically comes from many angles around the mobile. This scenario causes a range of Doppler shifts, known as the *Doppler spectrum*. The *maximum* Doppler shift corresponds to the local scattering components whose direction exactly opposes the trajectory of the mobile.

The channel filter applies path gains to the input signal, Signal in. The path gains are configured based on settings chosen in the fading channel object or block.

Differences between Noisy and Noiseless Channels.

Noiseless Channels:

- Noiseless Channels are Ideal Channel in which no frames are lost, duplicated or corrupted.
- Protocols used in Noiseless Channels for flow control are
  (i) Simplest Protocol and (ii)Stop and Wait Protocol
- In Noiseless Scheme, Channel is error free and doesn't require Error Control Mechanism.
- Noiseless Channels are Nonexistent.

Noisy Channels:

- Noisy Channels are having Noise of various ranges.
- Frames are lost, duplicated and corrupted in Noisy channels
- Protocols used in Noisy Channels for flow control are

(i) Stop and Wait Protocol Automatic Repeat Request
(ii) Go-Back-N Automatic Repeat Request
(iii) Selectively Repeat Automatic Repeat Request

- In Noisy Scheme, Channel is error prone and requires Error Control Mechanism.
- Noisy Channels are all Real Time Signals.

Differences between AWGN and (Rayleigh,Rician)Fading Channels.
AWGN noise: AWGN is a noise that affects the transmitted signal when it passes through the channel. It contains a uniform continuous frequency spectrum over a particular frequency band.
Rayleigh Fading: When no LOS path exists in between transmitter and receiver, but only have indirect path than the resultant signal received at the receiver will be the sum of all the reflected and scattered waves.
Rician Fading: It occurs when there is a LOS as well as the non-LOS path in between the transmitter and receiver, i.e. the received signal comprises on both the direct and scattered multipath waves.
(i)For Mobile Communication experiment we get the following Results:
For execution of digital communication using BFSK and MQAM and implementation of wireless communication used DS-CDMA and SFH-CDMA for AWGN, Rician Fading and Rayleigh fading the results are as follows:
The AWGN channel will be better realization than Rayleigh fading channel because the BER values lower than the BER value of fading use Rayleigh channel while usingRician fading better realization than AWGN and Rayleigh channels because BER value of Rician channel less than Rayleigh fading and upmost than AWGN channel.
(ii)For OFDMA Wimax the following results we get:
1. The performance of AWGN channel is the best of all channels as it has the lowest bit error rate (BER) under QAM, 16-QAM & 64-QAM modulation schemes. The amount of noise occurs

in the BER of this channel is quite slighter than fading channels.
2. The performance of Rayleigh fading channel is the worst of all channels as BER of this channel has been much affected by noise under QAM, 16-QAM & 64-QAM modulation schemes.
3. The performance of Rician fading channel is worse than that of AWGN channel and better than that of Rayleigh fading channel. Because Rician fading channel has higher BER than AWGN channel and lower than Rayleigh fading channel. BER of this channel has not been much affected by noise under QAM, 16-QAM & 64-QAM modulation schemes.

Acknowledgment: This material is based on the text book authored by Tolga M. Duman and Ali Ghrayeb, "Coding for MIMO Communication systems", John Wiley & Sons, West Sussex, England, 2007. Some additional material are taken and/or inspired by material from various paper and / or electronic resources.

Multiple choice Questions:

a) Channel Capacity is the tight ------ ---- on the rate at which information can be reliably transmitted over a communication channel.

i)Maximum ii) lower bound **iii) upper bound** iv)Minimum

b) At a SNR of 0 dB Signal power = ----- ----

i) Maximum power **ii) Noise power** iii) BW power iv) Minimum power

c) The data rate is directly proportional to the number of ----------

i) Noise level ii) Frequency Level iii) Amplitude level **iv) Signal levels**

d) Increasing the levels of a signal may ------ the reliability of the system

i) Reduce ii) Increase iii) not effect iv) improve

e) Maximum bit rate = 2 × Bandwidth × $\log_2 V$ is Nyquist bit rate for ---- ---(*V is the number of discrete levels in the signal*)

i) Signaling Channel ii) Imperfect Channel **iii) Perfect Channel** iv) Noisy Channel

f) Shannon's Capacity gives the theoretical maximum data rate or capacity of a --------

i) Signaling Channel ii) Imperfect Channel iii) Perfect Channel **iv) Noisy Channel**

g) Information is an ----------- in uncertainty or entropy.

i) increase ii) decrease iii) same iv) no effect

h) More certain or deterministic the event is, the less ---------- it will contain.
i) BW **ii) information** iii) Energy iv)invalid data
i) Information entropy tells how much --------- there is in an event
i) BW **ii) information** iii) Energy iv) invalid data
j) The symbol for entropy is an -----
i) Erlang ii) K iii) S iv) N
k) ---------- is the sum of the internal energy added to the product of the pressure and volume of the system
i) Erlang ii) Entropy **iii) Enthalpy** iv) BW
l) --------- reflects the capacity to do non-mechanical work and the capacity to release heat.
i) Erlang ii) Entropy **iii) Enthalpy** iv) BW
Q1. Focus on Capacity of Noisy Channel.
Data rate governs the speed of data transmission. A very important consideration in data communication is how fast we can send data, in bits per second, over a channel. Data rate depends upon 3 factors:

- The bandwidth available
- Number of levels in digital signal
- The quality of the channel – level of noise

Two theoretical formulas were developed to calculate the data rate: one by Nyquist for a noiseless channel, another by Shannon for a noisy channel.

In reality, we cannot have a noiseless channel; the channel is always noisy. Shannon capacity is used, to determine the theoretical highest data rate for a noisy channel:
Capacity = bandwidth * $\log_2(1 + SNR)$
In the above equation, bandwidth is the bandwidth of the channel, SNR is the signal-to-noise ratio, and capacity is the capacity of the channel in bits per second.
Bandwidth is a fixed quantity, so it cannot be changed. Hence, the channel capacity is directly proportional to the power of the signal, as SNR = (Power of signal) / (power of noise).
The signal-to-noise ratio (S/N) is usually expressed in decibels (dB) given by the formula:

$10 * \log_{10}(S/N)$
so for example a signal-to-noise ratio of 1000 is commonly expressed as:

$10 * \log_{10}(1000) = 30$ dB.

Examples:
Input1 : A telephone line normally has a bandwidth of 3000 Hz (300 to 3300 Hz) assigned for data communication. The SNR is usually 3162. What will be the capacity for this channel?
Output1 : $C = 3000 * \log_2(1 + SNR) = 3000 * 11.62 = 34860$ bps
Input2 : The SNR is often given in decibels. Assume that SNR(dB) is 36 and the channel bandwidth is 2 MHz. Calculate the theoretical channel capacity.
Output2 : $SNR(dB) = 10 * \log_{10}(SNR)$
$SNR = 10^{(SNR(dB)/10)}$
$SNR = 10^{3.6} = 3981$
Hence, $C = 2 * 10^6 * \log_2(3982) = 24$ MHz

Q2. Analyze the information Rate of Noisy Channel.

Shannon showed that it is possible to communicate at a positive rate and at the same time maintain a low error probability as desired. However, the rate is limited by a maximum rate called the channel capacity. If one attempts to send data at rates above the channel capacity, it will be impossible to recover it from errors. This is called ***Shannon's noisy channel coding theorem*** and it can be summarized as follows:

● A given communication system has a maximum rate of information – C, known as the channel capacity.

● If the transmission information rate R is less than C, then the data transmission in the presence of noise can be made to happen with arbitrarily small error probabilities by using intelligent coding techniques.

● To get lower error probabilities, the encoder has to work on longer blocks of signal data. This entails longer delays and higher computational requirements.

The theorem indicates that with sufficiently advanced coding techniques, transmission that nears the maximum channel capacity – is possible with arbitrarily small errors. One can intuitively reason that, for a given communication system, as the information rate increases, the number of errors per second will also increase.

Mathematically, the channel capacity C is given by

$$C = \lim_{n\to\infty} \frac{1}{n} \max_{p(x_1,x_2,\ldots,x_n)} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n).$$

For memoryless channels, i.e., if the input−output relationship can be written as

$$p(y_1, y_2, \ldots, y_n | x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p(y_i | x_i),$$

then the channel capacity can be simplified to a single letter representation given by

$$C = \max_{p(x)} I(X; Y),$$

where the mutual information $I(X; Y)$ is defined as

$$I(X; Y) = E\left[\log \frac{p(X, Y)}{p(X)p(Y)}\right],$$

Q3. Examine Capacity and Information of AWGN channel.

As a basic but important channel model, let us consider the capacity of a widely encoun- tered communication channel, namely, the discrete time AWGN channel. The input–output relationship is given by

$y = \sqrt{\rho}x + n$, where x is the channel input, y is the channel output, and n represents the additive noise term.

For this channel model the input distribution that maximizes the mutual information is zero mean complex Gaussian with variance 1/2 per dimension (independent for different uses of the

channel), and the resulting channel capacity is given by C =log(1+ρ).
It is clear that for an M-ary modulation scheme, the maximum achievable rate is limited by $\log_2 M$ bits per channel use. It is also observed that for low signal-to-noise ratios, the constrained information rate results are very close to the channel capacity obtained by using Gaussian inputs. This is explained as follows. The optimal input distribution is the one that induces Gaussian distribution at the channel output. However, if the signal-to-noise ratio is very low, the noise variance is high relative to the signal power, and regardless of the input selected, the channel output becomes approximately Gaussian.
In this section, the focus is on a band-limited real AWGN channel, where the channel input and output are real and continuous in time. The capacity of a continuous AWGN channel that is bandwidth limited to B Hz and average received power constrained to P Watts, is given by

$$C_{awgn}(P,B) = B\ log_2\left(1+\frac{P}{N_0 B}\right) \quad bits/s$$

Here, $N_0/2$ is the power spectral density of the additive white Gaussian noise and P is the average power given by

$$P = E_b R$$

where E_b is the average signal energy per information bit and R is the data transmission rate in bits-per-second. The ratio is the ***signal to noise ratio*** (SNR) per degree of freedom. Hence, the equation can be re-written as

$$C_{awgn}(P,B) = Blog_2(1+SNR) \quad bits/s$$

Here, C is the maximum capacity of the channel in bits/second. It is also called ***Shannon's capacity limit*** for the given channel. It is the fundamental maximum transmission capacity that can be achieved using the basic resources available in the channel, without going into details of coding scheme or modulation. It is the best performance limit that we hope to achieve for that channel. The above expression for the channel capacity makes intuitive sense:

● Bandwidth limits how fast the information symbols can be sent over the given channel.

● The SNR ratio limits how much information we can squeeze in each transmitted symbols. Increasing SNR makes the transmitted symbols more robust against noise. SNR represents the signal quality at the receiver front end and it depends on input signal power and the noise characteristics of the channel.

● To increase the information rate, the signal-to-noise ratio and the allocated bandwidth have to be traded against each other.

● For a channel without noise, the signal to noise ratio becomes infinite and so an infinite information rate is possible at a very small bandwidth.

● We may trade off bandwidth for SNR. However, as the bandwidth B tends to infinity, the channel capacity does not become infinite – since with an increase in bandwidth, the noise power also increases.

The Shannon's equation relies on two important concepts:

● That, in principle, a trade-off between SNR and bandwidth is possible

● That, the information capacity depends on both SNR and bandwidth

Shannon Capacity

• The maximum mutual information of a channel. Its significance comes from Shannon's coding theorem and converse, which show that capacity is the maximum error-free data rate a channel can support.
• Capacity is a channel characteristic - not dependent on transmission or reception techniques or limitation.
• In AWGN, C = B log2(1 + γ) bps, where B is the signal bandwdith and γ = S/N is the received signal-to-noise power ratio

Q4.Examine Capacity and Information Rates of Noisy Channels.

In **flat fading**, the coherence bandwidth of the **channel** is larger than the bandwidth of the signal. Therefore, all frequency components of the signal will experience the same magnitude of **fading**. In frequency-selective **fading**, the coherence bandwidth of the **channel** is smaller than the bandwidth of the signal.

5. Capacity of Flat-Fading Channels:

• Depends on what is known about the channel.
• Three cases: 1) Fading statistics known; 2) Fade value known at receiver; 3) Fade value known at transmitter and receiver.
• When only fading statistics are known, capacity is difficult to compute. Only known results are for Finite State Markov channels, Rayleigh fading channels, and block fading.

6. Fading Known at the Receiver:

• Capacity given by $C = \int_0^\infty B\log_2(1+\gamma)p(\gamma)d\gamma$ bps, where p(γ) is the distribution of the fading SNR γ.
• By Jensen's inequality this capacity always less than that of an AWGN channel.
• "Average" capacity formula, but transmission rate is fixed.

7. Capacity with Fading Known at Transmitter and Receiver

• For fixed transmit power, same capacity as when only receiver knows fading.
• By Jensen's inequality, fading reduces capacity w.r.t. AWGN for fixed transmit power.
• Transmit power as well as rate can be adapted.
• Under variable rate and power
$C = \max_{P(\gamma):\int P(\gamma)p(\gamma)d\gamma=S} \int_0^\infty B\log_2\left(1+\frac{P(\gamma)\gamma}{S}\right)p(\gamma)d\gamma,$, where P(γ) is power adaptation

8. Optimal Power and Rate Adaptation

• Optimal adaptation found via Lagrangian differentiation.
• Optimal power adaptation is a "water-filling" in time: power $P(\gamma) = \gamma_0^{-1} - \gamma^{-1}$ increases with channel quality γ above an optimized cutoff value γ0.
• Rate adaptation relative to $\gamma \geq \gamma_0$ is $B\log_2(\gamma/\gamma_0)$: also increases with γ above cutoff.
• Resulting capacity is $C = \int_{\gamma_0}^\infty B\log_2(\gamma/\gamma_0)p(\gamma)d\gamma.$.
• Capacity with power and rate adaptation not much larger than when just receiver knows channel, but has lower complexity and yields more insight into practical schemes.
• Capacity in flat-fading can exceed the capacity in AWGN, typically at low SNRs.

5. Channel Inversion

• Suboptimal transmission strategy where fading is inverted to maintain constant received SNR. • Simplifies system design and is used in CDMA systems for power control.
• Capacity with channel inversion greatly reduced over that with optimal adaptation (capacity equals zero in Rayleigh fading).
• Truncated inversion: performance greatly improved by inverting above a cutoff $\gamma 0$.

6.Capacity of Frequency Selective Fading Channels

• Capacity for time-invariant frequency-selective fading channels is hard: result is a "water-filling" of power over frequency.
• For time-varying ISI channels, capacity is unknown.
• Approximate by dividing up the bandwidth subbands of width equal to the coherence bandwidth (same premise as multicarrier modulation).
• We assume independent fading in each subband.
• Capacity in each subband obtained from flat-fading analysis. Power is optimized over both frequencies and time.

Main Points

• Capacity of flat-fading channels depends on what is known about the fading at receiver and transmitter.
• Capacity when only the receiver knows the fading is the same as when the transmitter also knows but does not adapt power.
• Capacity-achieving transmission scheme uses variable-rate variable-power transmission with power water-filling in time.
• Power and rate adaptation does not significantly increase capacity, and rate adaptation alone yields no increase. These results may not carry over to practical schemes.
• Channel inversion practical but has poor performance. Performance improved by truncating.
• Capacity of frequency-selective fading channels obtained by breaking up wideband channel into sub bands (similar to multicarrier).

Q5. Estimate the capacity of MIMO Channels.

3. MIMO Systems

• MIMO systems have multiple antennas at the transmitter and receiver.
• The antennas can be used for capacity gain and/or diversity gain.
• MIMO system design and analysis can be complex since it requires vector signal processing.
• The performance and complexity of MIMO systems depends on what is known about the channel at both the transmitter and receiver

4. MIMO Channel Decomposition

• With perfect channel estimates at the transmitter and receiver, the MIMO channel decomposes into RH independent parallel channels, where RH is the rank of the channel matrix (min(Mt , Mr) for Mt transmit and Mr receive antennas under rich scattering).
• With this decomposition there is no need for vector signal processing.
• Decomposition is obtained by transmit precoding and receiver shaping.

5. MIMO Channel Capacity: Static Channels

• Capacity depends on whether the channel is static or fading, and what is known about the channel at the transmitter and receiver.
• For a static channel known at the transmitter and receiver capacity is given by

$$C = \max_{P_i:\sum_i P_i \le P} \sum_i B\log_2\left(1+\frac{\sigma_i^2 P_i}{\sigma_n^2}\right) = \max_{P_i:\sum_i P_i \le P} \sum_i B\log_2\left(1+\frac{P_i\gamma_i}{P}\right).$$

This leads to a water-filling power allocation in space. • Without transmitter knowledge, outage probability is the right metric for capacity.

• In the limit of a large antenna array (Massive MIMO), even without TX CSI, random matrix theory dictates that the singular values of the channel matrix converge to the same constant.

Hence, the capacity of each spatial dimension is the same, and the total system capacity is

C = min(Mt , Mr)B log(1+ρ).

So capacity grows linearly with the size of the antenna arrays in Massive MIMO systems.

6. MIMO Channel Capacity: Fading Channels

• In fading, if the channel is unknown at transmitter, uniform power allocation is optimal, but this leads to an outage probability since the transmitter doesn't know what rate to transmit at:

$$P_{out} = p\left(\mathbf{H} : B\log_2 \det\left[\mathbf{I}_{M_r} + \frac{\rho}{M_t}\mathbf{H}\mathbf{H}^H\right] > C\right).$$

• Capacity with both transmitter and receiver knowledge of the fading is the average of the capacity for the static channel, with power allocated either by an instantaneous or average power constraint. Under the instantaneous constraint power is optimally allocated over the spatial dimension only. Under the average constraint it is allocated over both space and time.

7. Beamforming

• Beamforming sends the same symbol over each transmit antenna with a different scale factor.

• At the receiver, all received signals are coherently combined using a different scale factor.

• This produces a transmit/receiver diversity system, whose SNR can be maximized by optimizing the scale factors (MRC).

• Beamforming leads to a much higher SNR than on the individual channels in the parallel channel decomposition.

• Thus, there is a design tradeoff in MIMO systems between capacity and diversity.

Main Points

• MIMO systems exploit multiple antennas at both TX and RX for capacity and/or diversity gain.

• With both TX and RX CSI, multiple antennas at both transmitter and receiver lead to independent parallel channels.

• With TX and RX CSI, static channel capacity is the sum of capacity on each spatial dimension.

• Without TX CSI, use outage as capacity metric.

• For large arrays, random gains become static, and capacity increases linearly with the number of TX/RX antennas.

• With TX and RX CSI, capacity of MIMO fading channel uses waterfilling in space or space/time - leads to min(Mt , Mr) capacity gain.

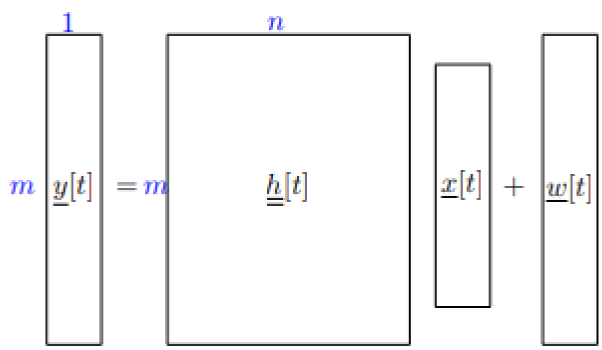
• Beamforming transforms MIMO system into a SISO system with TX and RX diversity.

Beamform along direction of maximum singular value

Q6. Determine the Capacity of Non Coherent MIMO Channels.

In non-coherent wireless communications, the receiver (neither the transmitter) are not aware of the actual realization of the channel coefficients. The encoding and decoding should be performed by using only the statistics of the coefficients.
Flat Fading Flat fading is referred to the model in which the changes in channel are so slow that the channel coefficients are assumed to be constant over each block of communications.
i.e., h[t] = h for t = 1, 2, · · · , T.
In MIMO communications there are n transmitting antennas and m receive antennas each receiving a linear combination of the transmitted signals from all TX antennas. The channel is modeled as following:

$$\begin{aligned} y_j[t] &= \sum_{i=1}^{n} h^*_{i,j}[t]x_i[t] + w[t], \qquad \text{for } t = 1, \cdots, T \\ \underline{y}[t] &= \underline{\underline{h}}[t]\underline{x}[t] + \underline{w}[t] \end{aligned}$$

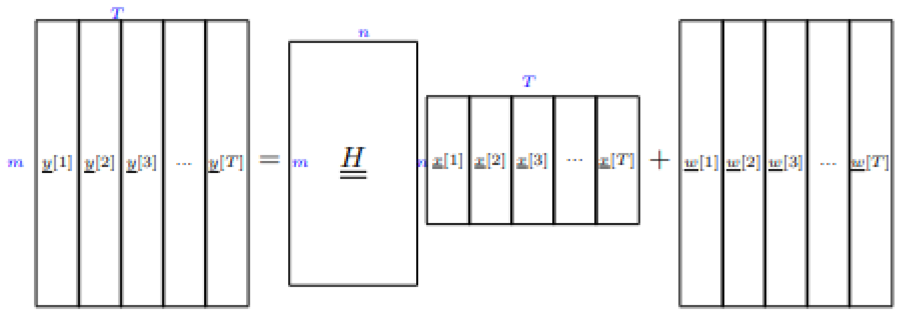


In *flat fading* MIMO communications, the matrix $\underline{\underline{h}}[t]$ is constant over a block of communication i.e., $\underline{\underline{h}}[t] = \underline{\underline{H}}$ for $t = 1, 2, \cdots, T$.

$$\underline{y}[t] = \underline{\underline{H}}\ \underline{x}[t] + \underline{w}[t] \qquad \text{for } t = 1, \cdots, T$$

$$\underline{\underline{Y}}_{m\times T} = \underline{\underline{H}}_{m\times n} . \underline{\underline{X}}_{n\times T} + \underline{\underline{W}}_{m\times T}$$

Degrees of Freedom (DOF=d) The received signal in a block of length T in all antennas, Y ∈ Cm×T lives in a space of complex dimension of mT. But not all these dimensions are affected by the input X. The number of dimensions in received signal that are only affected by the input X (i.e.,



the dimension of subspace that Y lives in, fixing all parameters and changing only X) is called the degrees of freedom of the system. The importance of this parameter is in the fact that in high SNR regime, degrees of freedom of the system is the pre-log factor of capacity of the system
i.e., C(SNR) ≈ d log SNR Intuitively, in high SNR regime, the constraint over the input power is not the limiting factor.
The limiting factor in communications is the number of dimensions of the received signal that we have control over by changing and designing the input signal.

Non-coherent Flat Fading MIMO Communications in high SNR

Channel Capacity

Theorem: If $R \in \mathcal{C}^{M\times T}$ is isotropically distributed (i.e., $\forall Q, p(R) = p(RQ)$), then

$$h(R) = h(C_R) + \log|G(T,M)| + (T-M)\mathbb{E}[\log\det RR']$$

- $h(C_R) + \log|G(T,M)|$ is differential entropy of R in $\mathcal{C}^{M\times M} \times G(T,M)$.
- $(T-M)\mathbb{E}[\log\det RR']$ is the Jacobian term for the coordinate change.

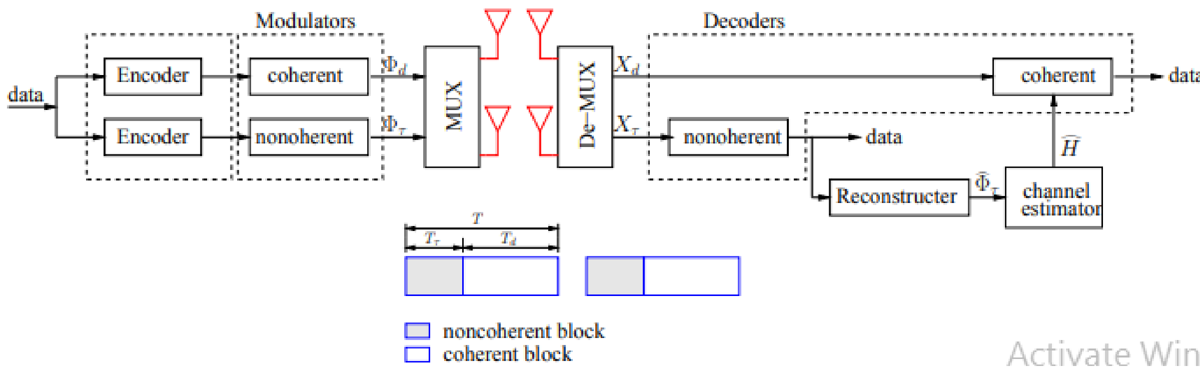
Using the above theorem, mutual information $I(X;Y)$ could be approximated in new corrdinate as following:

$$I(X;Y) = h(Y) - h(Y|X)$$

$$\begin{aligned} h(Y|X) &= m\mathbb{E}[\log\det A^2] + m^2\log(\pi e) + m(T-m)\log(\pi e\sigma^2) \\ h(Y) &\approx h(HX) \\ &= h(C_{HX}) + \log|G(T,M)| + (T-m)\mathbb{E}[\log\det HA^2H'] \\ &= h(C_{HX}) + \log|G(T,M)| + (T-m)\mathbb{E}[\log\det HH' + \log\det A^2] \end{aligned}$$

To maximize the mutual information, it is clear that the optimal power allocation is constant for all antennas and the information is conveyed through Ω_X.

Thus, $\mathbf{X} = \mathbf{A\Theta}$ where $P(\mathbf{A} = \sqrt{T}I_M) = 1$ and $\mathbf{\Theta}$ is a unitary matrix with uniform distribution over random unitary matrices in $\mathcal{C}^{T\times T}$.



Q7. Model the constrained signaling in MIMO Communication.
Since the capacity of a MIMO channel (in the presence of Gaussian noise) is achieved with Gaussian inputs, using such constellations will prevent us from achieving the channel capacity in general. For instance, if we use BPSK signaling, regardless of the signal-to-noise ratio (and the fading statistics), we can transmit at most Nt bits per channel use. However, it is important to find the limits of reliable transmission when constrained signaling is employed as well. This was rather easy to compute for the case of AWGN channels via a single integration, however for the case of fading channels, in particular for MIMO systems, it is not straightforward to do the computations analytically. In order to compute the information rates with constrained signaling, we go back to the original capacity expression, that is,

$$C = \max_{p(\boldsymbol{x})} H(\boldsymbol{Y}) - H(\boldsymbol{N}).$$

We note that the maximization should be done using the specific input constraints. This calculation is in general difficult, however we note that

$H(\boldsymbol{Y}) = E[-\log p(\boldsymbol{Y})],$ where the expectation is over the statistics of the random vector Y. This expectation can be easily calculated using Monte Carlo techniques, i.e., by generating a large number of realizations of the channel output (of course, using the input constraints), computing the log of the joint probability of the realization, and averaging them.
We observe that for low signal-to-noise ratios, the information rates with constrained i.u.d. (independent uniformly distributed, or i.u.d.) inputs are very close to the channel capacity. This is be- cause, when the noise variance is very high, the channel outputs with the constrained inputs are approximately Gaussian distributed, thus the resulting mutual information is almost optimal. However, when the signal-to-noise ratio is increased, the channel capacity grows with the signal-to-noise ratio while the information rates are limited by their maximum values determined by the constellation size and the number of transmit antennas, hence the large gap between them. It is interesting to observe that, there is an almost linear increase of the achievable information rates (or, constrained capacity) with the number of antennas n.
Q8. Explain Ergodic capacity for Fading Channel.
The ergodic capacity is the maximum mutual information between the input and output if the code spans an infinite number of independent realizations of the channel matrixH.
Ergodic capacity is the upper bound of the capacity on the statistics channel (i.e. time-varying channel). It can be evaluated by averaging the capacity is obtained at a particular time instance on a fading channel over an infinite time interval.

For fading channels, there are two capacity definitions that are important: ergodic (Shannon type) capacity and outage capacity for non-ergodic channels.
In this case, the channel coefficients vary over time, and it is possible to average over their statistics by coding over large blocks of data. For this scenario, the channel capacity is obtained by taking the expected value of capacity of an AWGN channel with the signal-to-noise ratio |h|2ρ (over the statistics of the channel coefficient h). That is, defining z =| h|,

$$C_{fading} = \int_0^\infty \log\left(1 + z^2\rho\right) p(z)dz, \tag{3.9}$$

or, equivalently by averaging over the p.d.f. of the instantaneous signal-to-noise ratio ρ' of the channel as

$$C_{fading} = \int_0^\infty \log\left(1 + \rho'\right) p(\rho')d\rho', \tag{3.10}$$

where ρ=| $h|^2\rho$.
Since the additive noise is independent Gaussian, the capacity is achieved by transmitting a very long sequence of independent Gaussian inputs with zero mean, and a fixed variance (determined by the power constraint). Since the transmitter does not have the channel state information, all the symbols are transmitted with the same variance. At the receiver, the resulting channel outputs are classified with respect to their fading states. This can be done since the fading coefficients are known at the receiver. Therefore, equivalently, for each state of the channel, long blocks of independent identically distributed Gaussian inputs are employed, i.e., for a given signal-to-noise ratio, a transmission rate of log(1+ρ) is obtained, and the ergodic capacity is simply their average.
The information rate can be computed by simply averaging the expression over the instantaneous signal-to-noise ratio, $\rho|h|^2$. Clearly, Monte Carlo techniques can also be employed as in the case of AWGN channels.
We observe that the channel fading deteriorates the capacity and information rates considerably compared with the case of AWGN channels, as expected. Also, the information rates with BPSK and QPSK inputs are very close to the channel capacity for low signal-to-noise ratios.
Q9. Demonstrate the Outage Capacity and Outage Information Rates for Fading Channel.
The outage capacity is defined as the maximum data rate that can be achieved given a specified outage probability, i.e. the probability that a signal is not properly decoded. ... Therefore, there is an optimal outage probability value which maximizes the outage capacity.
The outage capacity is of interest for slowly fading channels, where the SNR of the channel is a random variable that can take realizations close to zero.
In Information theory, outage probability of a communication channel is the probability that a given information rate is not supported, because of variable channel capacity. Outage probability is defined as the probability that information rate is less than the required threshold information rate.
In Quasi static Rayleigh channel Model the ergodic or Shannon type capacity of the channel is simply zero. This is because, regardless of the frame length, there is a non-zero probability that the instantaneous mutual information between the channel input and the channel output is below any given fixed rate. Therefore, another capacity definition is needed, namely, the

outage capacity.

Consider a quasi-static fading channel, for a given channel realization, or instanta- neous signal-to-noise ratio, the channel capacity is a random variable.

For instance, for a quasi-static Rayleigh flat fading channel, the instantaneous channel capacity as a function of the signal-to-noise ratio is given by

$C(\rho) = \log(1+\rho)$, where ρ is an exponential random variable.

Therefore, for any given rate R, there is a probability that any coding scheme will not be supported reliably over this channel. That is, the outage probability is given by

$P_{out} = P(C(\rho) < R)$.

To interpret this in another way,for any given outage probability level, there is an outage capacity associated with it, with the interpretation that when the system is not in outage (which occurs with probability $1-P_{out}$), this particular transmission rate can be supported. A similar definition holds for the outage information rates. That is, when a specific channel input is employed, the outage probability is given by

$P_{out} = P(I(\rho) < R)$ for a given rate R where the mutual information is computed for the given signal-to-noise ratio level using specific input constraints.

Q10. Investigate on Deterministic MIMO channel.

For deterministic MIMO channels, the assumption is that the channel gain matrix H is fixed. This may be the case for instance for fixed wireless links when the variations in the environment are negligible.

Using singular value decomposition (SVD), we can write

$H = UVH$,

where U and V are $N_t \times N_t$ and $N_r \times N_r$ unitary matrices (i.e., $UHU = I_{Nt}$, $V HV = I_{Nr}$), and is an $N_t \times N_r$ non-negative diagonal matrix whose diagonal elements are the singular values of the matrix H.

Let $\sigma_1, \sigma_2, \ldots, \sigma_\nu$ denote the non-zero singular values, clearly, $\nu \leq \min\{N_t, N_r\}$. Without loss of generality, we assume that the singular values are sorted in a non-increasing order, i.e., $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_\nu$.

Using the SVD of H, as an equivalent channel model, we can write as

$\tilde{y} = \sqrt{\rho}\tilde{x} + \tilde{n}$, where $\tilde{x} = xU$, $\tilde{y} = yV$ and $\tilde{n} = nV$.

Since the matrices U and V are invertible, knowing $\tilde{x}$ and $\tilde{y}$ is the same as knowing x and y. Also, the power constraint on the new input vector is the same as the one on the original input vector since

$E[\tilde{x}\tilde{x}H] = E[xUUHxH] = E[xxH] \leq 1$.

Furthermore, the new noise vector has the same statistics as the original one (since V is unitary and n has independent Gaussian elements). Therefore, equivalent to the original channel model, we have the following set of parallel channels

$$\tilde{y}_1 = \sqrt{\rho}\sigma_1\tilde{x}_1 + \tilde{n}_1,$$
$$\tilde{y}_2 = \sqrt{\rho}\sigma_2\tilde{x}_2 + \tilde{n}_2,$$
$$\vdots$$
$$\tilde{y}_\nu = \sqrt{\rho}\sigma_\nu\tilde{x}_\nu + \tilde{n}_\nu,$$
$$\tilde{y}_{\nu+1} = \tilde{n}_{\nu+1},$$
$$\vdots$$
$$\tilde{y}_{N_r} = \tilde{n}_{N_r},$$

where the power constraint on the input is Nt i=1 E[|˜ xi|2]≤1 and the noise terms areindependent complex Gaussian with variance 1 /2 per dimension. Clearly, this is nothing but a set of usual parallel Gaussian channels with independent noise terms having identical variances.

Equal Transmit Power Allocation Let us first assume that although the MIMO channel is deterministic (fixed H), the transmitter does not have access to the channel matrix (but the receiver does). Therefore the transmitter cannot optimize its power allocation among its antennas. In this case, the capacity of the MIMO channel is given by

C =max p(x) I(X;Y),

(3.21) =max p(x) H(Y)−H(Y|X).

When the channel input is given, the remaining uncertainty in the output is simply the entropy of the noise. Thus we can write

C =max p(x) H(Y)−H(N).

As the noise components are independent complex Gaussian with variance 1/2 per dimension, the entropy of N is

H(N) = Nr log(πe).

Therefore, the capacity is achieved if the entropy of the channel output is maximized. On the other hand, for the entropy of the channel output, we can write

H(Y) ≤logdet(πeRy), where Ry denotes the covariance matrix of the output vector y, and we have equality if y is complex Gaussian. Clearly the channel output vector y is complex Gaussian if the channel input vector x is complex Gaussian.

Noting that Ry = E[yHy]= ρHHRxH +INr,

we can write the channel capacity as

C =max Rx logdet(ρHHRxH +INr),

where the maximization is subject to the constraint E[xxH]=trace(Rx) ≤1.

Since we assume that the transmitter does not have the channel state information, with the given trace constraint, it selects an input with covariance Rx = 1 Nt INt. Namely the capacity-achieving input vector is independent complex Gaussian with equal power on each of the antennas. Therefore, the channel capacity is given by

C =logdetINr + ρ Nt HHH

Since the product HHH is positive semi definite with positive eigen values (λ1,λ2,..., λν) as the squares of the non-zero singular values of H (i.e., λ1 = σ2 1,λ2 = σ2 2,...,λν = σ2 ν), it can be diagonalized using a unitary matrix W as HHH = WWH, where is a diagonal matrix containing the eigenvalues λ1,λ2,...,λv (and possibly zero eigenvalues). Noting that

$$\log\det\left(I_{N_r} + \frac{\rho}{N_t} H^H H\right) = \log\det\left(I_{N_r} + \frac{\rho}{N_t} W \Lambda W^H\right),$$

$$= \log\det\left(W\left(I_{N_r} + \frac{\rho}{N_t}\Lambda\right) W^H\right),$$

we obtain

$$C = \sum_{i=1}^{v} \log\left(1 + \frac{\rho}{N_t}\lambda_i\right),$$

as the MIMO channel capacity.

It is easy to interpret the final result using the equivalent set of parallel channels as described previously. Since the optimal input with equal power allocation among the transmit antennas is independent complex Gaussian, and the transformation from x to ˜ x is unitary, the optimal choice of ˜ x is also independent complex Gaussian. Therefore, for capacity-achieving signaling, independent Gaussian inputs over the set of Nr parallel channels are used (each with power 1 Nt), and the overall capacity is simply the sum of the capacities of these parallel channels.

Q11.List out the properties of following:

iv) Noisy Channels

v) AWGN Channels

vi) Fading Channels

Noisy Channel:

Noisy coding theorem

The scenario for the noisy coding theorem is the same as that for the noiseless coding theorem, with the addition of a noisy channel between the compressor (renamed as the encoder) and the decompressor (renamed as the decoder). For many years, it was believed that as the transmission rate through a noisy channel is increased, the error rate also increased, no matter how slow the rate was. However, what Shannon showed, in this second theorem of his, is that as long as the transmission rate is below a certain maximum value, known as the capacity of the channel *C*, messages could be transmitted with *zero* error, asymptotically in the message length. This capacity is C=maxp(X)I(X;Y)

$$\frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}.$$

The noisy coding theorem states that the rate *C* is *optimal*, i.e. that rates larger than *C* are not possible without inducing unrecoverable errors as *n* → ∞, and *achievable*, meaning a scheme exists which can transmit at any rate less than or equal to *C*, in the limit as *n* → ∞.

FANO'S INEQUALITY AND THE DATA PROCESSING INEQUALITY

Fano's inequality is used to find a lower bound on the error probability of any decoder; it relates the average information lost in a noisy channel to the probability of the categorization error. We first introduce the concept of a Markov chain, then state Fano's inequality, and finally prove the data processing inequality, a result with many applications in classical information theory.

Random variables, X, Y, and Z, form a Markov chain denoted as X → Y → Z, if the conditional probability distribution of Z depends only on Y (it is independent of X):

The random variables, X1,X2,X3,…,X_n, could represent the states of a system and, as we know, the Markov property means that the past and the future state are conditionally independent given the present state; this implies that all terms in this sum except the first are equal to zero.

Consider now a random variable, X, with the probability density function, $p_X(x)$, and let $|\chi|$ denote the number of elements in the range of X. Let Y be another random variable related to X, with the conditional probability, $p_{Y|X}(y|x)$. Intuitively, we expect that the error when we use Y to estimate X will be small when the conditional entropy, $H(X|Y)$, is small. Our intuition is quantified by Fano's inequality.

AWGN Channels

Additive white Gaussian noise (**AWGN**) is a basic noise model used in Information theory to mimic the effect of many random processes that occur in nature. The modifiers denote specific characteristics:

- ***Additive*** because it is added to any noise that might be intrinsic to the information system.
- ***White*** refers to the idea that it has uniform to power across the frequency band for the information system. It is an analogy to the color white which has uniform emissions at all frequencies in the visible spectrum.
- ***Gaussian*** because it has a normal distribution in the time domain with an average time domain value of zero.

Wideband noise comes from many natural noise, such as the thermal vibrations of atoms in conductors (referred to as thermal noise or Johnson–Nyquist noise), shot noise, black-body radiation from the earth and other warm objects, and from celestial sources such as the Sun. The central limit theorem of probability theory indicates that the summation of many random processes will tend to have distribution called Gaussian or Normal.

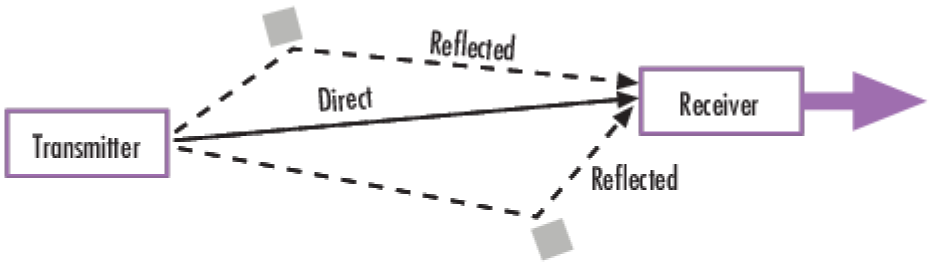
AWGN is often used as a channel model in which the only impairment to communication is a linear addition of wideband or white noise with a constant spectral density (expressed as watts per hertz of bandwidth) and a Gaussian distribution of amplitude. The model does not account for fading, frequency selectivity, interference, nonlinearity or dispersion.

The AWGN channel is a good model for many satellite and deep space communication links. It is not a good model for most terrestrial links because of multipath, terrain blocking, interference, etc.

Fading Channels

Rayleigh and Rician fading channels are useful models of real-world phenomena in wireless communications. These phenomena include multipath scattering effects, time dispersion, and Doppler shifts that arise from relative motion between the transmitter and receiver. This section gives a brief overview of fading channels and describes how to implement them using the toolbox.

This figure depicts direct and major reflected paths between a stationary radio transmitter and a moving receiver. The shaded shapes represent reflectors such as buildings.



The major paths result in the arrival of delayed versions of the signal at the receiver. In addition, the radio signal undergoes scattering on a *local* scale for each major path. Such local scattering typically results from reflections off objects near the mobile. These irresolvable components combine at the receiver and cause a phenomenon known as *multipath fading*. Due to this phenomenon, each major path behaves as a discrete fading path. Typically, the fading process is characterized by a Rayleigh distribution for a non line-of-sight path and a Rician distribution for a line-of-sight path.

The relative motion between the transmitter and receiver causes Doppler shifts. Local scattering typically comes from many angles around the mobile. This scenario causes a range of Doppler shifts, known as the *Doppler spectrum*. The *maximum* Doppler shift corresponds to the local scattering components whose direction exactly opposes the trajectory of the mobile.

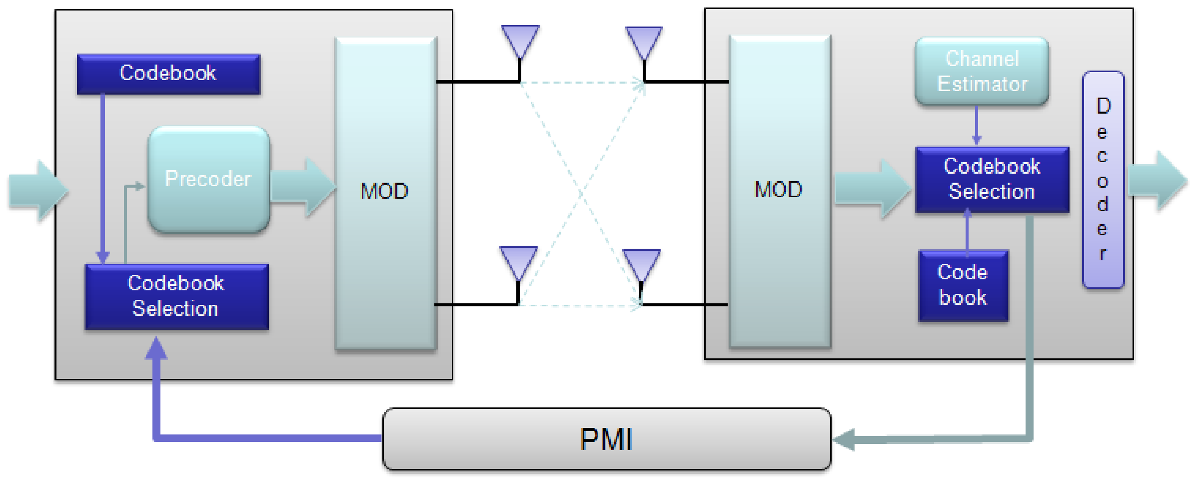
The channel filter applies path gains to the input signal, Signal in. The path gains are configured based on settings chosen in the fading channel object or block.

Q12. Elaborate on Signaling for MIMO Technology.
Rank Indicator is used in signaling Information which is as follows:
RI is an indicator showing how well multiple Antenna work. Each antenna in MIMO configuration works well if the signal from each antenna has NO correlation to each other". "No correlation" implies "no interference to each other".
Maxmum RI value is very closely related to the number of Antenna. Maximum RI is same as number of antenna on each side if the number of Tx antenna and Rx antenna is same. If the number of Tx and Rx are different, the one with less antenna is the same as Max achievable RI.

Max RI means "No Correlation between the antenna", "No interference to each other", "Best Performance".
For example, in case of 2x2 MIMO, the RI value can be 1 or 2. When the value 2 in this case means "No Correlation between the antenna", "No interference to each other", "Best Performance". If the value is 1, it implies that the signal from the two Tx antenna is percieved by UE to be like single signal from single Antenna, which means the worst performance.

PMI-Precoding Matrix Indicator:

For some MIMO implementation (e.g, TM4 in LTE = Closed Loop MIMO), you estimate the channel and select a specific precoding matrix and send it back to the receiver as shown below.



The algorithm by which UE select the codebook which is best fit for the channel at specific moment is as follows.

Codebook Element (Precoding Matrix)

Channel Information Matrix

Step 1 : Calculate $\Omega = W(H^H H)W^H$ for every elements in the code book

Step 2 : Select the codebook element which gives you the minimum value

Step 3 : Report the index of the selected codebook element (PMI) to the receiver

Once a specific codebook element (precoding matrix) is selected as shown here, that precoding matrix is used to transform the incoming bits.

Reciever should know this

Transmitter should know

$$\mathbf{y} = \mathbf{U}^H (\underbrace{\mathbf{U} \Sigma \mathbf{V}^H}_{\text{This is from channel (Information from Channel)}}) \mathbf{V}\mathbf{x}$$

- The three matrix at the center can be a known at least to the reciever since the reciever can estimate channel Matrix H from the received signa and calculate these matrix from H.
- The vector x is known to transmitter since it is just the data that's transmitted.
- The matrix U can be a known to the reciever since it can be derived from H.
- The issue is how to figure out the matrix V.

In the workaround, we use only a few/several pre-defined matrices for V and the receiver evaluate each of the candidates and pick the best one and inform the transmitter of the index of the candidate matrix in stead of the whole matrix contents. For example, in LTE TM4 they defined only 4 candidate V matrix (precoding matrix) for 2 x 2 MIMO and 16 candidates for 4 x 4 MIMO. The selected candidate may not be the best fit in terms of mathematics, but it can be a best-effort in terms of trade-off between mathematical accuracy and report overhead.

Condition Number:

In case, Rank Indicator is same but real communication performance is different. So in order to properly estimate the real performance of MIMO channel we may need another indicator, that indicator is Condition Number as defined as below

$$\kappa(H) = \frac{\lambda_{\max}}{\lambda_{\min}} \quad \begin{cases} \text{Optimal Scenario for Spacial Multiplexing is when } \kappa(H) \approx 1 \\ \text{Non-Optimal Scenario for Spacial Multiplexing is when } \kappa(H) >> 1 \end{cases}$$

Rank Indicator = Number of Nonzero value in Singular Matrix

$$\Sigma = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

It means this channel is the worst condition for MIMO.

$$\kappa(H) = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{\sqrt{2}}{0} \approx \infty$$

Then let's take the condition number. It is '1' as shown below. It means it is perfect condition for spatial multiplexing and you should have maximum performance from this channel.

$$\kappa(H) = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{1}{1} = 1$$

Q13. List out the differences between Noisy and Noiseless Channels.

Noiseless Channels:

- Noiseless Channels are Ideal Channel in which no frames are lost, duplicated or corrupted.
- Protocols used in Noiseless Channels for flow control are
  (ii) Simplest Protocol and (ii) Stop and Wait Protocol
- In Noiseless Scheme, Channel is error free and doesn't require Error Control Mechanism.
- Noiseless Channels are Nonexistent.

Noisy Channels:

- Noisy Channels are having Noise of various ranges.
- Frames are lost, duplicated and corrupted in Noisy channels
- Protocols used in Noisy Channels for flow control are

(i) Stop and Wait Protocol Automatic Repeat Request
(ii) Go-Back-N Automatic Repeat Request
(iii) Selectively Repeat Automatic Repeat Request

- In Noisy Scheme, Channel is error prone and requires Error Control Mechanism.
- Noisy Channels are all Real Time Signals.

Q14. List out the differences between AWGN and (Rayleigh,Rician)Fading Channels.
AWGN noise: AWGN is a noise that affects the transmitted signal when it passes through the channel. It contains a uniform continuous frequency spectrum over a particular frequency band.
Rayleigh Fading: When no LOS path exists in between transmitter and receiver, but only have indirect path than the resultant signal received at the receiver will be the sum of all the reflected and scattered waves.
Rician Fading: It occurs when there is a LOS as well as the non-LOS path in between the transmitter and receiver, i.e. the received signal comprises on both the direct and scattered multipath waves.
(i)For Mobile Communication experiment we get the following Results:
For execution of digital communication using BFSK and MQAM and implementation of wireless communication used DS-CDMA and SFH-CDMA for AWGN, Rician Fading and Rayleigh fading the results are as follows:
The AWGN channel will be better realization than Rayleigh fading channel because the BER values lower than the BER value of fading use Rayleigh channel while using Rician fading better realization than AWGN and Rayleigh channels because BER value of Rician channel less than Rayleigh fading and upmost than AWGN channel.
(ii)For OFDMA Wimax the following results we get:
1. The performance of AWGN channel is the best of all channels as it has the lowest bit error rate (BER) under QAM, 16-QAM & 64-QAM modulation schemes. The amount of noise occurs in the BER of this channel is quite slighter than fading channels.
2. The performance of Rayleigh fading channel is the worst of all channels as BER of this channel has been much affected by noise under QAM, 16-QAM & 64-QAM modulation schemes.
3. The performance of Rician fading channel is worse than that of AWGN channel and better than that of Rayleigh fading channel. Because Rician fading channel has higher BER than AWGN channel and lower than Rayleigh fading channel. BER of this channel has not been much

affected by noise under QAM, 16-QAM & 64-QAM modulation schemes.