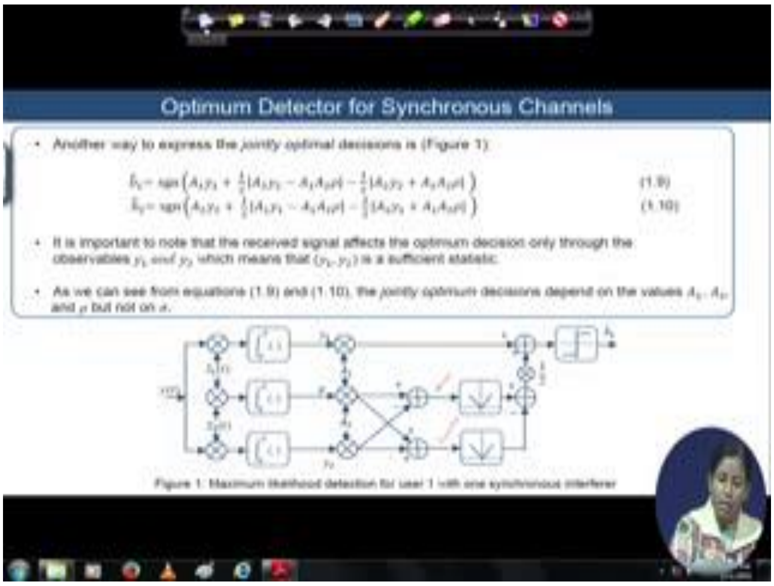
Spread Spectrum Communications and Jamming
Prof. Debarati Sen
G S Sanyal School of Telecommunications
Indian Institute of Technology, Kharagpur

Lecture – 60
Multiuser Detection – Part II

Hello students. We were discussing about the performance of this optimum multiuser detection. We will continue the concept in this module also. Here we will try to see how the optimal multiuser detection concept is actually deployed and implemented in the circuit.

(Refer Slide Time: 00:42)



In this part two, we will start with again that jointly optimal decision. In the last module we have learnt that the decision can be based on the individual maximizing, the individual a posteriori probability or either doing that or by the jointly maximizing the a posteriori probability. So, this jointly optimal decision we have learnt that this is actually much more preferable because they are less complex. Competition complexity is less and they can also give a very close minimal error actually in finding out the error probability finally, if I compare with the individual optimal decision. We will continue hence with the joint optimal decision in this module.

So, now we have seen one way of implementing the joint optimal decision process in the last module. It can be also deployed in some other way. So, let us see the figure and the concept here. So, this is incoming signal that you have received in the frontend of the receiver. We understand that we have bank of the matched filter concept, but we will restrict our discussion with two users scenario that we were doing in the last module. So, we it seems that we have two signatures with hand S 1 t and S 2 t and we have the co dilator architecture. So, we have not utilized the sampler followed by the co dilators and we have a knot because; that means, we are still actually in the continuous domain. We have not discretized it and finally, the architecture shows something like this. So, see the signature S 1 t is actually getting multiplied with the incoming signal in the upper path generating y 1. S 2 t is getting multiplied with the lower the path with the y t and generating the signal y 2 t. And remember we assumed that the two user signals are synchronized. So, it is a synchronized channel.

And then what we are doing additively in this structure in the receiver is also S 1 t and S 2 t are getting multiplied also here and; that means, the incoming. So, here S 1 t S 2 t multiplication is basically iteration over this duration, is basically generating that cross correlation factor rho and once this cross correlation factor is coming here, we understand and that we have an estimated value of A 1 and A 2 associated with the two signals transmitted by the two different users. So, y 1 will be multiplied here with this A 1 A 2 will be multiplied with y 2 in the lower zone, but also this A 1 A 2 will be multiplied by this cross correlation factor rho in between. These are the some of the fact, some mathematical operation going on the top of this and this is a multiplied finally, this architecture is finally, giving you the estimate about the b 1.

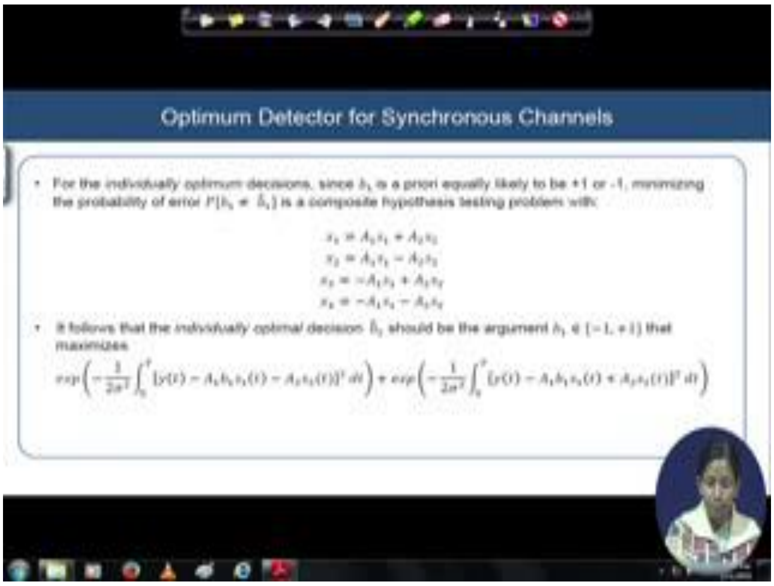
So, we understood in the last module the estimate of the likelihood estimate, joint estimate b 1 comma b 2 will be required if we try to go ahead with the joint optimal detection that will maximize the a posteriori, joint a posteriori probability at the output. And here that is why in our first task was to find estimate the b 1 and b 2. This architecture will give you the estimate of b 1. Similarly actually by changing the variables we can go ahead with estimating the b 2. The expressions will be given like this.

You have the b 1 estimate will be given by A 1 Y 1 plus half of this A 1 Y 2 minus A 1 A 2 rho, let us take a mod of it and subtract from this half of this one quantity b 2 t, so equivalently will be given by this. So, what finally, we are seeing here is deployed. See here you have a A 1 Y 1 which is this basically at the output here and half of A 2 Y 2 minus this guy. So, this factor is getting generated here, this factor is getting generated here. So, this term the plus A 2 Y 2 minus A 1 A 2 into rho that will be here and this guy A 2 Y 2 plus A 1 A 2 rho it will be the 2 additive terms are getting generated here. So, then actually the term is basically this plus and minus will be simply added and then multiplied with the factors, corresponding factors of half and this guy you are generating it were adding with this A 1 Y 1 and then it is a signal function basically taking care of actually I am telling you this, after this addition whether the value is positive or negative. So, that is the estimate about the b 1.

So, the whole architecture is implemented in this way and equivalently with b 2 we can generating the it taking the A 2 Y 2 and just reversing the variables instead of A 2 Y 2 I am using here A 1 Y 1. But remember it is important here to note that this received signal whatever this received signal affects the optimal decision of this b 1 and b 2, they are actually where the variables are. If I see actually the all expression of b 1 and b 2 I see that the terms like A 1 Y and rho are coming. Nowhere the variance of the noise I mean sigma is nowhere coming out.

So, the good part of this decision is joint optimal decision is and this architecture is that it will be depending up on your estimated value of A 1 a 2; that means received amplitude of each and every user, but and also the cross correlation matrix, but it will never be dependent on the noise variance. So, you are getting need of the noise affect, noise variance actually in the decision process. So, noise variance can never affect your decision process itself if you are following the joint optimal architecture.

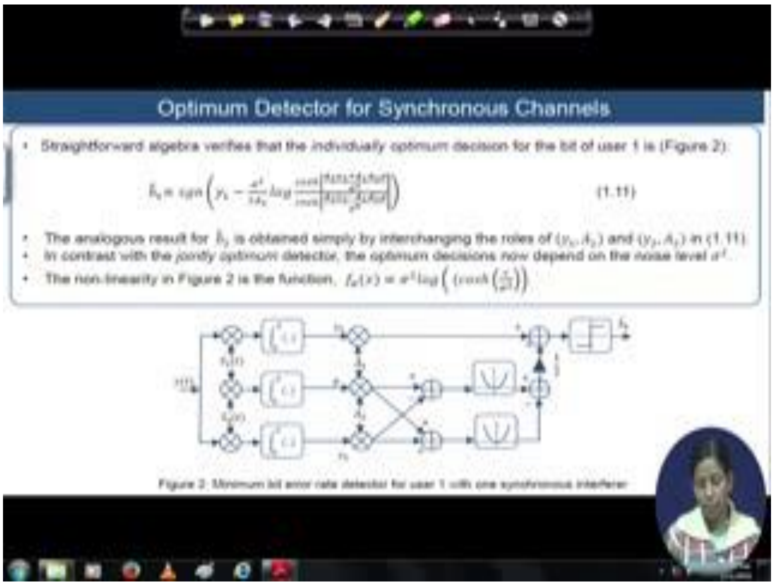
(Refer Slide Time: 07:19)



If it is an individual optimal decision, see then let us see how does it go? Now in this situations since my b 1 is an a priori equally likely to be either plus 1 or minus 1. So, minimizing remember individual optimal decision always try to minimize the error probability of an individual user. So, when I am talking about b 1. So, it will try to minimize the error a probability of error of user 1 and hence the probability of error of b 1 not equal to b 1 hat. So, it is minimizing that error. So, it is a composite that decision can be done by a composite hypothesis testing given by this x 1, x 2, x 3, x 4, where, x 1, x 2, x 3, we have discussed earlier they are the combination the all possible four combinations of receiving plus 1 and minus 1. And it follows that this individually optimal decision this b 1 hat should be argument b 1, that belongs to minus 1 and plus 1 and that should this b 1 hat should be estimated in such a fashion that it should maximize this guy, this bigger expression.

Basically actually this is a noise component once again I expressed earlier and this is the combined noise effect for the combinations of minus 1 minus 1 or plus minus 1 and if you can select the b 1 hat in such a way that this whole expression is minimized. Sorry this is maximized then fundamentally equivalently it will minimize the error probability. So, it means that it will be very close to the actual value of the b 1 transmitter and hence the error probability will be minimized.

(Refer Slide Time: 09:09)



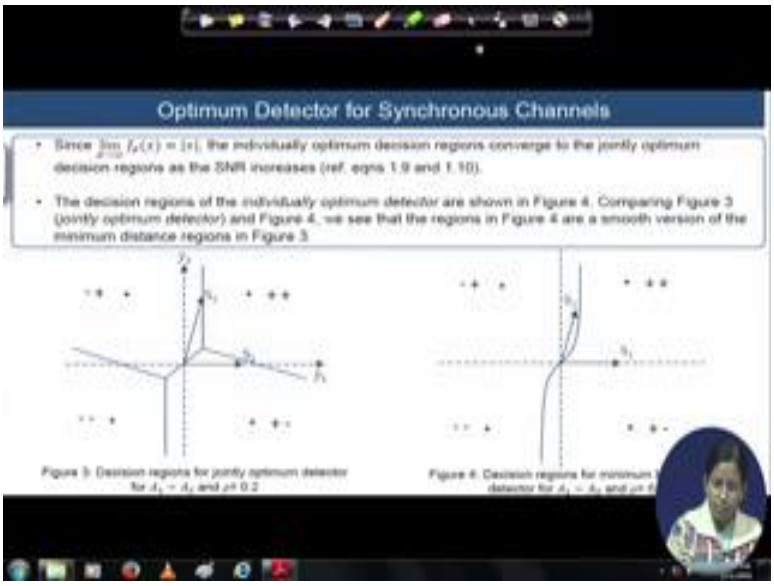
And straight forward algebra, if I try to put the there that will the individual optimal decision pattern and the estimates will be b 1 estimate for b 1 will be given by this. So, it is coming actually from the previous expression and there to be some straightforward calculations and we are ending up with the estimate of the b 1 given by equation 1.11. So, this estimate for this individually optimal decision can be implemented actually in circuit like the figure given in figure 2. What we are doing? We are taking the same signal S 1 T S 2 T multiplication is done. So, you have generated y 1 rho and y 2 like the earlier case because those three terms are heavily required. You have to also have in hand the estimates of the A 1 and A 2. So, you have the estimates.

So, now, this is just the combination of this A 1 Y 2 plus A 1 A 2 rho that you have to generate. And remember this non linear function that is given here, actually this is given by the fact of this sigma square log of cos hyperbolic x by sigma square which is coming here. So, this is the part that is implemented by this non linear function inside the equation and multiplication of half is anywhere required here the half factor is there. So, you are here and finally, after adding these two I am taking the, applying the signal function you will be getting the estimate of the b 1.

So, these are implementation of this equation of 1.11 that we have declared. We have decided and discussed, but remember 1 thing what is the basic difference that we will now see between 1.11 and 1.10? See here we are getting the parameters like A Y rho and also sigma. So, they are individual optimal decision is basically dependent on the noise variance. So, here is the basic difference between the individual optimal decision and the joint optimum decision. You need not to worry about the noise variance or the decision is completely free from the noise variance for the joint case. Whereas, the individual optimum decision will be largely dependent up on the noise variance and noise is a completely unpredictable one. It is a random variable we will have the statistical property associated with it, but estimation of that vector or having a priori knowledge of the sigma square is not feasible in practice.

So, these are the fundamental differences you should understand between the individual optimum decision and the joint optimum decisions and before taking on the how to, which should be applied in a particular scenario. The receiver taking the decision you should be having a clear knowledge about these two.

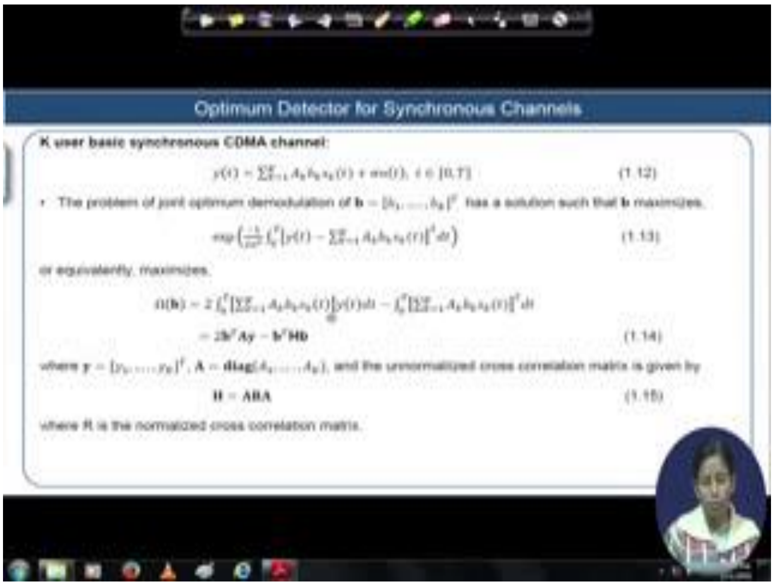
(Refer Slide Time: 12:08)



Remember, another thing that since this function I mean limit some function of sigma of x is equal to mod x if sigma tends to infinity. This individually optimum decision regions basically they converge to the jointly optimum decision regions. It was the jointly optimum decision region we have seen earlier and for actually some typical values of A 1 equal to A 2 and the co variance factor, cross correlation factor is equal to 0.2 and that will be coming here actually for when sigma tends to infinity.

That will be the region will be equivalently given by the figure 4, which I can see the comparing figure number 2, figure 3 and figure 4 I can see actually basically this individual optimum decision region is nothing, but a smooth function of this jointly optimal decision, when the sigma is very very high.

(Refer Slide Time: 13:21)



So, lots of similarities between these two are there. Now we prefer to come into the zone of K user basic synchronous CDMA channel. So, we extend our discussion from a 2 users to K number of the users. So, the equation that we have seen we will revise it in the context of the K users. Earlier we saw we have two terms here A 1 b 1 s 1 t plus A 2 b 2 s 2 t plus sigma n. Here it will be a combination of all the K users, number of the users going 1 to capital K and remember we are interested over the time interval of 0 to capital T. Capital T is the inverse of capital T will give you the data rate, transmission data rate.

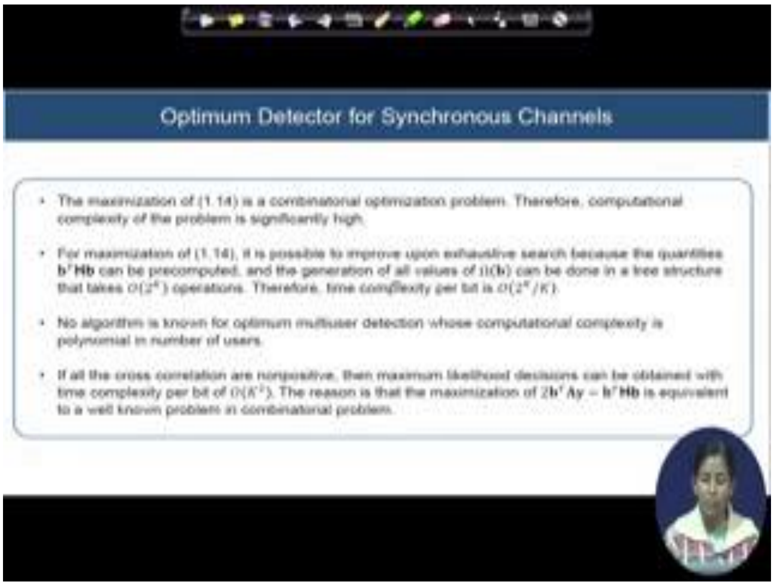
The problem of this joint optimum demodulation of this b, here remember actually for b we will have a complete vectors actually it is a matrix of b 1 to b K. So, if I try to have a joint optimum demodulation of this capital b of this bold b it has a solution such that this b maximizes now, this combined factor. I mean this y minus this combination of y minus this factor the first term will give you actually the noise variance. The part portion of the noise not the variance, it is the portion of the noise. So, basically when you are trying to do the joint optimum demodulation you are trying to search such value of the b that maximizes this expression 1.13.

So, basically it should minimize this noise. You are basically trying to maximize the b. I am trying to maximize this factor means you are trying to finally, minimize the error probability. Or equivalently we saw that we find a factor omega in the context of the 2 user case. That is here equivalently for b number of the users, for b number of the data to be optimum demodulation of the vector. We will be able to which should be equivalent given that you are trying to maximize this capital omega given by this expression.

Here actually we can have two factors of this omega. The first one actually equivalent in the vector form can be written as b transposed this A vector and some part than this can be contributed by the received vector y and this is again written as this guy where, this H that we have written. H is actually nothing, but a capital R into A. We have seen that the capital R earlier is normalized cross correlation matrix and capital A is a diagonal matrix and it is un normalized and. So, if I break this it will end up with this b transpose A capital R A into b. So, where we have replaced A capital R A into e by an and the vector capital H.

So, either you; that means, the for K basic user synchronized CDMA channels either you try to finding out the capital b vector sorry bold b vector means basically you are trying to find a solution such that, the selected value of the estimated value of this bold b should maximize either equation 1.13 or equivalently you maximize this equation 1.14.

(Refer Slide Time: 16:57)



But remember one thing, this maximization of this 1.14. This is a combinatorial optimization problem where, actually and combinatorial optimization problems we will have a very high complexity associated with it because some exhaustive search process actually is close to that it involved in the such mechanism itself, in the combinatorial optimization and. So, it may seem to us initially that this second process I mean equation 1.14 involves, it invokes actually very high computational complexity and it may not be feasible interact is to implement.

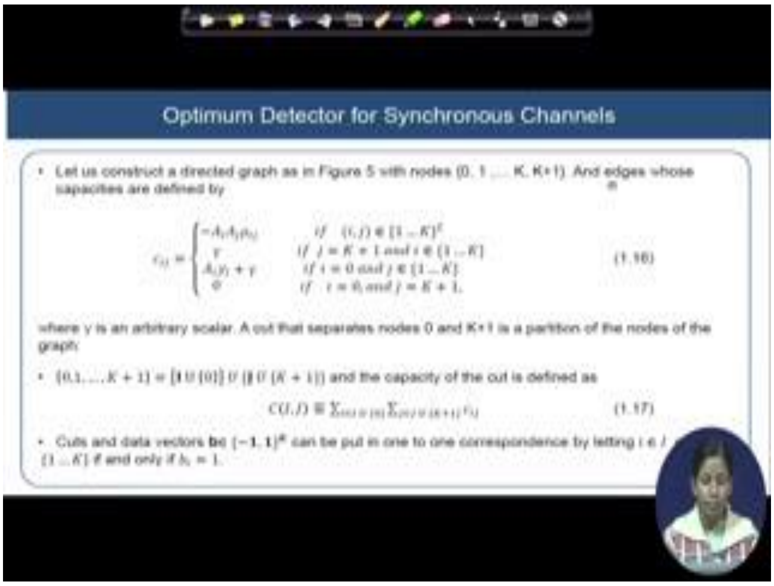
But fact is that those terms this b transpose A y and the sorry the second term b transpose H b that actually can be pre computed and it can be stored in such a way that all values of the generations of the all values of this sigma b it can be done first and then in a we can form a tree structure such that actually, it can take actually for every bit it can take actually 2 to the power K number of the operations.

So, remember one thing whenever we design a detector and we try to understand the performance analysis of this detector. The efficiency of a detector will be always given by the complexity of detecting 1 bit given the total number of the bits. So, how good our detector is? Or how efficient the detector is? Always is given by the time complexity involved per bit detection which, is given by this expression of 2 to the power that the total complexity involved. I mean in the operation how many number of the operations you are involving? Given by the total number of the bits and. So, we were always talk about this is a time complexity involved in the detection of the bit. So, whether we can what is the parameter at the output coming up for the output of this optimal detector is a point of the concern where the efficiency of the detector we will be defining based on these parameters always.

So, see actually there is no algorithm which is known for this optimum multiuser detection, whose computational complexity is polynomial in the number of the users. And if all these cross correlations are the non positive one then, this maximum likelihood decision it can be obtained by this time complexity per bit where, the operation number of the complexity complex operations that are involved will be given by the number of the bits square. And the reason is that the maximization of this term to be transpose A y minus b transpose of this H b which is involved here in the competition of this capital sigma b. It is equivalent to a well known problem in a combinatorial problem in such a situation.

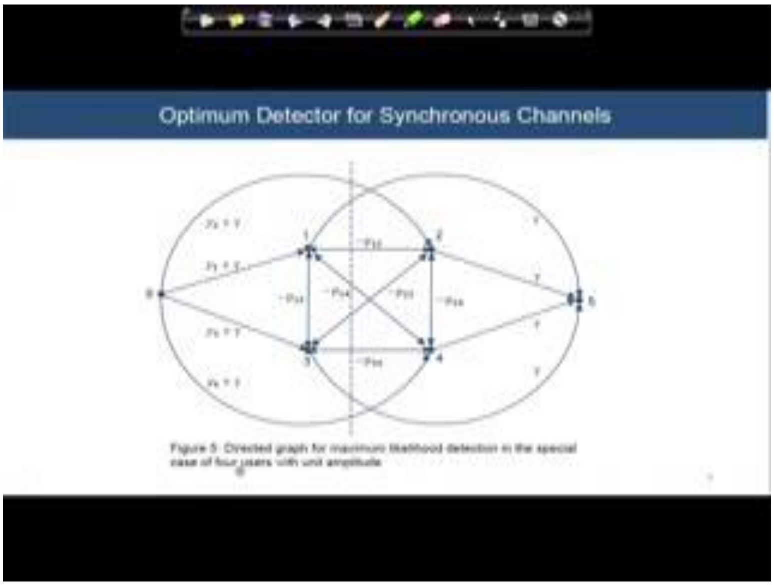
So, if we can guarantee that all the cross correlations are this non positive. If it is a non positive all the cross correlations then actually then in maximum likelihood decisions will be, the complexity of this competition of this capital sigma we will be little bit it will be actually given by this K square. So, this is all about actually how to define detectors efficiency? And how do we understand it? And how to relate it actually with this K user basic synchronous CDMA channel case where, the expression is related to the maximization of the sigma b whether, it is feasible in practice or if it is not. If it is there understand that this is a commentarial problem and it involves a huge complexity and how we can reduce the complexity by a typical tree architecture.

(Refer Slide Time: 20:54)



Now, let us construct go little bit more and let us construct a graph. A directed graph where, we understand that we are having certain nodes, all these nodes are corresponding to the number of users. So, we are having a users say from node number 0 to K plus 1 and all the nodes are 1 1 users and they are, there is a edge connecting the nodes and we call them in the graph they are the edges.

(Refer Slide Time: 21:25)

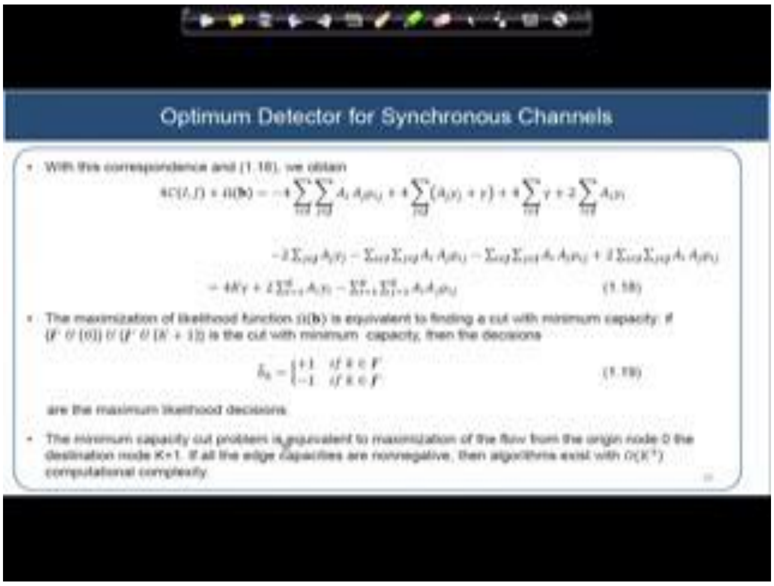


And it actually basically so we have from 0 to say K plus 1 number of the nodes. These are the users and they are communicating with each other. So, these are the channels between the user numbers 0 to user number 1. And the capacity of these channels we have defined here. So, the (Refer Time: 21:45) capacity of a channel C i j actually will be either any four of it, any one of this four steps depending up on the conditions that are defined in the equation 1.16. And remember we have introduced a new parameter gamma here other terms are known to us, but this gamma is an arbitrary scalar quantity and in this graph when once we actually put the nodes and construct the connection between them then it forms a graph. And then let us understand that what is a meaning of a cut inside a graph?

A cut is basically that which it is basically a cut that separates the node 0 and node K plus 1. And then the partition of the basically it is the cut means the partition of the nodes over the graph. If that happens then actually the all the number of the nodes the total number of the nodes it can be written like this. So, it is the matrix I, which is the union of the 0 node which is with the union of all the J nodes union with the remaining K plus 1 number of the nodes and their capacity will be of this cut can be written by the expression 1.17 from the graph theory it is coming.

So, the cuts and the data vectors this b e they can be put in 1 to 1 correspondence, by letting this small i here, by here this small i which belongs to this capital I and which is the subset of this 1 to K. So, if that cut is there and then the cuts and the data vectors it will be can have a 1 to 1 correspondence if and only if that b i is equal to plus 1.

(Refer Slide Time: 23:37)

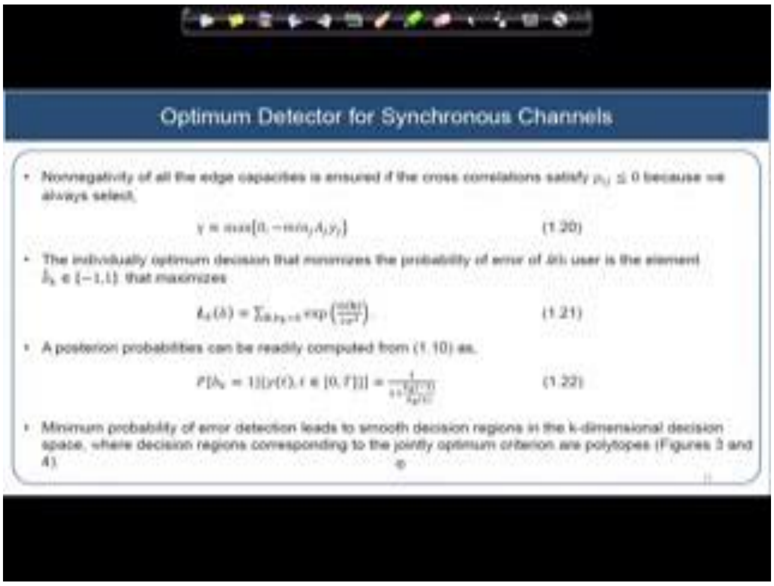


And some correspondence with this correspondence of this 1 to 1 and using this expression of this capacity allotment per edge and then we can actually come to an expression where, the sigma b plus 4 the capacity and the combination of the capacity as well as this sigma b is equivalently can be written as this bigger expression. Remember actually for each and every case what we are trying to do here is for the i-th one and for the i-th set we are having a after the cut you will have to set straight. One is denoted as i-th set and another is j-th set and we are trying to find out actually associating with the for each and every set we are trying to associate corresponding edge capacity and we are ending up here and so it can be further just simplified with the expression given here. The maximization of this likelihood function finally, here this function sigma b it is a finally, equivalent to finding a cut with the minimum capacity. If you understand that this is the cut with the minimum capacity then, the decision of this beta k will be either plus 1 if the k belongs to I star and or minus 1 if it belongs to if the k belongs to J star.

So, now question is actually now when we are finding out this expression of 1.18, the question is whether how will you finally maximize this likelihood function now boils down to the fact? That you find out the cut in such a way that it boils down with the minimum capacity. The cut actually is equivalently it is finding out a cut over the graph such a way that actually it is corresponds to a minimum capacity and if the minimum capacity is ensured then you can take a decision of this estimate of this beta k governed by this equation 1.18. And how I we came actually from this bigger equation to here? Is we the wherever actually there is no association with this capacity, the terms who are not related to this capacity we have excluded those terms and we have come up with the maximization of this likelihood function equivalently to the finding out the cut that minimizes the total capacity.

And these are the maximum likelihood decisions equivalently and the minimum capacity cut problem is equivalent to the maximization of also of the flow from the one origin node 0 to another decision node K plus 1. And if all the edge capacities are then in that case of non negative then, the algorithm such algorithm will you can device and find out the algorithms for which actually the operations, number of the operations that you need to find out to reach to the estimate of the beta k will be governed by the number of the users cube. So, that will be the computational complexity involved in estimating this beta k by this cut and finding out the cut with the minimum capacity.

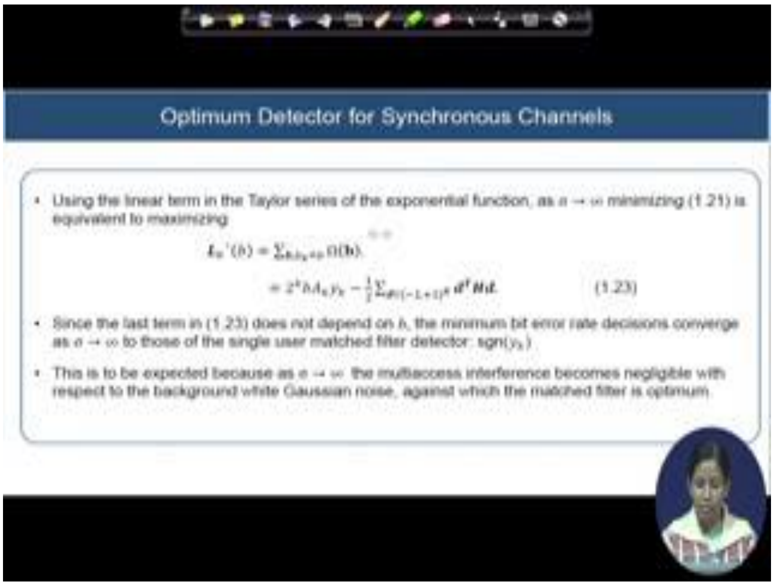
(Refer Slide Time: 26:55)



So, non negativity of all these edges, edge capacities is ensured if the cross correlation satisfy less than equal to 0 because we always select the value of the gamma which is the maximum of 0 or minimum of over the j of A j into Y j. So, the individually the optimum decision that minimizes this error probability of the k-th user finally, is finding out of the element beta hat k that belongs to either minus 1 or plus 1, that will maximize equivalently this expression 1.21.

So, finding out actually we started with the maximization of this capital gamma, capital omega and it boils down to the fact that equivalently it is maximization of this factor L k and where we understand that the a posteriori probability can be readily computed from 1.10 like this involving the value of this L k. And minimum error probability of this error detection then leads us to the smooth decision region in the k dimensional decision space and figure number 3 and 4 that we have shown earlier.

(Refer Slide Time: 28:12)



In this expression of this L k where, this exponential term is involved if I expand it with the Taylor series the expansion and we can actually also terminate it with the series and then the exponential function. And as the sigma goes to the infinity minimizing this 1.21 basically is equivalent to the maximization of this whole guy we understand and where this d belongs to this minus 1 and plus 1. And we understand this whole term of 1.23 as it never depends actually on the whole term does not whole term is not depending upon the factorial b; basically, the last term is or the second term is not basically depending upon the term b. So, the maximization of this L k star b will be mainly governed by the first term only.

And so, when actually sigma will be going to the infinity, so it will fall down to the single user matched filter detector, which is basically equivalent to the sign of Y k. And this is expected also because as sigma goes to the infinity we basically or multi axis interference becomes insignificant. When sigma is infinity means your noise dominating zone. So, that noise domination means you are at a very rho SNR zone where, your multiple axis interference very low compared to your noise variance. The contribution of the noise is mainly dominating the performance of the whole network and the receiver architecture and. So, it is a background white noise which will actually mainly actually detect the decision output at the output of this optimum detector and against. In such a situation again the matched filter will give you the optimum output because matched filter performs the best in presence of your noise AWGN noise.

And so in this whole conclusion is actually when the SNR is low you understand that it will be dominated, the performance will be dominating by the noise not the multiple axis interference of the network. So, the matched filter based detection will give you the optimum detection and if it is not you are in the higher zone of the SNR where, actually the multiple axis interference will be dominating you, then actually the detection should be preferable to go ahead with the jointly optimal decision technique and equivalently the formulation shows that you find out a beta hat following actually the maximization of this parameter given by 1.23.

And if you can find out a beta that maximizes this expression 1.23, basically that is the value of the, optimum value of beta b transmitted. So, that is the optimum detector structure in a higher SNR zone.