

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

Module - 31
Trellis Coded Modulation
Lecture - 31

Hello and welcome to our next lecture on Trellis Coded Modulation. Let us start with a brief outline.

(Refer Slide Time: 00:36)

Information Theory, Coding and Cryptography

Outline

- Ungerboeck's design rules
- Performance Evaluation
- Examples

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

So, we would primarily cover Ungerboeck's design rules today and understand how we can design very efficient trellis with good distance properties, and then we would spend some time looking at performance evaluation of TCM schemes over additive white Gaussian noise channels. Of course, we will look at some examples along the way.

(Refer Slide Time: 00:38)

Information Theory, Coding and Cryptography

Recap

- Combined Coding and Modulation
- Trellis Coded Modulation
- Free distance

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

Let us quickly see what we have done so far. We have understood what do we mean by coding and modulation and combining them then we introduced the notion of trellis coded modulation and we introduced the idea of free distance. We will soon see that free distance d_{free} will be the single most important design parameter for trellis coded modulation schemes.

(Refer Slide Time: 01:21)

Information Theory, Coding and Cryptography

Trade-off: Bandwidth and Performance

- In error control coding techniques, extra bits are added to the information bits in a **known** manner.
- However, the improvement in the bit error rate is obtained at the **expense of bandwidth**, because of these extra bits.
- This bandwidth expansion is equal to the reciprocal of the code rate.
- For example, a RS (255,223) code has a code rate, $R = 223/255 = 0.8745$ and $1/R = 1.1435$.

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

So, if you remember in error control coding, we introduced extra bits in a known manner to be recovered at the receiver end in order to come back from the errors.

Now, this addition of extra bits came at a cost of additional bandwidth. Therefore, error control schemes always required more bandwidth and we realized that this was inversely proportional to the code rate R.

(Refer Slide Time: 01:52)


Information Theory, Coding and Cryptography

Coding Gain

- The difference between the values of the SNR for the coded and uncoded schemes required to achieve the same error probability is defined as the **coding gain**, g .

$$g = SNR|_{\text{uncoded}} - SNR|_{\text{coded}}$$
- At high SNR, the coding gain can be expressed as

$$g_{\infty} = g|_{SNR \rightarrow \infty} = 10 \log \frac{(d_{free}^2 / E_s)_{\text{coded}}}{(d_{free}^2 / E_s)_{\text{uncoded}}}$$
- where g_{∞} represents the **asymptotic coding gain** and E_s is the average signal energy.
For uncoded schemes, d_{free} is simply the minimum Euclidean distance between the signal points.

 Indian Institute of Technology, Delhi
5
Ranjan Bose
Department of Electrical Engineering

And, what we decided that in trellis coded modulation scheme we can gain something out of nothing because we can leverage the gain by the error control coding scheme and the modulation scheme together.

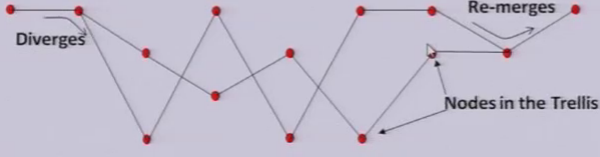
So, we defined for ourselves the coding gain where we found out that at a high SNR the coding gain asymptotic coding gain as SNR tends to infinity is defined as 10 log to the base 10 d_{free}^2 / E_s normalized coded scheme versus uncoded scheme. For the uncoded scheme d_{free} is simply the minimum Euclidean distance between the signal points. So, this is this gain that we get because of the trellis d_{free} that as to the advantage and please remember that we have now the trellis labeled by symbols rather than by the bits as in the convolutional encoder case.

(Refer Slide Time: 02:54)


Information Theory, Coding and Cryptography

Maximizing free distance

- Another approach is to assign the symbols to the branches in the trellis in a heuristic manner so as to increase the d_{free} .
- We know that an **error event** consists of a path diverging in one state and then merging back after one or more transitions
- The Euclidean distance associated with such an error event can be expressed as
$$d_{total}^2 = d_E^2 \text{ (diverging pair of paths)} + \dots + d_E^2 \text{ (re-merging pair of paths)}.$$



Nodes in the Trellis

 Indian Institute of Technology, Delhi

6

Ranjan Bose
Department of Electrical Engineering

So, we looked at how to maximize this free distance while designing good trellis coded modulation schemes for that we designed we decided that the error event is when we diverge like this and then we go through the trellis for a couple of hops, and then finally we merge back. So, this constitutes an error event. What happens is we transmit a sequence of bits and it corresponds to a path in the trellis for example, this path and at the decoding end suppose we are using Viterbi we decode another path in the trellis and therefore, this constitutes an error event.

Now, in order to maximize the d_{free} we would like to make sure that the two most closely resembling paths are separated maximally, for that we do not know what happens in between, but at least the diverging and the merging back paths need to be maximally apart in terms of the Euclidean distance. Here we look at the squared Euclidean distance because we take the total of all the branch labels.


(Refer Slide Time: 04:17)

Information Theory, Coding and Cryptography

Set Partitioning

- The mapping by set partitioning is based on successive partitioning of the expanded 2^{m+1} ary signal set into subsets with **increasing minimum Euclidean distances**.
- Each time we partition the set, we reduce the number of the signal points in the subset, but increase the **minimum** distance between the signal points in the subset.

7

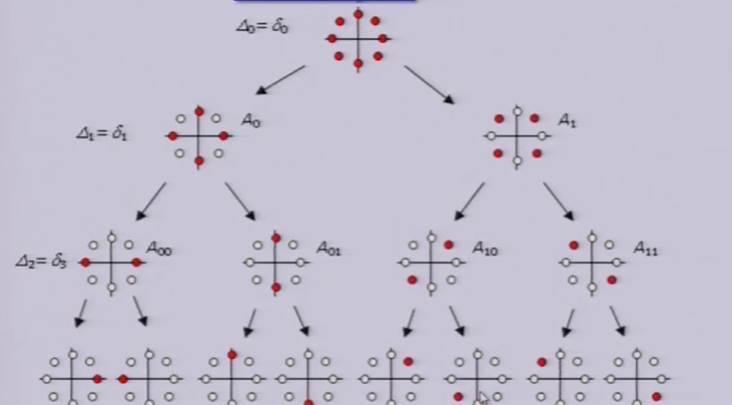
 Indian Institute of Technology, DelhiRanjan Bose
Department of Electrical Engineering

So, we introduce the concept of set partitioning where we consecutively partition a set of the constellation diagram into increasingly minimum Euclidean distances. And what we want to do is to associate the different symbols at different stages of the set partition tree with the branches and label the branches accordingly.


(Refer Slide Time: 04:42)

Information Theory, Coding and Cryptography

Example



8

 Indian Institute of Technology, DelhiRanjan Bose
Department of Electrical Engineering

So, this is a simple example of 8-PSK and how we do set partitioning. So, step one we get into two subsets each one has a larger Euclidean distance and then we continue this


further till we get to the maximally separated points. So, this is an example of a set partitioning of 8-PSK signal set.

(Refer Slide Time: 05:08)

Information Theory, Coding and Cryptography

Ungerboeck's Design Rules

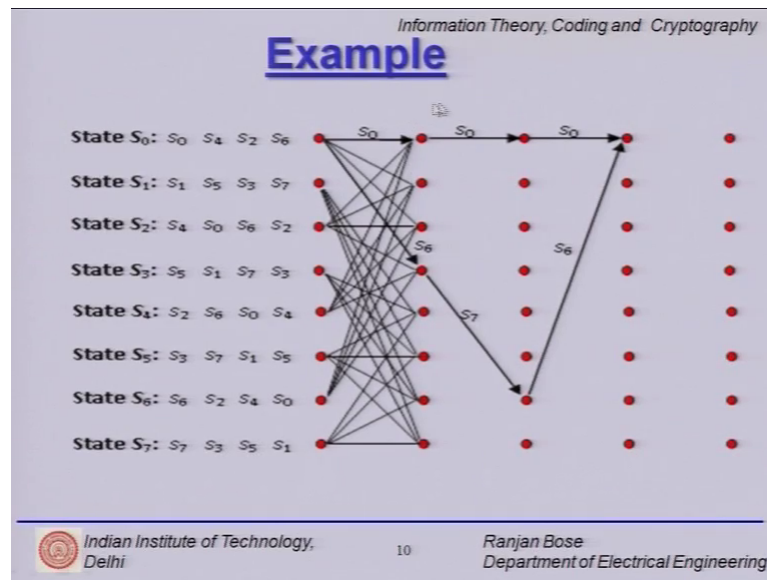
- In 1982 Ungerboeck proposed a set of design rules for maximizing the free Euclidean distance for TCM schemes.
- These design rules are based on heuristics.
- **Rule 1:** Parallel transitions, if present, must be associated with the signals of the subsets in the lowest layer of the set partitioning tree.
These signals have the minimum Euclidean distance
- **Rule 2:** The transitions originating from or merging into one state must be associated with signals of the first step of set partitioning.
- The Euclidean distance between these signals is at least
- **Rule 3:** All signals are used with equal frequency in the trellis diagram.

 Indian Institute of Technology,
Delhi9Ranjan Bose
Department of Electrical Engineering

Now, Ungerboeck's design rule tells us in a heuristic manner how to assign symbols to a trellis. This is rule number 1, where the parallel transitions if present must be associated with the signals in the subset of the lowest layer of the set partitioning tree, which means that if you were to have parallel transitions in your trellis then this is the lowest set and if you have parallel transitions then these opposite symbols s_0 and s_4 for example, should be assigned or s_2 and s_6 must be assigned to the parallel transitions and so and so forth.

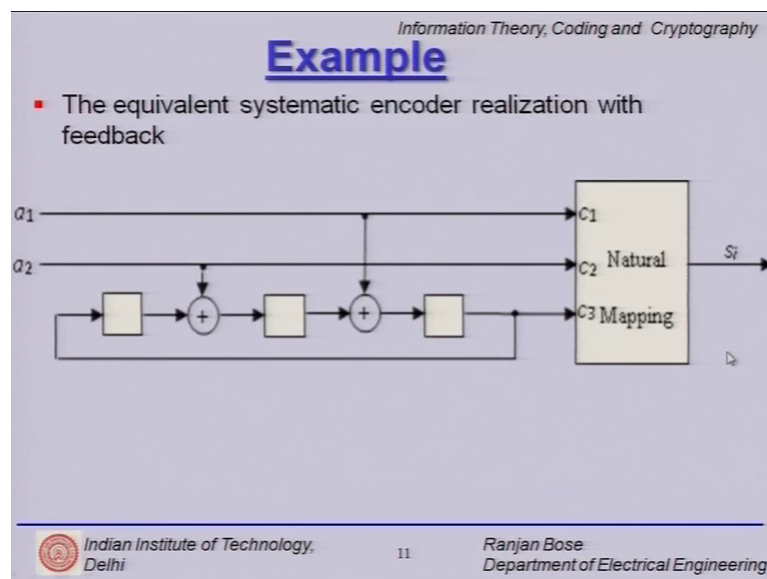
Now, rule number 2 says that the transitions originating from or merging into one-state must be associated with the signals of the first step of the. So, go one step higher first the most damaging ones are the parallel transitions. So, the symbols should be such that they are maximally apart. Then if we do not have parallel transitions then we look at the merging and diverging paths where they should be assigned to the next higher level. And finally, what we must do is try to ensure that all signals are used with equal frequency. This is to our advantage otherwise if we use certain signals too many times then you end up reducing the distance.

(Refer Slide Time: 06:33)



This is an example we did in the last class and how we have been able to assign the diverging paths, and the merging back paths from the sets of symbols which are maximally apart.

(Refer Slide Time: 06:52)



So, let us look at this example again it is best understood by this example. So, suppose we are looking at a rate 2 by 3 convolutional encoder followed by a natural mapping. If you remember natural mapping means 000 is mapped to symbol s_0 , 001 is mapped to symbol s_1 and so on and so forth.

So, clearly there are 3 bits at the output of this convolutional encoder and therefore, we must have 8-PSK in order to convert it into a symbol S_i . So, 2 bits come in goes through the convolutional encoder 3 bits come out they are mapped using the natural mapping and one symbol comes out. So, the equivalent trellis will have 8 states because there are 3 memory elements and there will be 2 paths emanating sorry 4 paths because they are 2 bits which are input. So, corresponding to 00, 011, 0 and 1 1 you will have 4 branches coming out and each branch will be labeled by a symbol.

(Refer Slide Time: 07:59)

Information Theory, Coding and Cryptography

Example


- Let us represent the output of the convolutional encoder in terms of the input and the delayed versions of the input.
-

$$c_1(D) = a_1(D),$$

$$c_2(D) = a_2(D),$$

$$c_3(D) = \left(\frac{D^2}{1+D^3}\right)a_1(D) + \left(\frac{D}{1+D^3}\right)a_2(D).$$

© 2007

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

So, the other way to visualize this is in terms of the delay versions. So, if you look at just the first C_1 it is directly a 1 and C_2 is directly a 2, but C_3 is a 1 with one delay and then a 2 goes through two delays and the sum goes through three delays. So, if you solve this then you can label C_1 as a 1, C_2 as a 2, because they were directly connected, but C_3 if you solve those intermediate equations can be represented in this delayed version.

(Refer Slide Time: 08:50)

Information Theory, Coding and Cryptography


Example

- Therefore, the generator polynomial matrix of this encoder is

$$G(D) = \begin{bmatrix} 1 & 0 & \frac{D^2}{1+D^3} \\ 0 & 1 & \frac{D}{1+D^3} \end{bmatrix}$$

and the parity check polynomial matrix, $H(D)$, satisfying $G(D) \cdot H^T(D) = \mathbf{0}$ is

- $H(D) = [D^2 \quad D \quad 1+D^3].$

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

So, we have an equivalent generator polynomial matrix for this encoder which we studied earlier in convolutional encoder as follows. So, you have a simple representation of this generator polynomial matrix, ok. This unity this one and one here and identity matrix in the beginning shows that it is kind of a systematic encoder.


So, clearly if we have $G(D)$ we can write out the $H(D)$ matrix the poly the parity check polynomial matrix $H(D)$ such that $G(D) \cdot H^T(D)$ should be equal to the 0 matrix. So, you can quickly make an observation and write as follows. So, if you multiply G with H^T then this D^2 is specter by this one D specter by this one so, numerator becomes $D^2 + D$ divided by $1 + D^3$ and you quickly realize that you have $G(D) \cdot H^T(D) = 0$. So, it is easy to make $H(D)$ from $G(D)$ and $G(D)$ is equally easy to realize from the visual observation.

(Refer Slide Time: 10:13)

Information Theory, Coding and Cryptography

Example

- We can re-write the parity check polynomial matrix
- $H(D) = [H_1(D) \ H_2(D) \ H_3(D)]$, where
 - $H_1(D) = D^2 = (000 \ 100)_{\text{binary}} = (04)_{\text{octal}}$
 - $H_2(D) = D = (000 \ 010)_{\text{binary}} = (02)_{\text{octal}}$
 - $H_3(D) = 1 + D^3 = (001 \ 001)_{\text{binary}} = (11)_{\text{octal}}$
- Its possible to make a Table that gives the encoder realization and asymptotic coding gains of some of the good TCM codes constructed for the 8-PSK signal constellation.
- Almost all of these TCM schemes have been found by exhaustive computer searches.
- The coding gain is given with respect to the uncoded QPSK.
- The parity check polynomials are expressed in the octal form.

 Indian Institute of Technology,
Delhi14Ranjan Bose
Department of Electrical Engineering

So, now you have this $H(D)$ matrix in terms of the $H_1(D)$, $H_2(D)$ and $H_3(D)$ and you can write it in terms of binary or octal as follows. So, D^2 . So, this stands for the coefficients of D raised to the power 0, if it is missing. So, it is 0 this is the coefficient for D raised to the power 1, which is missing. So, it is 0 and this is the coefficient for D raised to the power 2 which is present. So, it is 1. So, that is the binary representation and this is the octal representation.

Similarly, $H_2(D)$ it is the coefficient for D is there, coefficient for D raised to the power 0 is missing, coefficient is for D^2 is missing and therefore, it is 010 and binary and 02 in octal. The first three bits stands for the first digit secondary bits stands for the second digit and $H_3(D)$ is 1 plus D^3 if you see 1 plus D^3 . So, you have 1 here and this is the coefficients of D^3 ; so 11 in octal. So, I can represent this simply using the octal notations which is found in the literature, ok.

So, it is now possible to form a table which goes the encoder realization and asymptotic gains of some good TCM codes usually constructed from binary searches because there is no hard and fast design rules. So, we can always come up with a table, they have been found by exhaustive computer searches.


(Refer Slide Time: 11:58)

Information Theory, Coding and Cryptography

Some Good TCM schemes

TCM schemes using 8-PSK

No. of states	H ₁	H ₂	H ₃	d_{free}^2 / E_s	(dB)
4	-	2	5	4.00	3.01
8	04	02	11	4.58	3.60
16	16	04	23	5.17	4.13
32	34	16	45	5.75	4.59
64	066	030	103	6.34	5.01
128	122	054	277	6.58	5.17
256	130	072	435	7.51	5.75

 Indian Institute of Technology, Delhi
15
Ranjan Bose
Department of Electrical Engineering

So, we can write them out and document their gain asymptotic gain with respect to QPSK as follows. Here is a table of good TCM schemes of which we have already looked at this number of state 8 H 1 – 4, 2 and 11 in octal notation, where we calculated that this normalized free distance squared is 4.58 leading to a gain of 3.6 dB asymptotic coding gain, but that is not the only one we have, so many other possibilities and by now we understand how to write the encoder using the octal notation as follows. So, this is a set of good TCM schemes using 8-PSK.


(Refer Slide Time: 12:45)

Information Theory, Coding and Cryptography

Example

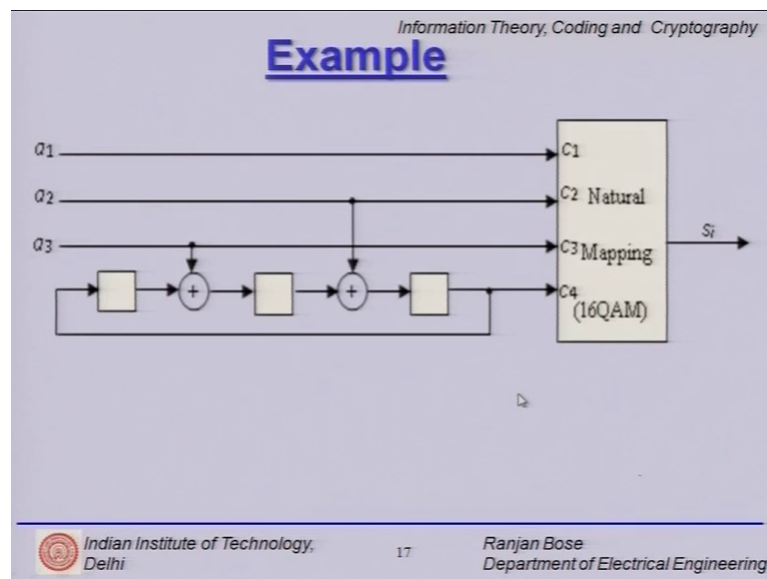
- We now look at a TCM scheme that involves 16QAM.
- The TCM encoder takes in 3 bits and outputs one symbol from the 16QAM constellation diagram.
- This TCM scheme has a throughput of 3 bits/s/Hz and we will compare it with uncoded 8-PSK, which also has a throughput of 3 bits/s/Hz.
- Let the minimum distance between two points in the signal constellation of 16QAM be d_0 as depicted in Fig. 7.12.
- It is assumed that all the signals are equiprobable.
- Then the average signal energy of a 16QAM signal is obtained as

$$E_s = \frac{1}{16} (2\delta_0^2 + 10\delta_0^2 + 10\delta_0^2 + 18\delta_0^2) = \frac{10}{4}\delta_0^2$$

 Indian Institute of Technology, Delhi
16
Ranjan Bose
Department of Electrical Engineering

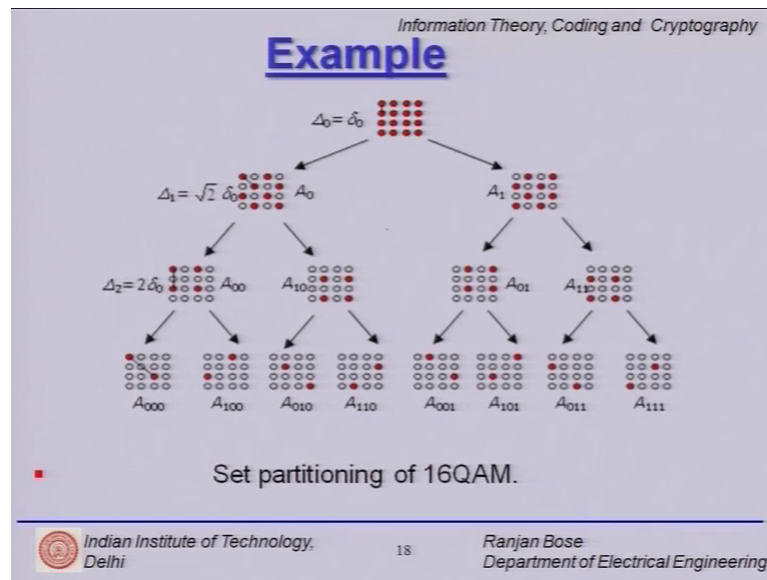
So, now let us look at another example using 16 QAM and this encoder takes in 3 bits and adds one more bits to gives it is a 4 bit output. So, the convolution encoder is a 3 by 4 consequently we cannot use 8-PSK we must use 16 QAM and what we do is that you can first find out what is the average signal energy for 16 QAM. If you remember all the symbols do not carry equal energy because they are not equally spaced from the center or the origin.

(Refer Slide Time: 13:25)



So, this is the example of the rate 3 by 4 convolutional encoder. It takes 3 bits and gives out 4 bits followed by a natural mapping, but this time I need to have a 16 QAM and consequently we get mapped symbol S_i out. Again, there are 3 memory elements leading to 8 states in the trellis. So, rate 3 by 4 convolution encoder natural mapping eight-state trellis.

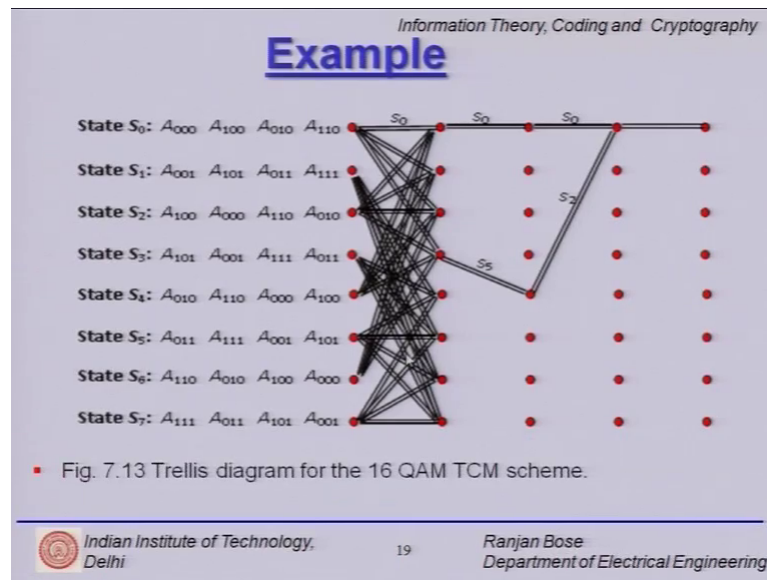
(Refer Slide Time: 13:57)



So, first we have to do set partitioning, but this time we start with the 16 QAM and if you see as a first step I divided into two subsets each one. We, throughout the alternate symbols leading to an increased Euclidean spacing between the symbols as follows and then we do not stop we continue further. So, we get even larger distances between the neighboring symbols. So, first two subsets and then two more for each so, four, total number of subsets and then eight subsets we have continuously increase the distance. So, this is the example for set partitioning of 16 QAM.

Now, we will use Ungerboeck's rule to assign the symbols from the correct level of the set partitioning tree to the diverging paths and the merging back paths.

(Refer Slide Time: 14:58)



So, this is an example of that convolutional encoder if you see there are parallel transitions here so. In fact, every node has eight outgoing branches of which there are four pairs of parallel transitions. So, we must apply Ungerboeck's rule for this as you can see there were three memory elements leading to eight states in the trellis, right and there were three input bits coming in leading to eight branches emanating from each node from 000 up to 111.

Here is assignment. So, here is the assignment of the symbols from the set partitioning tree that we just now saw. This is a tree and this is what we mean by A 0 and A 1 as the subsets this is A 00, A 10, A 01 and A 11. Similarly, these eight subsets A 000, A 100 up to A 111. So, we are going to assign the parallel transition from this lowest rung in the ladder,

So, these two must be assigned to the first parallel transition, these two symbols must be assigned to the second and then subsequently we can look at the different transitions. So, A 000 so, the two symbols from A 000 are assigned here A 100 are assigned to the next set of parallel transitions. Two symbols from A 010 to this one. And finally, two symbols from A 110 to this one alright, but we have also made sure that the diverging branches must be such that they get assigned from the next higher level.

So, the diverging branch should be such that the symbols are assigned from these parts. So, that is the case because there more than one ways to assign these two symbols to the

parallel branches. So, similarly the merging back branches merging back branches must be assigned to the next higher level. So, this is an example how Ungerboeck's design rule is used to assign symbols to the trellis with parallel transitions.

(Refer Slide Time: 17:38)


Information Theory, Coding and Cryptography

Example

- We have,

$$\delta_0 = 2\sqrt{\frac{E_s}{10}}$$

- The trellis has 8 states.
- Each node has 8 branches emanating from it because the encoder takes in 3 input bits at a time ($2^3 = 8$).
- The Ungerboeck design rules are followed to assign the symbols to the different branches.
- The branches diverging from a node and the branches merging back to a node are assigned symbols from the set A_0 and A_1 .
- The parallel paths are assigned symbols from the lowest layer of the set partitioning tree ($A_{000}, A_{001}, \text{etc.}$).

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


So, if you calculate now the d_{free} based on this then you can find out what is asymptotic coding gain.

(Refer Slide Time: 17:51)

Information Theory, Coding and Cryptography

Example

- The squared Euclidean distance between any two parallel paths is $d_p^2 = 8\delta_0^2$
- This is by design as we have assigned symbols to the parallel paths from the lowest layer of the set partitioning tree.
- The minimum squared Euclidean distance between non-parallel paths is $d_E^2 = d_1^2 + d_0^2 + d_1^2 = 5\delta_0^2$
- Therefore, the free Euclidean distance for the TCM scheme is $d_{free}^2 = \min [8\delta_0^2, 5\delta_0^2] = 5\delta_0^2 = 2E_s$.

 Indian Institute of Technology, Delhi
21
Ranjan Bose
Department of Electrical Engineering

This if you do the calculations about the squared Euclidean distance then the minimum squared Euclidean distance between non parallel paths is $5\delta_0^2$ and you

have if you look at the squared Euclidean distance between two parallel transitions is $8 \Delta^2$. So, we have been able to ensure that it is not the parallel transitions that are leading to d_{free}^2 .

So, if you look at d_{free}^2 it is the minimum of the diverging and the merging back branches. So, we calculated d_{free}^2 between this node and this node either through this branch or it diverges and then merges back. So, it is not the parallel transition, but the others which are causing it to have an error event and consequently we have the d_{free}^2 dictated by the non parallel path which is $5 \Delta^2 = 2 \sqrt{2} E_s$.

(Refer Slide Time: 18:56)

Information Theory, Coding and Cryptography

Example

- Note that the free Euclidean distance is determined by the non-parallel paths rather than the parallel paths.
- We now compare the TCM scheme with the uncoded 8-PSK, which has the same throughput.
- For uncoded 8-PSK, the minimum squared Euclidean distance is $2 - \sqrt{2} E_s$.
- Thus, the asymptotic coding gain for this TCM encoder is

$$g_c = 10 \log \frac{2}{2 - \sqrt{2}} = 5.3 \text{ dB.}$$

Indian Institute of Technology,
Delhi

22

Ranjan Bose
Department of Electrical Engineering

So, now that we have the minimum Euclidean distance corresponding to the d_{free} for the TCM scheme we want to find out the asymptotic coding gain. So, if we had not used rate 3 by 4 encoder, if we had this 3 bits coming and we had to transmit them by modulating in them we would use 8-PSK. So, we find out the squared Euclidean distance from 8-PSK which is $2 - \sqrt{2} E_s$ and now, this asymptotic coding gain is simply the ratio $\frac{2}{2 - \sqrt{2}}$ and this is log to the base 10, 10 times. So, it gives you a whopping 5.3 dB coding gain.

So, that trellis the complicated looking trellis also gave us a pretty good asymptotic coding gain. If you remember in electrical engineering even a coding gain of 3 dB is worth looking into. Now, here we have 5 point 3 dB asymptotic coding gain. So, this is


really a very good design, it is a good example of a rate 3 by 4 convolutional encoder coupled with the natural mapper.

(Refer Slide Time: 20:12)

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* 23 *Ranjan Bose
Department of Electrical Engineering*

Now, let us quickly spend some time looking at the decoding part. So, we have already made this observation that there is a one to one correspondence between the bit stream coming in and a path in the trellis. Only in this case the trellis paths are labeled by symbols and not by bits. So, the decoding problem is finding the most likely path of the trellis with respect to that which is received and most likely comes in terms of the minimum Euclidean distance.


So, we use the maximum likelihood criteria to do it and Viterbi algorithm is commonly used for this decoding technique.

(Refer Slide Time: 20:54)

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* 24 *Ranjan Bose
Department of Electrical Engineering*

So, for soft decision decoding of the received sequence is using Viterbi algorithm each trellis branch is labeled by the branch metric based on the observed received sequence. The only difference is in convolution decoding we used hard we used the bits and therefore, the hamming distance here we will use the Euclidean distance.


So, using the maximum likelihood decoder for the additive white Gaussian noise channels the branch metric is defined as the Euclidean distance, and the Viterbi decoder tries to find out a path in the trellis which is most closely resembling that is closest in terms of the Euclidean distance with respect to the received path.

(Refer Slide Time: 21:43)

Information Theory, Coding and Cryptography

Definition

- The **branch metric** for a TCM scheme designed for AWGN channel is the *Euclidean distance* between the received signal and the signal associated with the corresponding branch in the trellis.

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


So, a branch metric for a TCM scheme is now in terms of the Euclidean distance and we now try to find out the performance of TCM schemes in additive white Gaussian noise channel.

(Refer Slide Time: 21:56)

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, Delhi 26 Ranjan Bose
Department of Electrical Engineering

So, in the next few slides we will develop a mathematical tool. It will be slightly involved, but what we will do is try to get the message across in terms of what we are trying to do. So, we are now going to work with this free Euclidean distance.

Now, what we would define is the average number of nearest neighbors N as a function of d_{free} gives the average number of paths in the trellis with the free Euclidean distance d_{free} . So, d_{free} is actually the weakest link in the chain and we would like to find out how many weak links are there in the chain. So, this N as a function of d_{free} is used in conjunction with d_{free} to evaluate how good a trellis coded modulation scheme is.

(Refer Slide Time: 22:50)

Information Theory, Coding and Cryptography

Error Event

- Two finite length paths in the trellis form an **error event** if they start from the same state, diverge and then later s_n and \hat{s}_n merge back.
- An error event of length l is defined by two coded sequences


$$s_n = (s_n, s_{n+1}, \dots, s_{n+l+1})$$

$$\hat{s}_n = (\hat{s}_n, \hat{s}_{n+1}, \dots, \hat{s}_{n+l+1}),$$
- Such that

$$s_n = \hat{s}_n$$

$$s_{n+l+1} = \hat{s}_{n+l+1}$$

$$s_i \neq \hat{s}_i, i = n+1, \dots, n+l.$$


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

So, let us look at an error event definition. So, we have already defined that we have a sequence S_n , this is a vector of S_n, S_{n+1} and so on so forth, but we have l branches. So, S_{n+1} and we have a received sequence or the estimate that we try to get in terms of the guessed sequence that we sent, right. And we would like to get S_n equal to \hat{S}_n .

(Refer Slide Time: 23:28)

Information Theory, Coding and Cryptography

Definition

- The probability of an error event starting at time n , given that the decoder has estimated the correct transmitter state at that time is called the **error event probability**, P_e .

28

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

So, the probability of an error event starting at time n , given that the decoder has estimated the correct transmitter state. And the time is called the error event probability P_e and we will try to get an upper bound on this P_e .

(Refer Slide Time: 23:44)

Information Theory, Coding and Cryptography

Performance Evalⁿ over AWGN Ch.

- The performance of TCM schemes is generally evaluated by means of upper bounds on error event probability.
- It is based on the generating function approach.
- Let us consider again the Ungerboeck model for rate $m/(m+1)$ TCM scheme
- The encoder takes in m bits at a time and encodes it to $m+1$ bits, which are then mapped by a memoryless mapper, $f(\cdot)$, on to a symbol s_i .
- Let us call the binary $(m+1)$ -tuples \mathbf{c}_i as the labels of the signals s_i .
- We observe that there is a one-to-one correspondence between a symbol and its label.
- Hence, an error event of length l can be equivalently described by two sequences of labels

$$\mathbf{C}_l = (c_k, c_{k+1}, \dots, c_{k+l-1}) \text{ and } \mathbf{C}'_l = (c'_k, c'_{k+1}, \dots, c'_{k+l-1})$$

- where, $c'_k = c_k \oplus e_k, c'_{k+1} = c_{k+1} \oplus e_{k+1}, \dots$, and $\mathbf{E}_l = (e_k, e_{k+1}, \dots, e_{k+l-1})$ is a sequence of binary error vectors.

29

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

So, the performance of a TCM scheme is generally evaluated by means of the upper bound on the error event probability. We will soon realize why an upper bound is used simply because it is impossible to keep track of all the possible path that may result in an

error event. So, we look at the whole group of possible paths and we come up with and some kind of a union bound.

Now, the performance evaluation is based on the generating function approach that we have already studied in our earlier portion of trellis coded modulation and the convolutional encoder. So, here please note we are going to look at m over m plus one TCM scheme in general and this will be coupled with a mapper. So, encoder takes m bits and converts into m plus 1 bits and what we have is we have a binary m plus 1 tuple c_i as the label for signal s_i . So, we have these labels that we are going to work within the next few slides and the idea is to recover these labels and there is of course, a one to one correspondence with this c_i to s_i . We observe s_i and then we try to recover the c_i .

So, an error event of length l can be equivalently described by two sequences of labels C_l and C_l' , ok. So, we are going to work with this and error event is when they are not the same, ok. So, how do we describe it? Well, if they are not the same then an error has happened and we have this c_k binary addition e_k where this is the error event. Again, this is a binary label $c_k' = c_k \oplus e_k$ this is nothing but the original one that was transmitted plus an error binary event.

So, just like that we have a sequence of c_k, c_{k+1} and so and so forth we have a sequence of this error binary error vectors, ok.

(Refer Slide Time: 26:07)


Information Theory, Coding and Cryptography

Performance Evalⁿ over AWGN Ch.

- The mathematical symbol \oplus represents binary (modulo-2) addition.
- An error event of length l occurs when the decoder chooses, instead of the transmitted sequence C_l , the sequence C_l' which corresponds to a path in the trellis diagram that diverges from the original transmitted path and re-merges back exactly after l time intervals.
- To find the probability of error we need to sum over all possible values of l the probabilities of error events of length l (i.e., joint probabilities that C_l is transmitted and C_l' is detected).
- The upper bound on the probability of error is obtained by the following union bound

$$P_e \leq \sum_{l=1}^{\infty} \sum_{s_l} \sum_{s_l' \neq s_l} P(s_l) P(s_l, s_l')$$

- where $P(s_l, s_l')$ denotes the pairwise error probability (i.e., probability that the sequence s_l is transmitted and the sequence s_l' is detected).

 Indian Institute of Technology, Delhi
30
Ranjan Bose
Department of Electrical Engineering

So, basically in the performance evaluation over additive white Gaussian noise channel, we look at the upper bound on the probability of error simply by looking at this union bound, the probability of s_l and the pair wise error probability $P_{s_l, s_l'}$ where s_l' is not equal to s_l . So, we look at all such cases therefore, it is summation over all the cases where s_l' is not equal to s_l .

And, then we have all the possible symbols, right. So, we have a summation over all the possible symbols and then we have all possible path lengths of the error event. So, error event can be of one length. So, in one hop it diverges and merges back or two hops or three hops up to infinity, because the trellis is a semi infinite geometric structure. So, we have these three summations here. And therefore, we have a union bound on all the cases, where s_l is not the same as s_l' .

(Refer Slide Time: 27:21)

Information Theory, Coding and Cryptography

Performance Evalⁿ


- Assuming a one-to-one correspondence between a symbol and its label, we can write

$$P_e \leq \sum_{l=1}^{\infty} \sum_{C_l} \sum_{C_l' \neq C_l} P(C_l) P(C_l, C_l')$$

$$= \sum_{l=1}^{\infty} \sum_{C_l} \sum_{E_l \neq 0} P(C_l) P(C_l, C_l \oplus E_l)$$
- The pairwise error probability $P_2(C_l, C_l \oplus E_l)$ can be upper-bounded by the Bhattacharyya bound as follows

$$P_2(C_l, C_l \oplus E_l) \leq e^{-\left\{ \frac{1}{4N_0} |f(C_l) - f(C_l')|^2 \right\}}$$

$$= e^{-\left\{ \frac{1}{4N_0} \sum_{i=1}^l |f(C_{l,i}) - f(C_{l,i}')|^2 \right\}}$$
- where $f(\cdot)$ is the memoryless mapper.

 Indian Institute of Technology,
Delhi

31

Ranjan Bose
Department of Electrical Engineering

So, now we can do a quick set of algebraic steps and, we look at this upper bound. Now, since there is a one to one correspondence between a symbol and its label we have replaced the symbols by these labels. So, we have now the sequence of labels and the probability that C_l is not the same as C_l' . So, we do over all those possible cases and we have all the possible symbols and we have the C_l labels and again all the possible length of the path, but we have just now put together that this C_l' is nothing but the original C_l plus the error. So, it is simply written as follows.

So, now we can use the Bhattacharyya bound to limit upper bound this error event. What is this pairwise error event? Probability between C_1 being transmitted and C_1 plus some error being received is now less than this is a function of C_1 minus C_1 prime. It is the Euclidean distance square, where f is a memoryless mapper. So, this is how we are using the Bhattacharyya bound to get into this Euclidean distance concept.

(Refer Slide Time: 27:28)


Information Theory, Coding and Cryptography

Performance Evalⁿ

- Let $D = e^{-\frac{1}{4N_0}}$ (for additive white Gaussian noise channel with single sided power spectral density N_0), then

$$P_2(C_1, C_1 \oplus E_1) \leq D^{\|f(C_1) - f(C_1 \oplus E_1)\|^2} = D^{d_E^2(f(C_1), f(C_1 \oplus E_1))}$$
- where $d_E^2(f(C_1), f(C_1 \oplus E_1))$ represents the squared Euclidean distance between the symbol sequences
- Next, define the function

$$W(E_1) = \sum_{C_1} P(C_1) D^{\|f(C_1) - f(C_1 \oplus E_1)\|^2}$$

 Indian Institute of Technology,
Delhi

32

Ranjan Bose
Department of Electrical Engineering

So, now define capital D as e raised to the power minus 1 over 4 N naught. So, this will help us write this in a compact form. So, now, this pairwise error probability is upper bounded by this D raised to power this norm of $f C_1$ minus $f C_1$ prime, and this is nothing but the squared Euclidean distance between these two.

So, we now define this n function W of E_1 . So, please note it is a function only of E_1 and this is great, because we do not want this W to be a function of any particular path in the trellis it is only a function of the error event. Therefore, we are defining it as this. This is simply defined as follows and now we will plug in this W into the term for pairwise error probability.

(Refer Slide Time: 29:46)


Information Theory, Coding and Cryptography

Performance Evalⁿ

- We can now write the probability of error as

$$P_e \leq \sum_{l=1}^{\infty} \sum_{E_l \neq 0} W(E_l)$$
- From the above equation we observe that the probability of error is upper-bounded by a sum over all possible error events, (for additive white Gaussian noise channel with single sided power spectral density N_0), then E_l .
- Note that

$$d_E^2(f(\mathbf{C}_l), f(\mathbf{C}_l \oplus \mathbf{E}_l)) = \sum_{i=1}^l d_E^2(f(c_i), f(c_i \oplus e_i)).$$
- We now introduce the concept of an error state diagram which is essentially a graph whose branches have matrix labels.
- We assume that the source symbols are equally probable with probabilities $2^{-m} = 1/M$.

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

So, now pairwise error probability is now simply for all possible path lengths from 1 to infinity and error not being equal to 0. Earlier it was at C l is not equal to C l prime and for all this cases where error is not equal to 0 it looks at all the possible W's here, fine.

So, we now look at finally, how to get a handle on this upper bound on the right hand side.

(Refer Slide Time: 30:22)


Information Theory, Coding and Cryptography

Performance Evalⁿ

- The error weight matrix, $G(e_i)$ is an $N \times N$ matrix whose element in the p^{th} row and q^{th} column is defined as

$$[G(e_i)]_{p,q} = \frac{1}{M} \sum_{c_{p \rightarrow q}} D^{|f(c_{p \rightarrow q}) - f(c_{p \rightarrow q} \oplus e_i)|^2}$$
- if there is a transition from state p to state q , and

$$[G(e_i)]_{p,q} = 0,$$
- If there is no transition from state p to state q of the trellis, $c_{p \rightarrow q}$ where are the label vectors generated by the transition from state p to state q .

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

So, we now have this error weight matrix $G(e_i)$ is an N cross N matrix whose element in the p -th row and q -th column is defined as follows, ok. So, we define this error weight

matrix and see how clearly we are defining it we have this D which we have defined earlier, this is the mapper of f to q , right and minus c to q plus this error e . So, if you do this is talking about from transition from state p to q .

So, we are now looking at all the possible transitions. So, we are looking at all the possible transitions and how the error is incorporated when we go from one transition to other, right.


(Refer Slide Time: 31:25)

Information Theory, Coding and Cryptography

Performance Evalⁿ

- The summation accounts for the possible parallel transitions (parallel paths) between states in the trellis diagram.
- The entry (p, q) in the matrix \mathbf{G} provides an upperbound on the probability that an error event occurs starting from the node p and ending on q .
- Similarly, $(1/N)\mathbf{G}\mathbf{1}$ is a vector whose p^{th} entry is a bound on the probability of any error event starting from node p .
- Now, to any sequence $E_l = e_1, e_2, \dots, e_l$, there corresponds a sequence of l error weight matrices $\mathbf{G}(e_1), \mathbf{G}(e_2), \dots, \mathbf{G}(e_l)$.
- Thus we have

$$W(E_l) = \frac{1}{N} \mathbf{1}^T \prod_{n=1}^l \mathbf{G}(e_n) \mathbf{1}$$

 Indian Institute of Technology, Delhi
35
Ranjan Bose
Department of Electrical Engineering

So, if you do this math completely you can see that this $W(E_l)$ which is only a function of the error can be written as 1 over N then $\mathbf{1}$ this is a vector transpose product of $\mathbf{G}(e_n) \mathbf{1}$. So, this is what we can do some basic mathematics to come to this one.


(Refer Slide Time: 31:49)

Information Theory, Coding and Cryptography

Performance Evalⁿ

- where $\mathbf{1}$ is a column N -vector all elements of which are unity.
- We make the following observations:
- For any matrix A , $\mathbf{1}^T A \mathbf{1}$ represents the sum of all entries of A .
- The element (p, q) of the matrix $\prod_{i=1}^l G(\mathbf{e}_i)$ enumerate the Euclidean distance involved in transition from state p to state q in exactly l steps.

36

 Indian Institute of Technology, DelhiRanjan Bose
Department of Electrical Engineering

One I have said is a column vector of N length vector.


(Refer Slide Time: 31:57)

Information Theory, Coding and Cryptography

Performance Evalⁿ

- Our next task is to relate the above analysis to the probability of error, P_e .
- It should be noted that the error vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_l$ are not independent.
- The error state diagram has a structure determined only by the linear convolutional code and differs from the code state diagram only in the denomination of its state and branch labels ($\mathbf{G}(\mathbf{e}_i)$).
- Since the error vectors \mathbf{e}_i are simply the differences of the vectors \mathbf{c}_i , the connections among the vectors \mathbf{e}_i are the same as that among the vectors \mathbf{c}_i .

37

 Indian Institute of Technology, DelhiRanjan Bose
Department of Electrical Engineering

And so, we would like to finally, get to this P_e , the probability of error.

(Refer Slide Time: 32:08)

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

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

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, Delhi
38
Ranjan Bose
Department of Electrical Engineering

And, if you do this you can write out in terms of that modified state diagram T of D, where D is as earlier e raise to the power minus 1 over 4 N naught.

So, here T D is very easily written as 1 over N 1 vector transpose G 1 vector and this in some sense you have seen earlier in the modified state diagram, and G matrix is defined as follows. T D is called a scalar transfer function or simply the transfer function of the error state diagram. So, once we learn how to calculate, and we will try to see an example, then it is very easy to get an upper bound on that.


(Refer Slide Time: 32:59)

Information Theory, Coding and Cryptography

Example

- Consider a rate $\frac{1}{2}$ TCM scheme with $m = 1$, and $M = 4$
- It takes one bit at a time and encodes it into two bits, which are then mapped to one of the four QPSK symbols.
- The two-state trellis diagram and the symbol allocation from the 4-PSK constellation is given in Fig 7.15.
- Let us denote the error vector by $e = (e_2 e_1)$.
- Then,

$$\mathbf{G}(e_2 e_1) = \frac{1}{2} \begin{bmatrix} D^{\|f(00)-f(00 \oplus e_2 e_1)\|^2} & D^{\|f(10)-f(00 \oplus e_2 e_1)\|^2} \\ D^{\|f(00)-f(01 \oplus e_2 e_1)\|^2} & D^{\|f(11)-f(00 \oplus e_2 e_1)\|^2} \end{bmatrix}$$

 Indian Institute of Technology, Delhi
39
Ranjan Bose
Department of Electrical Engineering

So, let us quickly go through an example to see how this works. We have a rate 1 by 2. So, m is equal to one m over M plus 1, TCM scheme and M is equal to 4 capital M. So, it takes one bit converted into two bits and two bits required QPSK to be used.

Now, the two-state trellis diagram is we will just show it and the error vector e will be as follows. So, this is an example how G e 2 e 1 can be written, ok. So, we have this 00, 10, 01 and 11.

(Refer Slide Time: 33:45)

Information Theory, Coding and Cryptography

Example

Two-state trellis diagram.

$$\begin{aligned}
 \mathbf{G}(e_2, e_1) &= \frac{1}{2} \begin{bmatrix} D\|f(00) - f(00 \oplus e_2 e_1)\|^2 & D\|f(10) - f(00 \oplus e_2 e_1)\|^2 \\ D\|f(00) - f(01 \oplus e_2 e_1)\|^2 & D\|f(11) - f(00 \oplus e_2 e_1)\|^2 \end{bmatrix} \\
 &= \frac{1}{2} \begin{bmatrix} D\|f(00) - f(e_2 e_1)\|^2 & D\|f(10) - f(\bar{e}_2 e_1)\|^2 \\ D\|f(01) - f(e_2 \bar{e}_1)\|^2 & D\|f(11) - f(\bar{e}_2 \bar{e}_1)\|^2 \end{bmatrix}
 \end{aligned}$$

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

And, here are the two-state trellis it is a simple example, where we have this QPSK and the four symbols are being used and you can easily write G e 2 e 1 as follows.

(Refer Slide Time: 34:05)

Information Theory, Coding and Cryptography

Example

- Where $\bar{e} = 1 \oplus e$.
- The error state diagram for this TCM scheme is given by

- The error state diagram.
- \mathbf{I}_2 is the 2×2 identity matrix.
- In this case there are only three error vectors possible, {01, 10, 11}.

We have

$$\mathbf{G}(01) = \frac{1}{2} \begin{bmatrix} D^2 & D^2 \\ D^2 & D^2 \end{bmatrix}, \mathbf{G}(10) = \frac{1}{2} \begin{bmatrix} D^+ & D^+ \\ D^+ & D^+ \end{bmatrix}, \text{ and } \mathbf{G}(11) = \frac{1}{2} \begin{bmatrix} D^2 & D^2 \\ D^2 & D^2 \end{bmatrix}$$

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

So, finally, if you follow the steps we have these three matrices G 01, G 10, G 11 in terms of this D.

(Refer Slide Time: 34:25)

Information Theory, Coding and Cryptography

Example

- We obtain the matrix transfer function of the error state diagram as

$$\mathbf{G} = \frac{1}{2} \frac{D^6}{1-D^2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

- The scalar transfer function, $T(D)$, is then given by

$$T(D) = \frac{1}{2} \mathbf{1}^T \mathbf{G} \mathbf{1} = \frac{D^6}{1-D^2}$$

- The upper bound on the probability of error can be computed by substituting

$$D = e^{-\left\{ \frac{1}{4N_0} \right\}}$$

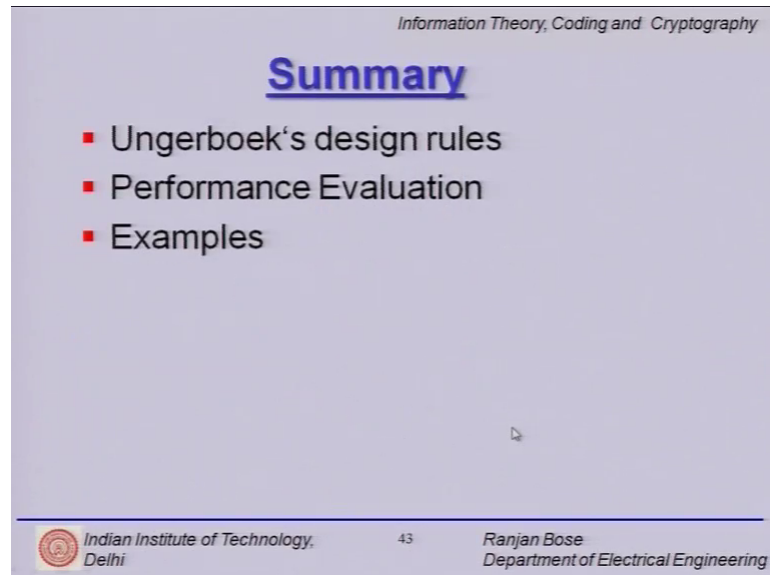
$$P_e \leq \frac{D^6}{1-D^2} \Bigg|_{D=e^{-\frac{1}{4N_0}}}$$

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

And, very quickly we can calculate this scalar transfer function T D and it comes out as D square D raised to the power 6 over 1 minus D squared we could have solved it using a traditional method also using dummy variables in the middle we have learnt how to solve this. So, T D comes out to be this. Once we have the T D in any way you would like to

calculate then you substitute this D equal to e raised to the power minus 1 over $4 N$ naught and you have the upper bound on the probability of error.

(Refer Slide Time: 34:58)



The slide is titled "Information Theory, Coding and Cryptography" at the top right. The main heading is "Summary" in blue. Below it, there is a bulleted list with three items: "Ungerboeck's design rules", "Performance Evaluation", and "Examples". At the bottom left is the IIT Delhi logo and text. At the bottom center is the number "43". At the bottom right is the name "Ranjan Bose" and "Department of Electrical Engineering".

Information Theory, Coding and Cryptography

Summary

- Ungerboeck's design rules
- Performance Evaluation
- Examples

Indian Institute of Technology, Delhi

43

Ranjan Bose
Department of Electrical Engineering

So, this gives a basic idea. So now, we would like to kind of summarize what we have done today. We have looked revisited actually Ungerboeck's design rules, where we would like to understand how the parallel transitions and diverging and merging back paths are assigned. Then, we looked at how to evaluate TCM schemes over additive white Gaussian noise channel. We made the observation that d_{free} is the single most important parameter for TCM schemes that we will use and it comes out that the probability of error upper bound is strictly dependent on this d_{free} notion. We also looked at certain examples to see how we can calculate this probability of error.

With that we come to this end of this lecture.

Thank you.