

Information Theory, Coding and Cryptography
Dr. Ranjan Bose
Department of Electrical Engineering
Indian Institute of Technology, Delhi

Module - 32
Trellis Coded Modulation
Lecture - 32

Hello and welcome to our next lecture on Trellis Coded Modulation.

(Refer Slide Time: 00:33)

Information Theory, Coding and Cryptography

Outline

- Computation of d_{free}
- TCM for fading channels
- Space Time Trellis Codes
- Slow Rayleigh Fading Scenario
- Fast Rayleigh Fading Scenario

Indian Institute of Technology, Delhi 2 Ranjan Bose
Department of Electrical Engineering

Let us start with a brief outline of today's talk. As we have seen that the d_{free} which is the free distance is one of the single most important parameters to characterize our trellis coded modulation scheme and we would look at some methods to compute d_{free} . Then we go to fading channels, commonly encountered in wireless communications, and we will see that the TCM design criteria for fading channels is different than a normal additive white Gaussian noise channels. Then we will logically shift to space time trellis codes, which is a logical extension going from TCM and we look at two distinct scenarios; slow Rayleigh Fading Scenario and fast Rayleigh Fading Scenario. We will look at the distance, product design criteria for these two cases.

(Refer Slide Time: 01:30)

Information Theory, Coding and Cryptography

Recap

- Combined Coding and Modulation
- Trellis Coded Modulation
- Free distance
- Ungerboeck's design rules
- Performance Evaluation

 Indian Institute of Technology, Delhi 3 Ranjan Bose
Department of Electrical Engineering

Let us do a quick recap; what we have learnt so far is the combined coding and modulation. Then we introduced the notion of trellis coded modulation, followed by what we mean by free distance; then we looked at Ungerboeck's heuristic design rules and spend some time looking at the performance evaluation of designed TCM schemes.

(Refer Slide Time: 01:52)

Information Theory, Coding and Cryptography

TCM Decoding

- We have seen that, like convolutional codes, TCM schemes are also described using **trellis diagrams**.
- Any input sequence to a TCM encoder gets encoded into a sequence of symbols based on the trellis diagram.
- **The encoded sequence corresponds to a particular path in this trellis diagram.**
- There exists a **one-to-one correspondence** between an encoded sequence and a path within the trellis.
- The task of the TCM decoder is simply to identify the **most likely path in the trellis**.
- This is based on the **maximum likelihood criterion**.
- As seen in the previous chapter, an efficient search method is to use the **Viterbi algorithm**.

 Indian Institute of Technology, Delhi 4 Ranjan Bose
Department of Electrical Engineering

We can recollect the TCM decoding is done using the trellis diagram, and we use the notion that any transmitted sequence is actually a path in the trellis. And therefore, we have to find out the corresponding path in the trellis in terms of the decoding algorithm.

And we do this using the maximum likelihood decoding criteria and to do so, we use the very well-known Viterbi algorithm.

We are used earlier the Viterbi algorithm for convolutional codes, but remember TCM is non-linear as opposed to convolutional codes which were linear codes.

(Refer Slide Time: 02:38)

Information Theory, Coding and Cryptography

TCM Decoding

- For **soft decision decoding** of the received sequences using the Viterbi algorithm, each trellis branch is labeled by the branch metric based on the observed received sequence.
- Using the maximum likelihood decoder for the additive white gaussian Noise (AWGN) channels, the branch metric is defined as the **Euclidean distance** between the coded sequence and the received sequence.
- The Viterbi decoder finds **a path through the trellis** which is closest to the received sequence in the Euclidean distance sense.

 Indian Institute of Technology, Delhi

5

Ranjan Bose
Department of Electrical Engineering

So for TCM Decoding, we are realized that soft decision decoding was used and we employed the Euclidean distance between the coded sequence and the received sequence, as a measure to find out the maximum likelihood sequence. And the Viterbi decoder tries this measure to find the most likely path in the trellis. So, we are using Euclidean distance as opposed to hamming distance, which was used for the convolutional encoder decoder case.

(Refer Slide Time: 03:10)

Information Theory, Coding and Cryptography

Performance Measures

- There are different performance measures for a TCM scheme designed for an AWGN channel.
- We have already discussed the asymptotic coding gain, which is based on **free Euclidean distance, d_{free}** .
- We will now look at some other parameters that are used to characterize a TCM code.
- **Definition** The **average number of nearest neighbours** at free distance, $N(d_{free})$, gives the average number of paths in the trellis with free Euclidean distance d_{free} from a transmitted sequence.
- This number is used in conjunction with for the evaluation of the error event probability.

 Indian Institute of Technology, Delhi6Ranjan Bose
Department of Electrical Engineering

We looked at some performance measures and we talked about d_{free} as a free Euclidean distance. And we also talked about the N_{free} which is also a function of d_{free} in terms of the average number of paths in the trellis with free Euclidean distance d_{free} . This number is used in conjunction with the evaluation of the error event probability.

(Refer Slide Time: 03:40)

Information Theory, Coding and Cryptography

Performance Evalⁿ over AWGN Ch.

$$P_e \leq T(D) \Big|_{D=e^{-1/4N_0}}$$

where

$$T(D) = \frac{1}{N} \mathbf{1}^T \mathbf{G} \mathbf{1}$$

- And the matrix

$$\mathbf{G} = \sum_{l=1}^{\infty} \sum_{E_l=0} \prod_{n=1}^l \mathbf{G}(e_n)$$

is the matrix transfer function of the error state diagram.

- $T(D)$ is called the **Scalar Transfer Function** or simply the **Transfer Function** of the error state diagram.

 Indian Institute of Technology, Delhi7Ranjan Bose
Department of Electrical Engineering

Then we went on to define this, probability of error in terms of the scalar transfer function $T(D)$. We have learnt how to calculate $T(D)$, using a modified state diagram and

we figured out that finally we can upper bound the error event probability P by this following expression.

(Refer Slide Time: 03:55)

Information Theory, Coding and Cryptography

Performance Evalⁿ

- A tighter upper bound on the error event probability is given by (exercise)

$$P_e \leq \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{d_{free}^2}{4N_0}} \right) e^{\frac{d_{free}^2}{4N_0}} T(D) \Big|_{D=e^{-1/2N_0}}$$

- As asymptotic estimate on the error event probability can be obtained by considering only the error events with free Euclidean distance

$$P_e \approx \frac{1}{2} N(d_{free}) \operatorname{erfc} \left(\sqrt{\frac{d_{free}^2}{4N_0}} \right)$$

 Indian Institute of Technology, Delhi
 8
Ranjan Bose
Department of Electrical Engineering

And if you look at the asymptotic estimate on the error event probability, we get the following expression; probability of error approximately half times N as a d function of d free error function complement under root d free squared over 4 n naught.

(Refer Slide Time: 04:33)

Information Theory, Coding and Cryptography

Bit Error Probability

- The bit error probability can be upper bounded simply by weighting the pairwise error probabilities by the number of incorrect input bits associated with each error vector and then dividing the result by m. Therefore,

$$P_b \leq \frac{1}{m} \frac{\partial T(D, I)}{\partial I} \Big|_{I=1, D=e^{-1/2N_0}}$$

- Where $T(D, I)$ is the **Augmented Generating Function** of the modified state diagram.
- The concept of the modified state diagram was introduced in the chapter on Convolutional Codes.
- A tighter upper bound can also be obtained for the bit error probability, and is given by

$$P_b \leq \frac{1}{2m} \operatorname{erfc} \left(\sqrt{\frac{d_{free}^2}{4N_0}} \right) e^{\frac{d_{free}^2}{4N_0}} \frac{\partial T(D, I)}{\partial I} \Big|_{I=1, D=e^{-1/2N_0}}$$

 Indian Institute of Technology, Delhi
 9
Ranjan Bose
Department of Electrical Engineering

So, from that we went on to discuss, the bit error probability and we showed that, it can be written as a derivative for this modified T as a function of D and I. So, here T D

comma I is the augmented generating function for the modified state diagram. Again when we had talked about convolutional code, we had introduced this function called augmented generating function.

Now what is important is to find out that the same function can be used to get an upper bound on the bit error probability. Here the derivative is evaluated I , I is equal to 1 and D is equal to e raise to power minus 1 over 4 N naught. A tighter upper bound can also be obtained for the bit error probability and is given by the following expression.

So, you see that d_{free} is figuring prominently in all of these expressions, again establishing the fact that it is the single most important parameter. Of course, for the probability of error expression, the noise power also starts figuring in the expressions.

(Refer Slide Time: 05:50)

Information Theory, Coding and Cryptography

Computing d_{free}

- We have seen that the **Euclidean free distance**, d_{free} , is the single most important parameter for determining how good a TCM scheme is for AWGN channels.
- It defines the asymptotic coding gain of the scheme.
- Earlier we saw that the generating function can be used to calculate the **Hamming free distance** d_{free} .
- The transfer function of the error state diagram, $T(D)$, includes information about the distance of all the paths in the trellis from the all zero path.
- If $T(D)$ is obtained in a closed form, the value of d_{free} follows immediately from the expansion of the function in a power series.

Indian Institute of Technology, Delhi 10 Ranjan Bose, Department of Electrical Engineering

Now, since d_{free} is such an important parameter, we spent a few slides talking about computing the d_{free} ok. Now please note that this statement is correct uh, when we are talking about additive white Gaussian noise channels. When we go to fading channels, we will have a slightly different parameter that we will shortly talk about.

So, the generating function can be used to calculate the hamming d_{free} distance earlier. Now the transfer function of the error state diagram $T(D)$ also includes information about the distance of all paths in the trellis from the all zero paths. So, the question is, can we use that information to calculate d_{free} for our case? The question is d_{free} may or may

not be available in a closed form expression all the time. In the event that TD is obtained in a closed form, the value of d_{free} follows immediately from the expansion of the function in a power series. We have seen that earlier.

(Refer Slide Time: 07:06)

Information Theory, Coding and Cryptography

Computing d_{free}

- The generating function can be written as

$$T(D) = N(d_{free})D^{d_{free}^2} + N(d_{next})D^{d_{next}^2} + \dots$$

- where d_{next}^2 is the second smallest squared Euclidean distance.
- Hence the smallest exponent of D in the series expansion is d_{free}
- However, in most cases, a closed form expression for $T(D)$ **may not be available**, and one has to resort to numerical techniques.

 Indian Institute of Technology, Delhi
11
Ranjan Bose
Department of Electrical Engineering

Now, if you see; if you do not really have a closed form solution, you can write as an expansion as follows. We had seen this earlier. Well, it comes directly from the closed form expression, but in general we can write like this. So, what we do is TD is nothing, but the number of paths with d_{free} times, this shows that the d_{free} is obtained for these paths and then there is a next expression where this is d_{free} plus something. So, it is the d_{next} and this $N(d_{next})$ corresponds to the number of paths that contribute to those d_{next} ok. So, d_{next} is the second smallest squared Euclidean distance.

So, the smallest exponent of D in the series is in the expansion is d_{free} , but we have made this point earlier that we may not have a closed form expression all the time and one has to resort to numerical methods.

Student: Sir, what is this d_{next} exactly I mean relate this and $d_{free} + 1$ then it will be $d_{free} + 2$ or something.

(Refer Slide Time: 08:24)

Information Theory, Coding and Cryptography

Computing d_{free}

- The generating function can be written as

$$T(D) = N(d_{free})D^{d_{free}^2} + N(d_{next})D^{d_{next}^2} + \dots$$

- where d_{next}^2 is the second smallest squared Euclidean distance.
- Hence the smallest exponent of D in the series expansion is d_{free}
- However, in most cases, a closed form expression for $T(D)$ **may not be available**, and one has to resort to numerical techniques.

 Indian Institute of Technology, Delhi

11

Ranjan Bose
Department of Electrical Engineering

The question being asked is what is this d_{next} ? So, clearly d_{free} is not an integer this time. Had it been, the hamming distance measured and we had taken in convolutional code, the generating function $T(D)$. Then this D raised to the power; if this was the 5, then d_{next} could be 6 or 7. But in this case d_{free} could be 3.722 and then the d_{next} could be 4.312. Some value which is the next larger value.

So, we have no constraint on this being integers here, because we are talking about the free distance of a trellis coded modulation scheme. So, whatever it is. In general, there will be some larger distance and that should correspond to certain number of paths that lead to it. So those suppose, there are 4 paths which contribute a distance Euclidean distance which is d_{next} , then n as a function of d_{next} times d raised to power d_{next} . And then there will be a next term, it will be $d_{next} + next$ and so and so forth.

(Refer Slide Time: 09:43)

Information Theory, Coding and Cryptography

Computing d_{free}

- Next, consider the function

$$\phi_1(D) = \ln \left[\frac{T(eD)}{T(D)} \right]$$

- $\phi_1(D)$ decreases monotonically to the limit d_{free}^2 as $D \rightarrow 0$.
- Therefore we have an upper bound on d_{free}^2 provided $D > 0$.
- In order to obtain a lower bound on d_{free}^2 consider the following function

$$\phi_2(D) = \frac{\ln T(D)}{\ln D}$$

 Indian Institute of Technology, Delhi
12
Ranjan Bose
Department of Electrical Engineering

So, let us consider the function; phi 1 as a function of D as log of the ratios t e D over T D all right. Now if you see this, phi 1 D decreases monotonically to the limit d free squared as D tends to 0. Therefore, we have an upper bound on free d free squared provided D is greater than 0, in order to obtain a lower bound on d free squared.

Consider the following function phi 2 which is the ratio of log T raise to power S sorry log T as a function of D divided by log of D. So, we have this function phi 1 and phi 2 and we will use them to get an estimate of d free.

(Refer Slide Time: 10:40)

Information Theory, Coding and Cryptography

Computing d_{free}

- Taking logarithm on both sides we get,

$$d_{\text{free}}^2 \ln D = \ln T(D) - \ln N(d_{\text{free}}) - \ln \left[1 + \frac{N(d_{\text{free}})}{N(d_{\text{next}})} D^{d_{\text{next}}^2 - d_{\text{free}}^2} = \dots \right]$$

- If we take $D \rightarrow 0$, provided $D > 0$, we obtain

$$\frac{\ln T(D)}{\ln D} = d_{\text{free}}^2 - \varepsilon(D)$$

- $\varepsilon(D)$ is a function that is greater than zero, and tends to zero monotonically as $D \rightarrow 0$.
- Thus, if we take smaller and smaller values of $\varepsilon(D)$ we can obtain values that are extremely close to d_{free}^2 .

 Indian Institute of Technology, Delhi
13
Ranjan Bose
Department of Electrical Engineering

So, if you take log on both sides of the equation that we saw just now, then we have $\ln D$ and the d_{free}^2 comes in front and then we have this expansion going up next. So, if we take D tending to the 0 and D greater than 0, we would get roughly $\ln T$ as a function of D divided by $\ln D$. This comes to d_{free}^2 minus ϵD .

So, the point is if we can make this ϵD as small as we can right. So, then, we can have a very good estimate on squared free distance. So, we take smaller and smaller values of ϵD and we can values we can get our values which are extremely close to squared free distance. So, this is one way to determine or at least get an estimate of squared free distance from the generating function $T(D)$.

(Refer Slide Time: 12:00)

Information Theory, Coding and Cryptography

Points about d_{free}

- It should be kept in mind that even though d_{free}^2 is the single most important parameter to determine the quality of a TCM scheme, two other parameters are also influential:
- The **error coefficient** $N(d_{\text{free}})$: A factor of two increase in this error coefficient reduces the coding gain by approximately 0.2 dB for error rates of 10^{-6} .
- The **next distance** d_{next} : is the second smallest Euclidean distance between two paths forming an error event.
- If d_{next} is very close to d_{free} , the SNR requirement for good approximation of the upper bound on P_e may be very large.
- So far, we have focused primarily on AWGN channels.
- We found that the best design strategy is to maximize the free Euclidean Distance, d_{free} , for the code.



Indian Institute of Technology,
Delhi

14

Ranjan Bose
Department of Electrical Engineering

So, couple of things that we must keep in mind, while talking about d_{free} . So, the squared free distance is the single most important parameter to determine the quality of a TCM scheme. So, if you have to compare two schemes. The first thing we should do is to compute d_{free} . Please note d_{free} calculation does not have to be done from the all 0 path because we are dealing with a non-linear code here; however, two of the parameters are also important.

The error coefficient N as a function of d_{free} which tells us that, how many paths which are there in the trellis which contribute to this d_{free} ? So, question is d_{free} appears to be the weakest link in the chain. How many weak links are there in this chain? Ok and after

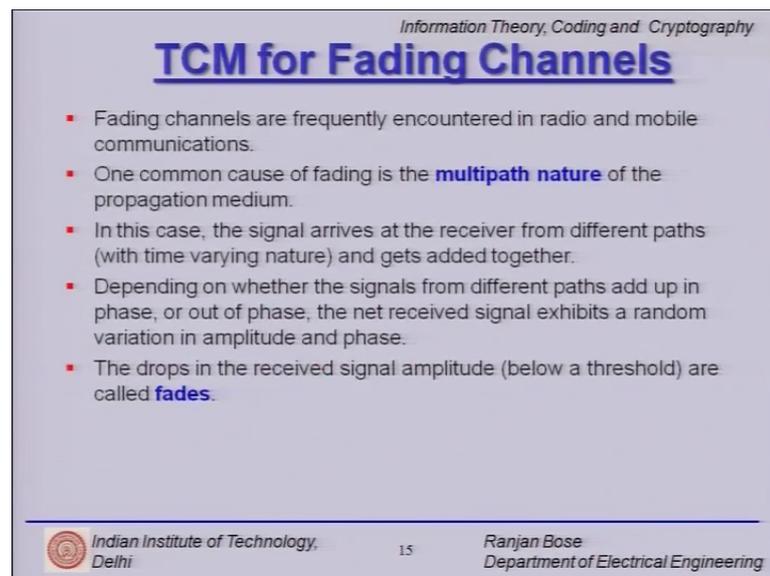
d_{free} , it is not always that we will always be in a bad shape; that we always get an error between the two closest paths. We can also have the d_{next} coming into play.

So, the second smallest Euclidean distance between two paths forming an error event is also important as a second order effect. Now question is how far are d_{free} and d_{next} ? If d_{next} is pretty large with respect to d_{free} , then d_{next} should stop mattering; however, if d_{next} is pretty close, why are we having this discussion while we did not have this for convolutional codes? That is because the d_{next} for the convolutional coder was always an integer away. So, if d_{free} was 2 or 4; then the d_{next} would be at least 1 or 2 more, but we can have d_{free} as 5.3 and d_{next} could be 5.4.

So, they can be fairly comparable to each other. So, if d_{next} is very close to d_{free} and it can be as close as you want or your luck may have it. The SNR requirement for good approximation on upper bound on P_e may be large.

Now so far, we have only talked about free distance and additive white Gaussian noise channels and we have found out of the best strategy is to maximize the free Euclidean distance for the code. But things may not be the same for fading channels. So, let us quickly talk about TCM design for fading channels. Before that let us take a minute to quickly recap, what do we mean by fading channels.

(Refer Slide Time: 14:38)



Information Theory, Coding and Cryptography

TCM for Fading Channels

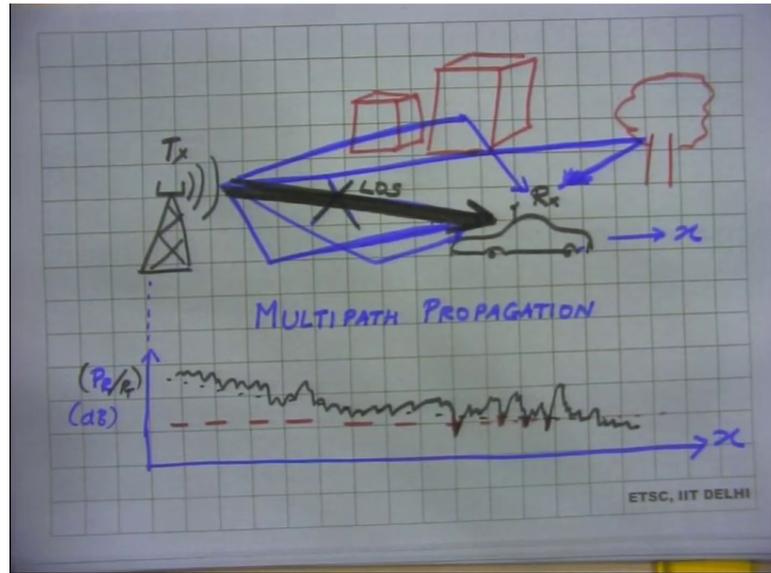
- Fading channels are frequently encountered in radio and mobile communications.
- One common cause of fading is the **multipath nature** of the propagation medium.
- In this case, the signal arrives at the receiver from different paths (with time varying nature) and gets added together.
- Depending on whether the signals from different paths add up in phase, or out of phase, the net received signal exhibits a random variation in amplitude and phase.
- The drops in the received signal amplitude (below a threshold) are called **fades**.

Indian Institute of Technology, Delhi

15

Ranjan Bose
Department of Electrical Engineering

(Refer Slide Time: 14:54)



So, if you consider a transmitter and a receiver, then we have the signals going from the transmitter to the receiver; however, it is possible that it reaches the receiver through reflections and the reflections, it can come from many kinds of objects. So, you could have different buildings or other reflectors which will lead to the signal reaching from the transmitter to the receiver through multipath.

So, currently we have drawn a scenario where we have multipath propagation. But the interesting thing is, when a signal reaches from the transmitter to the receiver through several paths, they add up at the receiver. Currently in phase out of phase, in such a manner that the resultant vector gives the actual value of the signal received. Now because of these paths reaching the receiver with delays because longer paths means more delay and hence a phase shift; that means, that they can add up and have a destructive interference also leading to the loss in the signal strength.

So, if you plot on this axis, the distance from the transmitter to the receiver and if you look at the power received and make a plot of the power received and if you do a logarithmic part you will have something like this. What it means is that; so if this is some kind of a plot in dB. So, we can normalize this.

So, if you see that even though as you move away from the transmitter and there appears to be a general downwardly trend. In the short term, there are major fluctuations in the received signal. And this is because that the signals received here add up destructively and the vector m gives a very low value.

So, this is typically of a profile of a fading channel and these are your fades. And if this is the receiver threshold over which the signal should work, then there would be many times when you go below the threshold and you may have a loss and communications.

So, there is the general nature of the fading channel and we would like to design a TCM scheme based on the characteristics of a fading channel. Now if there is a line of sight here, then we know that this is a Rician fading channel. However, if there is no line of sight, then we go to Rayleigh fading channel.

So, we will look at these scenarios separately. We come back to our slides now and we observe that fading channels are frequently encountered in radio and mobile communications and the cause of fading is the multipath nature as we discussed just now. And this is because the receiver gets signals from different paths with different path lengths. And if the signals add up in phase or out of phase, the net received signal exhibits a random variation in amplitude and phase which are the sudden drops are called fades.

(Refer Slide Time: 19:48)

Information Theory, Coding and Cryptography

TCM for Fading Channels

- Lets consider the performance of trellis coded M -ary Phase Shift Keying (MPSK) over a fading channel.
- We know that a TCM encoder takes in an input bit stream and outputs a sequence of symbols.
- In this treatment we will assume that each of these symbols s_i belong to the MPSK signal set.
- By using complex notation, each symbol can be represented by a point in the complex plane.
- The coded signals are **interleaved** in order to spread the burst of errors caused by the slowly varying fading process.
- These interleaved symbols are then pulse-shaped for no **inter-symbol interference** and finally translated to RF frequencies for transmission over the channel.

 Indian Institute of Technology,
Delhi

16

Ranjan Bose
Department of Electrical Engineering

So, we now talk about TCM for fading channels. Let us consider the performance of trellis coded M -ary Phase Shift Keying MPSK over a fading channel all right. We know we know that a TCM encoder takes in an input bit stream and outputs a sequence of symbols. In this treatment, we will assume that each of these symbols s_i belong to the MPSK signal set.

So, this is clear we have done this before and we will use a complex notation. So, each symbol is represented by a point on the complex plane. Now the coded symbols are interleaved in order to spread out the burst errors caused by slowly varying fading process and this interleaved symbols are then pulse shape for no inter symbol interference. So, those are the nitty-gritty details for communications, but will focus pretty much on the fading characteristic of the channel.

(Refer Slide Time: 20:40)

Information Theory, Coding and Cryptography

TCM for Fading Channels

- The channel corrupts these transmitted symbols by adding a **fading gain** (which is a negative gain, or a positive loss, depending on one's outlook) and AWGN.
- At the receiver end the received sequences are demodulated and quantized for soft decision decoding.
- In many implementations, the channel estimator provides an estimate of the channel gain, which is also termed as the **channel state information**.
- Thus we can represent the received signal at time i as

$$r_i = g_i s_i + n_r$$

Indian Institute of Technology,
Delhi

17

Ranjan Bose
Department of Electrical Engineering

Now, the channel corrupts these signals by adding a fading gain right and it also had has additive white Gaussian noise. So, the fading gain is a multiplicative gain. So, at the receiver end, the received sequences are demodulated and quantized for soft decision decoding. And in many implementation, the channel estimator provides an estimate of the channel gain. So, we would either have the channel state information built into it or sometimes we will have to go without it for cheaper designs and then we say the CSI the channel state information is not present. So, they are two distinct scenarios.

So, at any instant, i time instant i the received signal r_i is equal to g_i times s_i , s_i is the transmitted signal g_i is the multiplicative gain which is the fading coefficient plus the noise n_i . This is a very simplistic model that we will be using for our case.

(Refer Slide Time: 21:45)

Information Theory, Coding and Cryptography

Fading Channels

- $r_i = g_i s_i + n_i$
- where n_i is a sample of the zero mean gaussian noise process with variance $N_0/2$ and g_i is the complex channel gain, which is also a sample of a complex gaussian process with variance σ_g^2 .
- The complex channel gain can be explicitly written using the phasor notation as follows

$$g_i = a_i e^{j\phi_i}$$

- where, a_i and ϕ_i are the amplitude and phase processes respectively.

 Indian Institute of Technology, Delhi

18

Ranjan Bose
Department of Electrical Engineering

So, starting with r_i is equal to $g_i s_i + n_i$, we say that n_i is a sample of the zero mean Gaussian noise process with variance $N_0/2$; g_i is the complex channel gain right which is also sample of a complex Gaussian process. So, the g_i can be written as a_i times $e^{j\phi_i}$ because g_i is complex gain. It not only changes the amplitude, but also the phase remember, s_i is complex.

So, complex into complex gives another complex. So, a_i and ϕ_i are the amplitude and phase processes respectively and they can be modelled differently. For example, depending upon, whether you have a line of sight or not a_i could be Rayleigh or Rician and ϕ_i is typically uniformly distributed over 0 to 2π

(Refer Slide Time: 22:43)

Information Theory, Coding and Cryptography

Fading Channels

- **Assumptions**
- The receiver performs coherent detection,
- The interleaving is ideal, which implies that the fading amplitudes are statistically independent and the channel can be treated as memory less.
- Thus we can write

$$r_i = a_i s_i + n_i$$

- We know that for a channel with a diffused multipath and no direct path the fading amplitude is Rayleigh distributed with pdf

 Indian Institute of Technology, Delhi 19 Ranjan Bose
Department of Electrical Engineering

So, let us see; if we can start with some assumptions. Let us say that the receiver performs coherent detection, not a bad assumption. Interleaving is ideal. So, we have the fading and amplitude are statistically independent and the channel can be treated as memory loss channel. So, we have r_i is equal to $a_i s_i + n_i$ and we know that for a channel with diffused multipath with no direct line of sight communication between the transmitter and receiver the fading can be rightly modelled as Rayleigh distributed.

(Refer Slide Time: 23:22)

Information Theory, Coding and Cryptography

Rayleigh Fading Channels

- We know that for a channel with a **diffused multipath** and no direct path the fading amplitude is **Rayleigh** distributed with pdf

$$p_A(a) = 2ae^{-a^2}$$

With $a \geq 0$.

 Indian Institute of Technology, Delhi 20 Ranjan Bose
Department of Electrical Engineering

(Refer Slide Time: 23:30)

Rician Fading Channels

- For the case when there exists a direct path in addition to the multipath, Rician fading is observed.
- The pdf of the Rician fading amplitude is given by

$$p_A(a) = 2a(1+K)e^{-(K+a^2(K+1))}I_0(2a\sqrt{K(1+K)})$$

- where $I_0(\cdot)$ is the zero-order, modified Bessel function of the first kind and K is the Rician parameter defined as follows.



So, we write the pdf of the Rayleigh fading channel as follows. For the case where there exists a direct path in addition to the multipath Rician fading is observed and the following expression gives the pdf of the Rician fading.

Now K is the Rician fading parameter and we will define it shortly and I is the zero ordered modified Bessel function of the first kind. So, this K will play an important role and let us understand what K is.

(Refer Slide Time: 24:00)

Rician Fading Channels

- The **Rician Parameter** K is defined as the ratio of the energy of the direct component to the energy of the diffused multipath component.
- For the extreme case of $K = 0$, the pdf of the Rician distribution becomes the same as the pdf of the Rayleigh distribution.
- We now look at the performance of the TCM scheme over a fading channel.
- Let $\mathbf{r}_l = (r_1, r_2, \dots, r_l)$ be the received signal.
- The maximum likely decoder, which is usually implemented by the Viterbi decoder, chooses the coded sequence that most likely corresponds to the received signals.
- This is achieved by computing a metric between the sequence of received signals, \mathbf{r}_l and the possible transmitted signals, \mathbf{s}_l .
- As we have seen earlier, this **metric** is related to the conditional channel probabilities



K is the Rician parameter. What is it? It is the ratio of the energy of the direct component of the signal to the diffused multipath component all right. So, if you go back

to the diagram, we drew and we talk about the direct path. So, there is a certain amount of energy that is coming through the direct path call the line of sight. And then, there are so many other reflected paths which are carrying this energy to the receiver.

So, the ratio of the energy coming from the direct path to the non line of sight, the diffused path is the parameter K. So, clearly if there is no line of sight; then K is 0 and so, K is equal to 0 implies that we are looking at a standard Rayleigh fading channel. So, a special case K is equal to 0 comes back to the Rician Rayleigh fading distribution.

So, now let us look at the performance of a TCM scheme over the fading channel. Let us say that the received signal be r_1, r_2 up to r_l . This is a vector. The maximum likelihood decoder which we will say, we will implement using the Viterbi decoder; chooses the coded sequence that is most likely to correspond to the received signals. So, this is achieved by computing a metric between the sequences of received signal r_l . Here is the received signal and the possible transmitted signals s . And we have seen that metric is related to the conditional channel probabilities.

(Refer Slide Time: 25:57)

Information Theory, Coding and Cryptography

Rician Fading Channels

$$m(\mathbf{r}_l, \mathbf{s}_l) = \ln p(\mathbf{r}_l | \mathbf{s}_l)$$

- If the channel state information is being used, the metric becomes

$$m(\mathbf{r}_l, \mathbf{s}_l; \hat{\mathbf{a}}_l) = \ln p(\mathbf{r}_l | \mathbf{s}_l, \hat{\mathbf{a}}_l)$$

- Under the assumption of ideal interleaving, the channel is memoryless and hence the metrics can be expressed as the following summations

$$m(\mathbf{r}_l, \mathbf{s}_l) = \sum_{i=1}^l \ln p(r_i | s_i)$$

- and

$$m(\mathbf{r}_l, \mathbf{s}_l; \hat{\mathbf{a}}_l) = \sum_{i=1}^l \ln p(r_i | s_i, \hat{a}_i)$$

Indian Institute of Technology,
Delhi

23

Ranjan Bose
Department of Electrical Engineering

So this metric, we can take as log of the probability that r_l ; this vector is received given s_l is transmitted. So, now, we can say that we have the access to the channel state information. So, the fading coefficients are known. So, at least the estimates are given. So, this metric also has r_l comma s_l semicolon \hat{a}_l right and this is the metric that comes in though.

We have the probability of r_l given s_l and I have got the estimates right. So, under the assumption that the ideal interleaving is happening, we can assume that the channel is memory less. So, each symbol do not have any implication on the subsequent symbol. Therefore, we have this product because if they are independent, then the probabilities would multiply, but since there is a log. Now that is a summation there.

So, we have, now i is equal to 1 through L ; $\ln p r_l$ given s_l because we have declared that each symbol received is independent of this earlier symbols. And if you have the channel state information, you can write this summation i is equal to 1 through L ; $\ln p r_l$ given s_l comma \hat{a}_l . So, these are the two matrix with which we will work.

(Refer Slide Time: 27:30)

Information Theory, Coding and Cryptography

Rician Fading Channels

- First we consider the scenario where the channel state information is known, i.e., $\hat{a}_i = a_i$
- The metric can be written as

$$m(r_i, s_i; \hat{a}_i) = -|r_i - a_i s_i|^2$$
- Therefore, the pairwise error probability is given by

$$P_2(s_l, \hat{s}_l) = E_{a_l} [P_2(s_l, \hat{s}_l | a_l)]$$
- where E is the statistical expectation operator.

$$P_2(s_l, \hat{s}_l | a_l) = P[m(r_l, \hat{s}_l; a_l) \geq m(r_l, s_l; a_l) | a_l]$$

Indian Institute of Technology, Delhi
24
Ranjan Bose
Department of Electrical Engineering

Now first we consider the scenario, when the channel state information is known correctly all right. So, the metric can now be written as minus r_i minus $a_i s_i$ because a_i is known right. So, received and we have to see whether s_i into a_i and this is the distance measure, we have. So, the pair wise error probability P_2 between s_l and s_l hat is given by E_{a_l} . This is the expectation operator; statistical expectation operator and this is $P_2(s_l, s_l$ hat given a_l . On what does this P_2 stand for?

Well, it talks about the sequence s_l and s_l hat which is estimate; given the channel sate information as the metric of s_l hat right with r_l greater than or equal to between r_l and s_l . So, how do I choose the most likely? So wherever, which maximizes the metric given of course, this one. So, it is pretty straightforward, if you write it out.

(Refer Slide Time: 28:54)

Information Theory, Coding and Cryptography

TCM for Rician Fading Channels

- Using the Chernoff bound, the pairwise error probability can be upper bounded as follows.

$$P_2(s_i, \hat{s}_i) \leq \prod_{i=1}^L \frac{1+K}{1+K + \frac{1}{4N_0} |s_i - \hat{s}_i|^2} \exp \left[-\frac{K \frac{1}{4N_0} |s_i - \hat{s}_i|^2}{1+K + \frac{1}{4N_0} |s_i - \hat{s}_i|^2} \right]$$

- For high SNR, the above equation simplifies to

$$P_2(s_i, \hat{s}_i) \leq \prod_{i \in \eta} \frac{(1+K)e^{-K}}{\frac{1}{4N_0} |s_i - \hat{s}_i|^2}$$

- where η is the set of all i for which $s_i \neq \hat{s}_i$
- Let us denote the number of elements in η by L_η then we can write

 Indian Institute of Technology, Delhi
 25
Ranjan Bose
Department of Electrical Engineering

So, we use Chernoff bounds and we can write the upper bound on the pair wise error probability as follows. So, this is the pair wise error probability between s_1 and \hat{s}_1 . This is the expression that comes out where K is the Rician parameter. So, if you see for high SNR scenario right; so well, this is low. So, this expression is large and you can neglect this, when the denominator you are left with this. So, you can write out $P_2(s_1, \hat{s}_1)$ as a product where here; this η is the set of all i for which s_i is not equal to \hat{s}_i . So, that is why we are talking about the error probability.

So, let us denote the number of elements in η by L_η ; so, L_η . So, we are defining this term L_η . What is L_η ? Where is the number of elements in this? So, number of cases where s_i is not equal to \hat{s}_i . So, error happens.

(Refer Slide Time: 30:14)

Information Theory, Coding and Cryptography

TCM for Rician Fading Channels

$$P_2(s_i, \hat{s}_i) \leq \frac{((1+K)e^{-K})^\eta}{\left(\frac{1}{4N_0}\right)^{l_\eta}} d_p^2(l_\eta)$$

- where

$$d_p^2(l_\eta) \leq \prod_{i \in \eta} |s_i - \hat{s}_i|^2$$

is the **Squared Product Distance** of the signals $s_i \neq \hat{s}_i$.

- The term l_η is called the **Effective Length** of the error event $s_i \neq \hat{s}_i$

 Indian Institute of Technology,
Delhi
26
Ranjan Bose
Department of Electrical Engineering

So, if you write it down, then from here you can see in terms of the product i is an element of η . Therefore, you get here raised to power l_η because only those number of terms multiply. So, you have got this expression now.

So, it brings us to a very interesting conclusion. You can see that we have introduced this term, d_p^2 as a function of l_η . What is d_p^2 ? d_p^2 is the product of this squared Euclidean distance; where i is an element of η . All those paths where you do not have s_i is equal to \hat{s}_i . So, this d_p^2 ; is the squared product distance of the signal s_i not equal to \hat{s}_i . This will be a very very important parameter for fading channels; specifically we are talking about Rician fading channel.

So, the term l_η is called the effective length of the error event s_i not equal to \hat{s}_i . We will use these two extensively for the next few slides, but it is so happens that from the derivation, we have come to this product; the squared product distance.

(Refer Slide Time: 31:40)

Information Theory, Coding and Cryptography

High SNR Scenario

$$P_e \leq \sum_{l_\eta} \sum_{d_p^2(l_\eta)} \alpha(l_\eta, d_p^2(l_\eta)) \frac{((1+K)e^{-K})^{l_\eta}}{\left(\frac{1}{4N_0}\right)^{l_\eta} d_p^2(l_\eta)}$$

where $\alpha(l_\eta, d_p^2(l_\eta))$ is the average number of code sequences having the effective length l_η and the squared product distance $d_p^2(l_\eta)$.

 *Indian Institute of Technology, Delhi* 27 *Ranjan Bose*
Department of Electrical Engineering

So, if you talk about the high SNR situation, you can write in the upper bound as P_e less than or equal to this summation over this, effective length and summation over all the d_p squared with effective length. And then you have this expression where this first term, α as a function of effective length and squared product distance is the average number of code sequences having the effective length l_η and the square product distance d_p square.

So, please note, we are doing sequence by sequence decoding and we have come across a set of sequences which have the effective length l_η and a corresponding square product distance d_p squared.

(Refer Slide Time: 32:43)

Information Theory, Coding and Cryptography

Error Event Probability

- The error event probability is actually dominated by the smallest effective length l_η and the smallest product distance $d_p^2(l_\eta)$.
- Let us denote the smallest effective length by L and the corresponding product distance by $d_p^2(L)$.
- The error event probability can then be asymptotically approximated by

$$P_e \approx \alpha(L, d_p^2(L)) \frac{(1 + K)e^{-K}}{\left(\frac{1}{4N_0}\right)^L d_p^2(L)}$$

 Indian Institute of Technology,
Delhi28Ranjan Bose
Department of Electrical Engineering

So, the error event probability is actually dominated by the smallest effective length l_η and the smallest product distance d_p^2 . So, this is an important observation. So, let us denote this smallest effective length by l and the corresponding product distance by $d_p^2(l)$. Since, these are the two important parameters which dictate the error event probability. Then asymptotically the error event probability can be written as approximately $P_e \approx \alpha$, we have talked about this. And in the numerator, we have $1 + Ke^{-K}$ raised to power L . This corresponds to the noise power and squared product distance.

So, the effective length L and the squared product distance, prominently figure out in the error probability expression and so, we should have our designed criteria based on these

(Refer Slide Time: 33:55)

Information Theory, Coding and Cryptography

Error Event Probability

$$P_e \approx \alpha(L, d_p^2(L)) \frac{(1+K)e^{-K}}{\left(\frac{1}{4N_0}\right)^L d_p^2(L)}$$

- We make the **following observations**
- The error event probability asymptotically varies with the L^{th} power of SNR.
- This is similar to what is achieved with a time diversity technique.
- Hence, L is also called the **time diversity** of the TCM scheme.
- The important TCM design parameters are the time diversity, L , and the product distance $d_p^2(L)$.
- This is in contrast to the free Euclidean distance parameter for AWGN channel.

 Indian Institute of Technology,
Delhi29Ranjan Bose
Department of Electrical Engineering

So, we have taken down the same expression on this slide and we make the following observations. So, the error event probability asymptotically varies with the L^{th} power of SNR here; this term L^{th} power of SNR. This is similar to what is achieved by time diversity techniques. This is very important.

So, this observation tells me that here, we are getting a diversity gain because of the TCM scheme right. So, L is also called the time diversity of the TCM scheme. We are not explicitly using any time diversity schemes. If you remember, time diversity means that you can transmit the same information in different instances in time; assuming that the spacing between these times is not correlated and you get time diversity gain.

Here L is the time diversity of the TCM scheme; simply because we have an effective length L and so, a design criteria should be such that we maximize this effective length L so as to gain time diversity in terms of the TCM scheme.

The important TCM design parameter are the time diversity L and the product distance d_p^2 . So, we will use them for our design or comparison of two TCM schemes in fading channel. Please note, we are not talking about d_{free} . There are two different parameters that we have come across; one is a time diversity L and the other is the product distance d_p^2 . This is in contrast to the free Euclidean distance parameter for the additive white Gaussian noise channel. So, what works for fading channel, may not work for AWGN channel and vice versa.

(Refer Slide Time: 36:00)

Information Theory, Coding and Cryptography

Error Event Probability

$$P_e \approx \alpha(L, d_p^2(L)) \frac{\left((1+K)e^{-K} \right)^L}{\left(\frac{1}{4N_0} \right)^L d_p^2(L)}$$

- TCM codes designed for AWGN channels would normally fare poorly in fading channels and vice-versa.
- For large values of the Rician parameter, K , the effect of the **free Euclidean distance** on the performance of the TCM scheme becomes dominant.
- At low SNR, **again, the free Euclidean distance** becomes important for the performance of the TCM scheme.

 Indian Institute of Technology, Delhi30Ranjan Bose
Department of Electrical Engineering

So, this observation we have made and for large values of the Rician parameters K , the effect of free distance on the performance scheme becomes dominant; that is because if you go to very large value of K , then it starts resembling the AWGN channel; the diffused components whether away and it there is pretty much a line of sight, direct line of sight between the transmitter and receiver. And their Gaussian noise plays dominant role and therefore, free Euclidean distance becomes important. You can also see that when K becomes large this terms kinds of vanishes and we are pretty much left with the Euclidean distance part.

Now at low SNR, things again change. Because now at low in SNR part, this component right, we will play start playing a role and then you will have; even though the value of K is not large, then the SNR starts playing this role. So, again the free Euclidean distance becomes important for TCM designing.

(Refer Slide Time: 37:20)

Information Theory, Coding and Cryptography

Design Rules

- Thus the basic *design rules* for TCMs for fading channels, at high SNR and for small values of K , are
- maximize the effective length, L , of the code, and
- minimize the minimum product distance $d_p^2(L)$.
- Consider a TCM scheme with effective length, L , and the minimum product distance $d_{p_1}^2(L)$.
- Suppose the code is redesigned to yield a minimum product distance, $d_{p_2}^2(L)$ with the same L .
- The increase in the coding gain due the increase in the minimum product distance is given by

$$\Delta g = SNR_1 - SNR_2 \Big|_{P_{e_1} = P_{e_2}} = \frac{10}{L} \log \frac{d_{p_2}^2(L) \alpha_1}{d_{p_1}^2(L) \alpha_2}$$

 Indian Institute of Technology,
Delhi
31
Ranjan Bose
Department of Electrical Engineering

So, let us look at the basic design rules for TCM for fading channels at highest SNR's and small values of K ok. Why small values of K ? If the K becomes too large, then it is no longer really making a fading channel; that are the design rules? Maximize the effective length L of the code and minimizing the minimum product distance d_p^2 ok. These are the two designs.

So, if you consider as TCM scheme with effective length L and their minimum product distance d_p^2 ; suppose that the code is redesigned to yield a minimum product distance d_p^2 with the same L . Then increase in the coding gain due to the increase in their minimum product distance is given as follows. So, this is the coding gain and we have the same effective length, but the product distance changes. So, if you have been able to decrease the product distance. That means, your gain will improve and therefore, you can have a relative coding gain as follows.

(Refer Slide Time: 38:39)

Information Theory, Coding and Cryptography

Coding Gain

$$\Delta g = SNR_1 - SNR_2 \Big|_{P_{e1} = P_{e2}} = \frac{10}{L} \log \frac{d_{p_2}^2(L) \alpha_1}{d_{p_1}^2(L) \alpha_2}$$

- where $\alpha_i, i = 1, 2$, is the average number of code sequences with effective length L for the TCM scheme i .
- We observe that for a fixed value of L , increasing the minimum product distance corresponding to a smaller value of L is more effective in improving the performance of the code.

 Indian Institute of Technology, Delhi 32 Ranjan Bose
Department of Electrical Engineering

So, here the parameter alpha; alpha 1 and alpha 2 are the average number of code sequences with effective length L. So, it also matters how many sequences in the trellis diagram have that effective length. So, we observe that for a fixed value of L, increasing the minimum product distance corresponding to a smaller value of L is more effective in improving the performance of the code because here the smaller value of the L will give you a better.

So, at on one side you have to increase L, but when you have tweaking the code; then the gain will be more, if the increase is for a smaller L because L is also here. And please note that the product distance squared is in the logarithm whereas, this is inversely proportional.

(Refer Slide Time: 39:45)

Information Theory, Coding and Cryptography

Without CSI

- So far we have assumed that the channel state information was available.
- A similar analysis as carried out for the case where channel state information was available can also be done when the information about the channel is unavailable.
- In the *absence* of channel state information, the metric can be expressed as

$$m(r_i, s_i; \hat{a}_i) = - | r_i - s_i |^2$$

 Indian Institute of Technology, Delhi
33
Ranjan Bose
Department of Electrical Engineering

Now so far, we have assumed that the channel state information was indeed available, but we can carry out a analysis when the CSI is not available. So, what will change is the metric? Earlier we had alpha here, the gain parameter. Now, we do not have. So, the best way, it is r i minus s i. So, this is the metric we define.

(Refer Slide Time: 40:08)

Information Theory, Coding and Cryptography

Without CSI

- After some mathematical manipulations it can be shown that

$$P_2(s_l, \hat{s}_l) \leq \frac{(2e/l_\eta)^{l_\eta} \left(\sum_{i \in l_\eta} |s_i - \hat{s}_i|^2 \right)^{l_\eta}}{(1/N_0)^{l_\eta} d_p^4(l_\eta)} (1+K)^{l_\eta} e^{-l_\eta K}$$

- Using arguments discussed earlier, the error event probability P_e can be determined for this case when the channel state information is not available.

 Indian Institute of Technology, Delhi
34
Ranjan Bose
Department of Electrical Engineering

So, we can do similar mathematical analysis and we can get this pair wise error probability upper bounded by the following expression ok. Again the effective length and K show a prominent prominently in the expression. But here please note that we have d

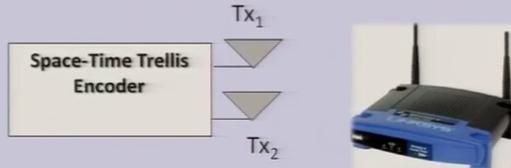
p raise to power 4 ok. So, it is even more sensitive here. So, the error event probability P can be determined for this case, when the channel state information is not available.

(Refer Slide Time: 40:43)

Information Theory, Coding and Cryptography

Space Time Trellis Codes

- **Space Time Trellis Codes** (STTC) are coding techniques designed for multiple-antenna transmissions.
- STTCs are TCM schemes in which every branch of the trellis is labeled by N_t **signals** corresponding to the signals transmitted by the N_t **transmit antennas**.



The diagram shows a box labeled 'Space-Time Trellis Encoder' on the left. Two lines extend from the right side of the box to two antenna symbols labeled 'Tx₁' and 'Tx₂'. To the right of the antennas is a photograph of a blue wireless router with two antennas.

 Indian Institute of Technology, Delhi35Ranjan Bose
Department of Electrical Engineering

So, with that we kind of lay grounds for a next step which is called the space time Trellis Codes; which is in one way an extension of trellis coded modulation schemes. But what we do is we now look at expressions and cases where we have two transmit antennas.

So, if you see, earlier we had only a 1 antenna to be the trellis and input sequence came in, we looked at that trellis on each branch. We had a single symbol labeled and the transmitter transmitted at. What if we label every branch with a pair of symbols? So, we have now two transmitters. So, symbol one is sent from transmitter 1 and symbol two is sent from transmitter 2. It is a very logical extension and I can extend it to n antennas, N_t antennas; standing in front n transmit antennas.

So, what happens, if we do this ok? This kind of a scheme is called space time trellis codes because we also not only use the space coordinate, but also time coordinates. So, in different time instances, we have different pairs being sent from transmitter 1 and transmitter 2. So, we are curious to finding out, what happens.

So, now, we go to multi antenna systems. So, we have TCM schemes where every branch is labelled with N_t signals. In this example, there are two transmitter antennas.

So, every branch will have two symbols, if you have N_t transmitter antennas; I will have N_t symbols.

(Refer Slide Time: 42:34)

Information Theory, Coding and Cryptography

Example

4-PSK

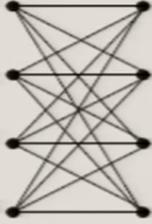


00 01 02 03

10 11 12 13

20 21 22 23

30 31 32 33



A 2-STTC, 4-PSK, 4 states, 2 bits/s/Hz

The **edge label** is denoted by $c_1 c_2$ where
 c_1 is transmitted from antenna 1 and
 c_2 is transmitted from antenna 2.



Indian Institute of Technology,
Delhi

36

Ranjan Bose
Department of Electrical Engineering

Let us look at a very simple example. So, there are four states in this trellis and it is sufficient to draw the first portion of the trellis diagram because it will repeat infinitely after this. There are 4 emanating branches from every node. So, there should be 4 labels here. Now this is symbol S ; S_0, S_1, S_2, S_3 . So, this is the 4-PSK constellation, QPSK.

Now in a normal TCM schemes each branch will be labelled with either s_0 or for a short $0 S_2$ or $2 S_4$. Or here it is QPSK, so $S_0 S_1 S_2 S_3$; so, they will be labelled here. But here you see there is a pair. So, $0 0$ means, if this branch is to be traversed; then transmitter antenna 1 should sent S_0 , transmitter antenna 2 should sent S_0 .

If we go on this branch, then transmitter antenna 1 should send S_0 , transmitter antenna 2 should send S_1 and so and so forth. So, this is a mapping scheme. So, if the label is $c_1 c_2$ for any edge, then c_1 is transmitted from antenna 1 and c_2 is transmitted from antenna 2; a very simple extension of the TCM scheme.

(Refer Slide Time: 44:00)

Information Theory, Coding and Cryptography

Decoding

- For decoding of a STTC code, **vector Viterbi algorithm** is required, assuming ideal channel state information at the receiver.
- Assuming that r_t^j is the received signal at receive antenna j at time t and α_{ij} is the path gain from transmit antenna i to receive antenna j , the branch metric for a transition labeled $q_t^1 q_t^2 \dots q_t^{N_r}$ is given by
$$\sum_{j=1}^{N_r} \left| r_t^j - \sum_{i=1}^{N_t} \alpha_{ij} q_t^i \right|^2$$
- The Viterbi algorithm is used to compute the path with the lowest accumulated metric.**

 Indian Institute of Technology, Delhi 37 Ranjan Bose
Department of Electrical Engineering

So for decoding, we would use vector Viterbi decoding right and we can assume that channel state information is present at the receiver. And let us see, what how do we define it. So, assuming that r_t^j is the received signal and the received antenna j at time t and α_{ij} is the path gain from the transmitter antenna i to receive antenna j and branch metric of the transmission is labelled as follows.

So, this is $q_1^t, q_2^t \dots$ up to q_t^t . So, then we have all the path gains in the transmission path times the received signal and minus r_t^j . So, the Viterbi algorithm is used to compute the path with the lowest accumulated metric. So, please note, what we are doing is defining the metric slightly differently taking into consideration both or all the N_t signals that are transmitted

(Refer Slide Time: 45:18)

Information Theory, Coding and Cryptography

Encoder for STTC

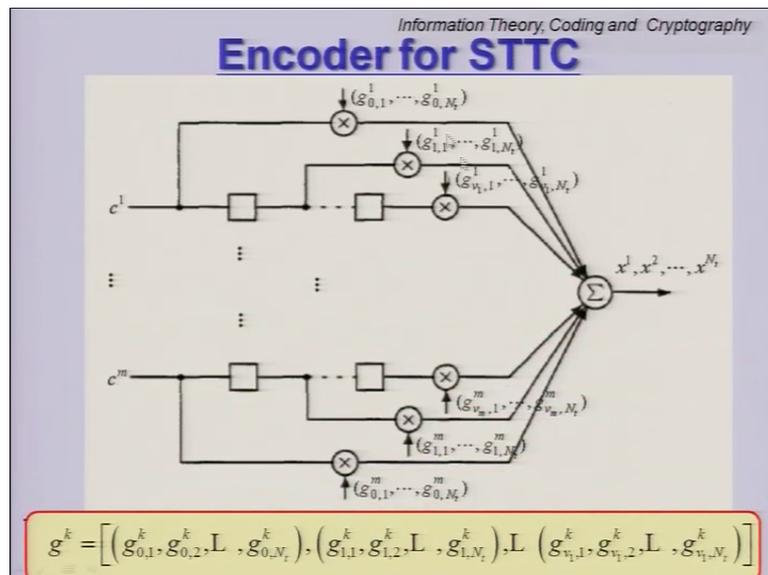
- Consider a generic encoder structure for a full rate **M-PSK STTC**.
- This encoder structure would be valid in general for a full rate STTC with any other M -ary signal constellation.
- The encoder consists of $m = \log_2 M$ feedforward shift registers into which m binary input sequences c^1, c^2, \dots, c^m are fed.
- The multiplier coefficient set for the k^{th} shift register is denoted by

$$g^k = \left[\left(g_{0,1}^k, g_{0,2}^k, \dots, g_{0,N_t}^k \right), \left(g_{1,1}^k, g_{1,2}^k, \dots, g_{1,N_t}^k \right), \dots, \left(g_{v_i,1}^k, g_{v_i,2}^k, \dots, g_{v_i,N_t}^k \right) \right]$$

Indian Institute of Technology, Delhi
38
Ranjan Bose
Department of Electrical Engineering

So, consider a generic encoder structure for a full rate M-PSK STTC; space time trellis code right. So, what we will do is in the next slide we will look at the generic structure of this and it will be better to start with the circuit diagram and or a block diagram of this encoder first.

(Refer Slide Time: 45:44)



So, we have c_1, c_2 up to c_m coming in and then we have these coefficients that we just now looked at. They are multiplied and then they add and then this is what is sent out.

This g^k is represented by as follows. This is the general encoder structure for a space time trellis code.

(Refer Slide Time: 46:19)

Information Theory, Coding and Cryptography

Encoder for STTC

$$g^k = \left[\left(g_{0,1}^k, g_{0,2}^k, \dots, g_{0,N_t}^k \right), \left(g_{1,1}^k, g_{1,2}^k, \dots, g_{1,N_t}^k \right), \dots, \left(g_{v_k-1,1}^k, g_{v_k-1,2}^k, \dots, g_{v_k-1,N_t}^k \right) \right]$$

- where

$$g_{j,i}^k, k = 1, 2, \dots, m, j = 1, 2, \dots, v_k, i = 1, 2, \dots, N_t$$
- is an element of M -PSK (or M -ary) constellation set and v_k is the memory order of the k^{th} shift register.
- The multiplier outputs from all shift registers are added modulo M , giving the encoder output $x = (x^1, x^2, \dots, x^{N_t})$
- The total memory order of the encoder, denoted by v , is given by

$$v = \sum_{k=1}^m v_k \quad v_k = \left\lfloor \frac{v+k-1}{\log_2 M} \right\rfloor$$

 Indian Institute of Technology, Delhi 40 Ranjan Bose
Department of Electrical Engineering

So the what we can do is, we can use that diagram to understand that the multiplier outputs from all the shift registers are added modulo M and we will give the encoder output as follows x^1, x^2 dot dot dot up to x^{N_t} and the total memory order of the encoder is denoted by v is as follows.

(Refer Slide Time: 46:52)

Information Theory, Coding and Cryptography

Design Criteria

- **Slow Rayleigh Fading:** For each input symbol the space-time encoder generates N_t code symbols $c_t^1, c_t^2, \dots, c_t^{N_t}$, which are simultaneously transmitted from the N_t transmit antennas.
- We define the code vector as $c_t = [c_t^1, c_t^2, \dots, c_t^{N_t}]^T$
- Suppose that the code vector sequence $C = \{c_1, c_2, \dots, c_J\}$ is transmitted.
- Consider the pairwise error probability (PEP) that the ML decoder decides erroneously in favour of the legitimate code vector sequence $\tilde{C} = \{\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_J\}$
- Consider a frame of length J and define the $N_t \times N_t$ error matrix A as

$$A(C, \tilde{C}) = \sum_{t=1}^J (c_t - \tilde{c}_t)(c_t - \tilde{c}_t)^H$$

 Indian Institute of Technology, Delhi 41 Ranjan Bose
Department of Electrical Engineering

So, we have 2 conditions; Slow Rayleigh Fading and fast Rayleigh fading. So, if you consider the slow Rayleigh fading, so for each symbol the space time encoder generated N_t code symbols c_1, c_2, \dots, c_{N_t} which are simultaneously transmitted from the N_t transmit antennas that we have seen.

Now we define the code vector as follows right. So, this is the code vector sequence which is transmitted and the ML decoder decides to erroneously find the \tilde{C} as follows. So, transmitted sequence, received sequence. So, frame of length J is considered. So, we have now an error matrix as follows which is an N_t cross N_t error matrix \mathbf{A} and we will use this for design criteria.

What is \mathbf{A} as a function of \mathbf{C} and $\tilde{\mathbf{C}}$? \mathbf{C} was sent $\tilde{\mathbf{C}}$ was received. It is we define it as $c_t - \tilde{c}_t$ times $c_t - \tilde{c}_t$ complex conjugate.

(Refer Slide Time: 48:10)

Information Theory, Coding and Cryptography

Design Criteria

- If ideal channel state information is available at the receiver, then it is possible to show that the pairwise error probability (PEP) of erroneous detection is given by

$$P(\mathbf{C} \rightarrow \tilde{\mathbf{C}}) \leq \left(\prod_{i=1}^r \lambda_i \right)^{-N_r} (E_s/4N_0)^{-rN_r}$$
- Where E_s is the symbol energy, N_0 is the noise power spectral density, N_r is the number of receive antennas, r is the rank of the error matrix \mathbf{A} and $\lambda_i, i=1, \dots, r$ are the non-zero eigenvalues of \mathbf{A} .
- Thus, a diversity gain of rN_r and a coding gain of $\left(\prod_{i=1}^r \lambda_i \right)^{1/r}$ is obtained.
- Minimizing maximum PEP leads to following design criteria ➔

Indian Institute of Technology, Delhi
42
Ranjan Bose
Department of Electrical Engineering

So, if ideal channel state information is available at the receiver, then it is possible to show that the pair wise error probability of their erroneous detection signal is given as follows. So, we do not go for the derivation, but this is the pair wise error probability that \mathbf{C} was sent and $\tilde{\mathbf{C}}$ was received, at least decoded incorrectly as follows. And we find very interestingly that the Eigen values of the \mathbf{A} matrix come into picture. N_r is the number of received antennas right and r is a rank of the error matrix \mathbf{A} which we are defined previously.

So, based on this to minimize the PEP which is the pair wise error probability, we have to play with the N_r and r the rank right. So, rank has to come into picture. So, please note that the diversity gain is given by so, this is the effectively the SNR. So, diversity gain of $r N_r$ is obtained and the coding gain is given by the product of all the λ_i 's which are the Eigen values of \mathbf{A}^{-1} over r . So, this r is playing a major role in terms of the diversity gain and the coding gain.

So, for the first time, it is very explicit that a space time trellis code is going to give you both a diversity gain and coding gain. So, diversity gain tantamount to the slope of the (Refer Time: 49:50) curve and coding gain is a kind of the left shift of the curve. So, we get both diversity gain and coding gain from space time trellis code. And now with this expression we can formulate the rank determined criteria

(Refer Slide Time: 50:06)

Information Theory, Coding and Cryptography

Rank-Determinant Criteria

- **Rank criterion:** In order to achieve maximum diversity advantage, the matrix $A(\mathbf{C}, \tilde{\mathbf{C}})$ has to be full rank over all possible \mathbf{C} and $\tilde{\mathbf{C}}$.
- **Determinant criterion:** Suppose $A(\mathbf{C}, \tilde{\mathbf{C}})$ is full rank for all \mathbf{C} and $\tilde{\mathbf{C}}$.
- In order to achieve maximum coding advantage, the minimum determinant of $A(\mathbf{C}, \tilde{\mathbf{C}})$ over all possible \mathbf{C} and $\tilde{\mathbf{C}}$ should be maximized.

Indian Institute of Technology,
Delhi

43

Ranjan Bose
 Department of Electrical Engineering

What does the rank criteria say? In order to achieve the maximum diversity advantage, the matrix $A(\mathbf{C}, \tilde{\mathbf{C}})$ has to be full rank because r was a rank. So, we have to find this \mathbf{C} and $\tilde{\mathbf{C}}$. We should design such that we should have this to be full rank and suppose A is full rank, then in order to achieve the maximum coding advantage, the minimum determinant of A over all possible \mathbf{C} and $\tilde{\mathbf{C}}$ should be maximized. So, this is known as the rank determined criteria that was so slow fading channel

(Refer Slide Time: 50:47)

Information Theory, Coding and Cryptography

Fast Rayleigh Fading

- Let $\rho(C, \tilde{C})$ denote the set of time instances $1 \leq t \leq J$ such that $|c_t - \tilde{c}_t| \neq 0$ and n_H be the number of such time instances.
- It can be shown for fast Rayleigh fading that

$$P(C \rightarrow \tilde{C}) \leq \prod_{t \in \rho(C, \tilde{C})} (|c_t - \tilde{c}_t|^2 E_s / 4N_0)^{-N_r}$$

Minimizing maximum PEP leads to following design criteria for fast fading

Indian Institute of Technology, Delhi
44
Ranjan Bose
Department of Electrical Engineering

And if you do a similar analysis for fast fading channels, you get a parallel set of design criteria. So, you can first show that: what is the pair wise error probability for fast Rayleigh fading channel. The expression comes out something like this. So, the rank r term is missing here. Now the minimizing the PEP leads to the following design criteria for fast fading channels.

(Refer Slide Time: 51:13)

Information Theory, Coding and Cryptography

Product-Distance Criteria

- **Distance Criterion:** In order to achieve the diversity $n_d N_r$, we require for any two codewords C and \tilde{C} that the minimum symbol-wise Hamming distance between C and \tilde{C} be at least n_H .
- **Product Criterion:** To obtain maximum coding advantage, maximize the minimum of the products

$$\prod_{t \in \rho(C, \tilde{C})} |c_t - \tilde{c}_t|^2$$

- over all pairs of distinct codewords C and \tilde{C} .

N_r

Indian Institute of Technology, Delhi
45
Ranjan Bose
Department of Electrical Engineering

First we have the distance criteria and then we have the product criteria. Therefore, they are called the product distance criteria for fast fading channel. In order to achieve the

diversity in $n \log N$ we require any two code words of C and C tilde such that the minimum symbol wise hamming distance between C and C tilde must be at least $n H$.

So, this will give you the diversity of $n \log N$; that is number 1 and the product criteria again coming from the PEP expression is to obtain the maximum coding advantage, we have to maximize the minimum of the products as follows. So, this can be done by simple search algorithms to find out good mapping. Finally, it comes down to the mapping. How do we assign the labels, the edge labels of the trellis? So, these this is the product distance criteria, with that we come to the end.

(Refer Slide Time: 52:10)

Information Theory, Coding and Cryptography

Summary

- Computation of d_{free}
- TCM for fading channels
- Space Time Trellis Codes
- Slow Rayleigh Fading Scenario
- Fast Rayleigh Fading Scenario
- Examples

Indian Institute of Technology, Delhi 46 Ranjan Bose
Department of Electrical Engineering

So, let us summarize. We have looked at how to compute d_{free} then we went to study TCM for fading channels. Then logically we went into space time codes, and looked at 2 scenarios slow, and fast Rayleigh fading channels, and we looked at a couple of examples.

With that we come to the end of this lecture.