Fundamentals of MIMO Wireless Communication

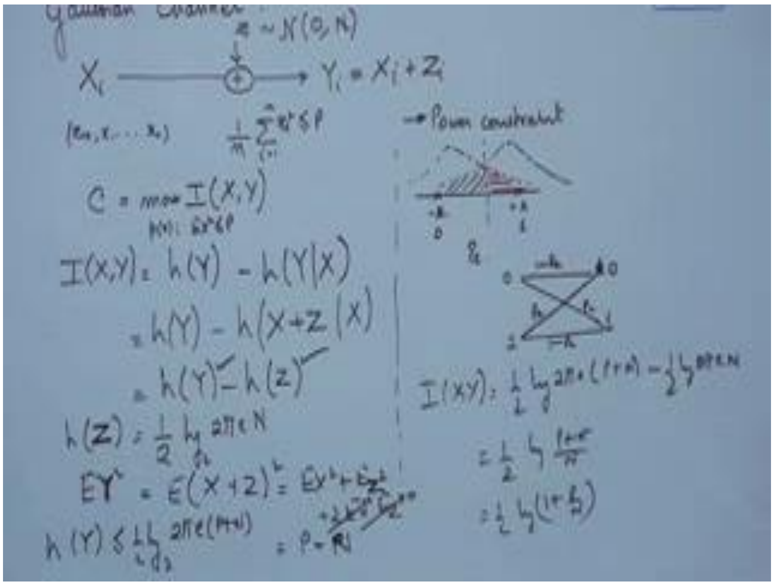
Prof. Suvra Sekhar Das
Department of Electronics and Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 35
Capacity of Deterministic MIMO Channels

Welcome to the course of fundamentals of MIMO wireless communications we have seen the expression of mutual information and we have also tried to understand the intuitive meaning of why that particular expression yields the expression of capacity. So, what we have last discussed is capacity can be expressed in terms of maximization of mutual information between source and the destination over all possible distributions at the source.

So, with that we move forward towards waveform channels and we would be interested in looking at the Gaussian channel. So, as to get the expression which is most commonly known and using that or equipped with all these things we will be able to easily understand the expression of typically flat a frequency flat MIMO channel.

(Refer Slide Time: 01:06).



So, let us begin our discussion with that. So, if we take the Gaussian channel. So, for the Gaussian channel we have the source which generates symbols x I and this symbols go through the channel where z I noise gets added and what is received is y I and this noise z I we will say that distributed according to normal distribution with 0 mean and certain noise power. So, which you can write it as n or n naught or n noise and we also have the constraint that if a codeword is designed in a way that the sequence of symbols x n that is transmitted we have the constraint that one upon n I equals to 1 to n x I squared is less than or equal to p. So, this is basically a power constraint this is a very, very important because most of the communication systems that we deal with they have a some constraints and these constraints amongst many constraint one of the most commonly understood constraint is a power constraint; that means, that there is limited amount of power available at the transmitter there can be many kinds of power constraints1of the most power constraint is a average power constraint.

So, what we are considering over here is average power constraint between the particular form. So, moving forward, this is the expression this is the one which constraints the power and if we have to find the capacity of the information theoretic capacity of this channel we would write c as max I x y. So, this is we have to explain over p of x such that e x square is less than or equal to p, given this constraint over all possible distributions due to maximize the capacity.

Now, if we consider b p s p channel we will just digress a little bit. Now if you consider b p s p modulation. So, this is let say minus s1and this is plus s1or plus a n minus a or if we consider plus minus a and plus a when signal is transmitted the when a is transmitted the probability of the received signal in PDF can be drawn in this way, because of noise when minus a is transmitted the PDF or the received amplitude can be drawn in this manner and if we consider a has been transmitted and find the area under this curve we are going to get the probability of error when, plus a is transmitted similarly when this particular when minus a is transmitted we could find the area under this tale and we could find what is the error probability.

So, basically what we can calculate is probability of error under this condition. So, if we would be able to calculate a probability of error then, we can say that suppose, if this is 0 and this is 1 because 0 is sent with the probability of error it becomes a one or rather with the probability of error p e it becomes 1 and with 1 minus pe it remains 0. If a one is sent it goes to one and 0 it goes to 0 with probability p n with one minus p becomes 1. So, basically waveform channel can be converted to a binary symmetric channel or vice-versa.

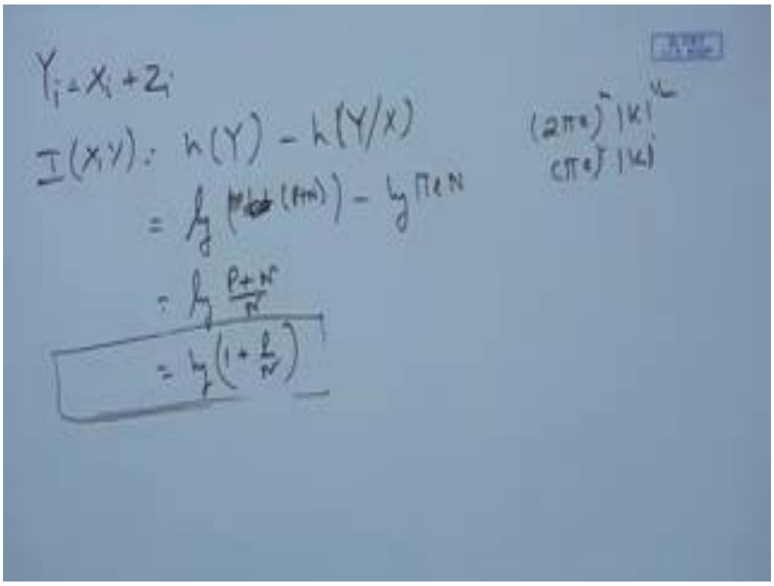
So, we this is suggest keeping in our mind. So, that when we are translating from discrete to ware form channels we can easily do in in this particular fashion, coming back to our expression over here if we would write the expression of I x y what we get is H of y minus of H y given x. So, this expression is what we have derived earlier as the expression of capacity. So, H of y minus H of y given x basically y I as we over here is x I plus z i. So, basically minus H of x plus z condition on x H of x condition on x is 0. So, basically it is H of z conditioned on x. So, z that is noise is independent of the source. So, basically each of z conditioned on x is basically H of z.

So, all we said is x conditioned on x is 0 H of z conditioned on of on x is basically H of z because these 2 are independent and H of z as we all know in case of real constraint we are going to have this as log base 2, 2 pi e n, n is the power of the noise component. So, this is what we have derived the entropy of Gaussian of a Gaussian distribution the differential interfere of Gaussian distribution all we have to do is now calculate this. So, if we look at e of y squared; that means, the variance or the power of y squared we could write it as e of x plus z squared which is e of x squared plus e of z squared plus2e x times e z.

Now, e of x is 0 it is a 0 mean symbol e of z is 0 because it is add a divide Gaussian noise with 0 mean. So, we are left with e x squared plus e z squared. So, e x squared we could restricted to p and e z squared is n. So, basically e of y squared is p plus n or this is the variance of Y you can say. So, if we have to calculate H of y we could say that it is less than or equal to half log 2, 2 pi e p plus n. So, we have this term now we have this term now. So, we have to calculate this. So, I of x y would be half log 2 pi e p plus n minus half log 2 pi e of n.

So, if you work it out it becomes half log p plus n upon n or half log one plus p by n this is the well known expression for the capacity of a Gaussian channel, if we move further beyond this and look at the situation when we have complex Gaussian.

(Refer Slide Time: 07:43)

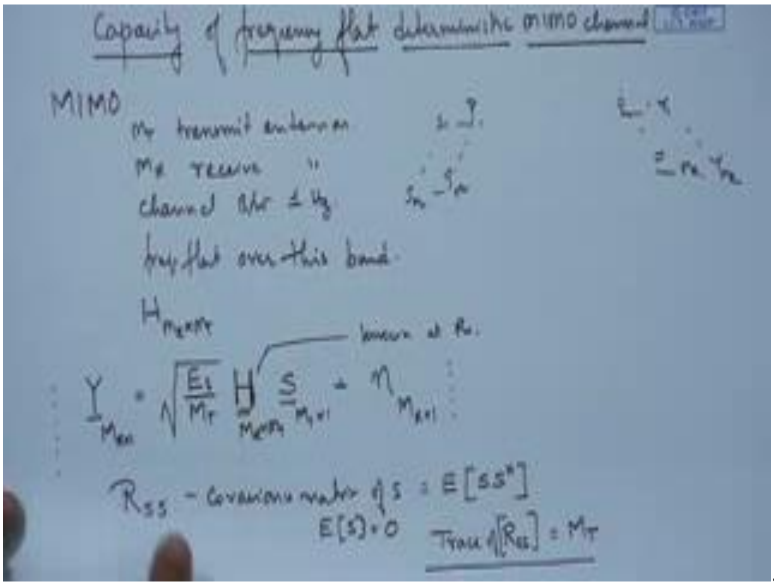


So, again in that case we have y I equals to x I plus z I and in this case we would say that z I is a Gaussian complex, Gaussian distributed in that case again these expressions do not change these expression remain the same.

So, here simply we would be writing log determinant of pi e pi e determinant of p plus n minus log pi e of n rather determinant is not there pi e p plus n because you would remember in the other case we had root 2 pi e to the power of n and in the other case it was pi e to the power of n and in real case it was k determinant of k to the power of half and here determinant of k. So, by virtue of that you get log p plus n upon n which is equal to log of one plus p upon n. So, when we have complex the expression would look like this which is a well known expression for capacity which we are using for Gaussian channel so; that means, whatever we have discussed does easily lead us to the expression of capacity for a Gaussian channel that is what we have just now seen.

Now, since we have seen what is the result for a SISO channel and now we will be moving on to finding the expression of capacity for a MIMO channel.

(Refer Slide Time: 09:17)



.

So, as we move forward we have now almost reached our goal; that means we want to find the capacity of frequency flat we are saying deterministic, I will explain the reason why. So, we are finding the capacity we understand the expression of we understand the meaning of capacity that is maximization of mutual information over all possible source distributions.

We have given an intuitive meaning of it when we talk about frequency flat we understand what is frequency flat we have described in details what leads to frequency flat when, we say deterministic; that means, it is given for a particular realization of a channel we know what are the details of MIMO channel. So, only deterministic what is knew in this case and all we are trying to say is it is known later on we will see when a channel is random channel what happens because of that. So, in this MIMO configuration we have m t transmit antennas we have m r received antennas and let our channel bandwidth the 1 hertz for simplicity or for e is equal always multiplied by the bandwidth to get it and it is assumed it is frequency flat over this band further, we have always written in capital H indicating the channel transfer matrix of size m r plus m t. So, we could write received signal y which is a vector m r plus1being equal to root over e x upon m t. When I write root over e x by m t this clearly means that the signal power is divided equally amongst the m t transmit antennas at this stage.

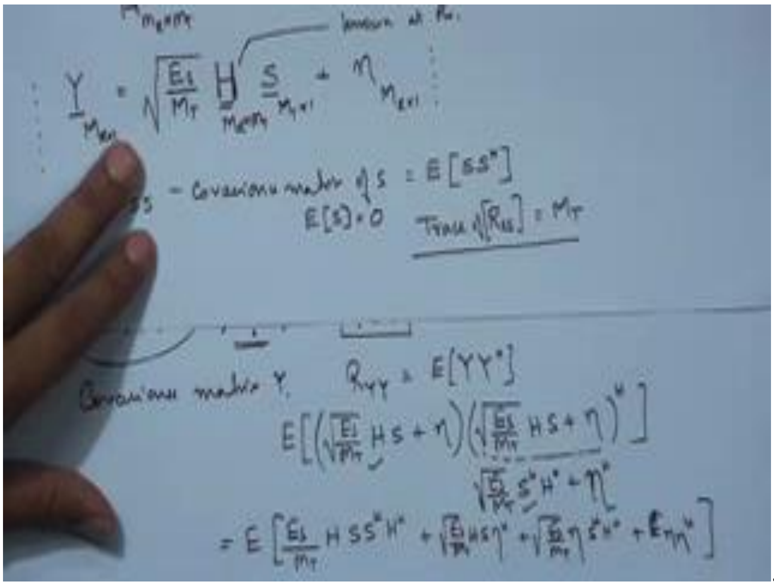
Later on this will get modified times H is a matrix indicated by double underline s is the vector m t plus1this is m r plus m t plus this is noise m r plus 1. So, this equation fit is in m r. So, basically m r number of noise samples corresponding to m r received antennas m t number of symbols and this is m r plus m t channel matrix and this is m r number of using the antennas.

So, basically redundant to draw this picture one to m t and again one to m r number of received antennas s1 up to s m t are getting transmitted y 1 up to t m r is getting received.

So, this is the system model and e s is the total average energy at the transmitted over a symbol period we also need to defined R ss which is the co variance matrix of the source symbol s which is defined as expectation of s times s summation and because we have e of s is equal to 0 using this, we put the constraint the trace of R ss is equal to m t. So, this is important we are putting this this particular constraint because we want total average constraint because, if you look at it we take this particular expression when we have E of ss hermitian over here we basically get R ss and if the trace of it is equal to m t the total transmit power over here would be e s because m t and m t would cancel out and that would keep the total transmit power.

So, this is the signal module which we proceed now we will assume that H is known at the receiver what rather what does it mean it means that ha the channel is estimated at the receiver and it is available and again we will make the assumptions there is perfect knowledge about the channel this perfect channel estimation at the receiver.

(Refer Slide Time: 13:59)

.

The MIMO channel capacity which we are talking about was given by a paper authored by Foschini in the year 96 and there was also wok by a Telake in 99 and there they discussed about this expression which is max of I between s and y. So, between s and y is a same expression over all possible distribution of the source symbol.

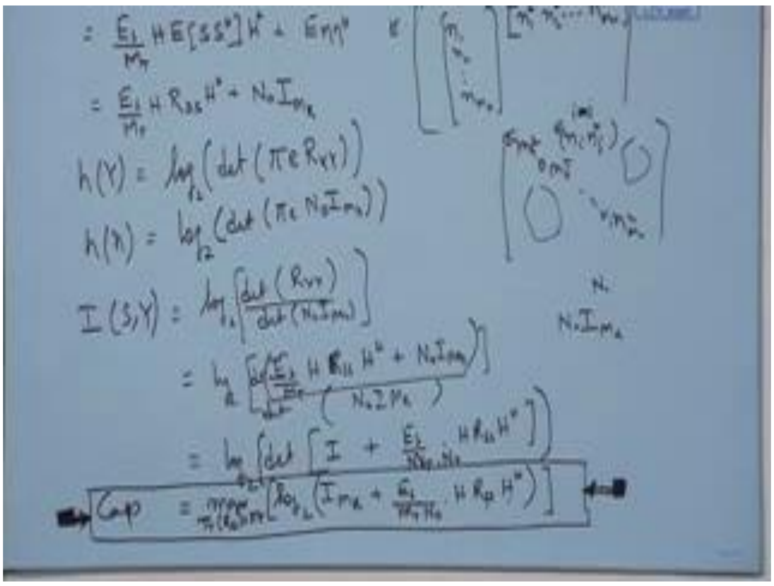
So, the overall concept remains a same, but the things what derived in in the papers by these people. So, I x y is of course, the mutual information and we know all the definitions now. So, it should not be difficult for us at this stage to work this out given whatever we have done in the previous lectures. So, this is the particular way which we have been explaining I x y H of y minus h. So, according to our definition we should use a small H and not capital H indicating differential entropy; that means, for continuous random variable. So, H of y given s we have just derived this thing for a w g n, where we at seen that y given s is basically this whole thing it given s.

Now, H is deterministic H is known. So, this is the random component at this time of point. So, H of s given s is 0 and it is H of the noise given s. So, basically what it leads us to is this particular 1 we could write as H of noise given s. Now again since noise and symbols are independent since these 2 are independent we could write this as H of n and this1would be H of y. So, we have this expression which we derived for the SISO case the expression at this stage looks similar.

So, in this case if you see H of n is defined nobody can do anything with it because it is about the noise characteristics. All you can do is to work with H of y and if you see the expression of capacity it has maximize this overall possible source distributions. So, when we say maximize I overall possible source distributions what you basically do is try to maximize I over I of s y and maximize this expression. Where you cannot do anything of H and n, what you can do is maximize this expression. So, to maximize this expression we need to look at the details of it. So, we begin with the covariance matrix of y right. So, covariance matrix of y is R YY which is equal to E of YY hermitian, E of YY hermitian.

So, E of YY hermitian to get E of YY hermitian we have to look at these so; that means, expectation of root over e s upon m t times H s plus noise times root over e s upon m t and H s plus noise hermitian. So, when I have a hermitian it could write it has root over e s by m t. I am just writing this part this part s hermitian H hermitian plus the noise hermitian. So, basically what we have is e of if I take the product e s upon m t because it is under square root sign and you have H s s hermitian H hermitian. So, that is a first product between this component and this component the we have a next component plus root over e s upon m t H s noise hermitian; that means, this and this component is done we have next this component root over e s upon m t times noise s hermitian H hermitian.

(Refer Slide Time: 18:36)



Now, we have the last component plus e of n n hermitian sorry e of n n hermitian. So, when we use the e operator inside; that means, we being the e operator inside what you have is you bring e inside. So, e s by m t would remain as it is since H is deterministic e operates only on s, H s s hermitian. Now you can see why we need the covariance matrix plus when I bring the e operation inside e of s is 0 e of n is 0. So, this term goes to 0 again e of n s are both 0. So, these are 0 and they are independent. So, what we are left with this this e of n n hermitian right.

So, this is the variance. So, we could write it in more convenient form H instead of writing E s s hermitian we would write R ss H hermitian plus this is the noise component. So, n one n 2 dot, dot up to n m r times n one conjugate n 2 conjugate up to n m r conjugate expectation of operator of this, if you take this product we will be getting n one squared n 2 squared up to n m r squared at the diagonal rather mod squared and the other terms of the cross terms n I n j conjugate I not equals to j and you are going to get expectation of these you can get expectation of these terms. So, all these terms would be 0 because noise in independent branches is are independent and each of these would give the same value. So, it is basically n naught or n. So, what you have is basically n naught times identity matrix because all these would be 0 of size m r.

So, we could as good as write this as n naught times I m r right. So, we have our expression of R YY and we now know how to write the expression of H y and we also have the expression of H n the expression of H of y in that case we could write it as log 2 determinant of pi E R YY. So, we have changed little bit over here we have said pi e and brought it inside the determinant this is as good as bringing pi e outside and raising it to be power of n because determinant of e times a matrix is basically that a raise to the power of n times the dimension or the size of the matrix.

So, this is no different equation this is all the same and H of n is log 2 determinant of pi e n naught I m r and rather this is a little bit more convenient form of writing this is bit is per second per hertz you could write it as way. So, I of s y we could write it as H y minus H n. So, basically log this minus log of this. So, it is basically log base 2 determinants of if you look at the size of this is m r plus m r this is also m r cross m r. So, basically pi e to the power of m r pi e to the power of m r they would cancel out we would have determinant of R YY upon determinant of n naught I m r right.

So, R YY let us expand it. So, I m r 1, 1 would be log base 2 that is I what we have here yeah we have over here e s by m t H R ss H hermitian plus this is n naught I m r right upon n naught I m r. So, what you have is basically determinant of this upon determinant of n naught I m r. So, this would lead to log base 2 determinant of I because of this1plus e s by m m t e s by m t times n naught times H R ss H hermitian. So, this would be expression of I x y. So, we could write the expression of capacity and this point would write the expression of capacity at this point as capacity is equal to maximize log determinant of I m r plus e s upon m t n naught times h, R ss H hermitian over all possible distributions of the source; that means, subject to the constraint trace of R ss is equal to m t.

So, basically you are maximizing over R ss this particular expression. So, this would be the expression of capacity of the MIMO channel and to derive this we have till now been able to explain the little characteristics of h. So, typical MIMO channel R ss we have defined and we have discussed the typical model which is for frequency flat and we have also used the definition of m dot the differential entropy of the received signal and we have derived that which particular entropy maximizes it.

So, basically we have done step by step finally, to arrive at this expression and what it appears, if you would look at the few steps that we have taken is pretty straight forward a given that we now understand all the exact terms that were defined in all previous lectures till now. So, a we would like to conclude this particular lecture at this point stating that we have arrived at the expression of capacity I would like you to take a look at this expression of the MIMO channel capacity which is a very, very important and based on this we will proceed on to carry out certain more expressions which would be able to give us a hint about what is the meaning of this particular expression.

So, since we arrived at the most important part of this course, which is talking about the capacity of the MIMO channel and we have seen the expression it now remains for us to explore this expression this particular expression open it up and see what happens, when this when the channels information that is H which we said is known. We say that let it be known only at the receiver then we will say that now let us we change the assumption and we say that there is a feedback channel because of the feedback channel the channel is known at the transmitters.

So, what extra can be done what is the impact on capacity we have also studied the correlation matrix or the co correlation effect on the channel or how can correlation be module. So, using whatever we have studied on the correlation model of MIMO channels, we will then bring out the impact of correlation on the capacity of MIMO channel and finally, we will look at some architectures which will be giving us the capability to trade off between the diversity gain that can be a achieve we have analyze diversity gain, we will see a special multiplex gain or the MIMO capacity gain and will see how can we tradeoff between achieving high capacity and bettor error probability or a better reliability link a towards the end of this particular course.

Thank you.