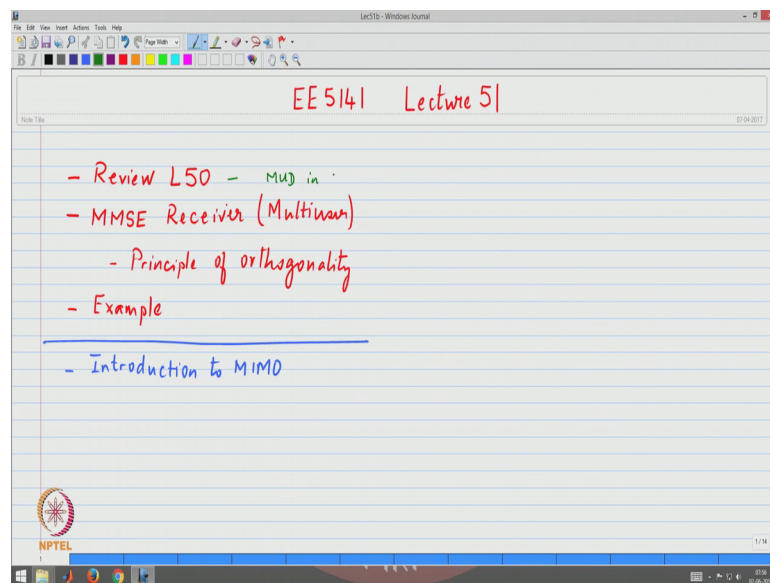


Introduction to Wireless and Cellular Communication
Prof. David Koilpillai
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture – 50
CDMA Receivers
CDMA Multiuser Detectors – Part 2

Good morning and welcome to lecture 51, in last lecture we have been covering some of the topics pertaining to multiuser detection in CDMA. So, we would like to begin today's lecture with a quick review of lecture 50 where we have been talking about the multiuser detection in CDMA and in today's lecture, we will introduce a new form of receiver a sub optimal receiver, but one that performs very well in the context of multi user environments that is called the minimum mean squared error receiver or MMSE receiver for multi user environments and this will be built on the well known principle of orthogonality in optimal filtering theory.

(Refer Slide Time: 00:37)

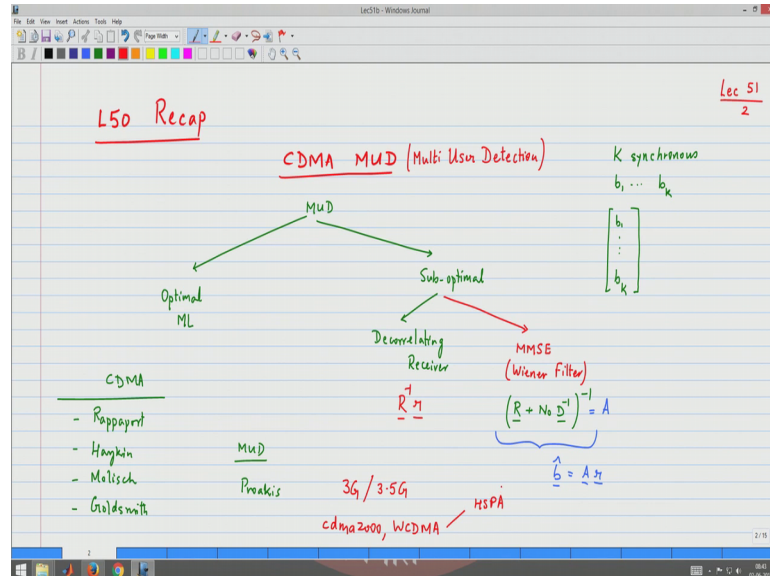


So, we would like to incorporate that in the context of an example, and with that we will complete our discussion of the introduction to CDMA.

The next unit that we will be covering will be the topic on multiple antennas; multiple antennas at the receiver multiple antennas at the transmitter. So, that would be refer to as a multiple input multiple output or a MIMO system. So, we move from the CDMA

context into the multiple antenna environment and that would be the next or the lateral half of today's lecture.

(Refer Slide Time: 01:43)



So, quick summary of the multi user detection that you have been so far; so, the problem of multiuser detection if you were to think of it as a family of solutions, the first solution that we refer to was the optimal or the maximum likelihood receiver, where we computed through exhaustive search what was the vector that was transmitted by each of the k users.

So, our assumptions are that there are upper K synchronous users, each of them transmitting a binary bit b_1 through b_k , and our goal is to estimate the vector b_1 through b_k at every instant of time that would be our multiuser detection problem statement. And we saw that the optimal solution would be the maximum likelihood solution, and it is a solution that turns out to be computationally intensive it is exponential and complexity in terms of the number of users that are present in the system, and also depends complexity increases with the size of the constellation that you are transmitting.

So, this would be a receiver that we would like to use, but if its complex. On the other hand we have also looked at suboptimal receivers. So, in the last lecture we introduced a very useful solution one that is widely used that is decorrelating receiver. So, this we showed is a receiver that has lower complexity, but at the same time achieves good

performance, it has some limitations which we talked about yesterday and we will also continue to discuss it today.

What we would like to do is present another sub optimal solution, but one that performs even better than a decorrelating receiver without a significant increase in complexity. So, that is a second suboptimal receiver, but one that performs close to the MMSE is also referred to as a wiener filter or a wiener filter based solution, we will refer to it as the MMSE solution.

(Refer Slide Time: 03:58)

The slide content is as follows:

Lec 51
3

Optimal ML receiver

$$\max_{\hat{\mathbf{b}}} P(\mathbf{r}_t | \hat{\mathbf{b}}) \quad \mathbf{r}(t)$$

$$\hat{\mathbf{b}} = \underline{\mathbf{b}}_j : \max_{\underline{\mathbf{b}}_j} \left(\underline{\mathbf{b}}_j^H \mathbf{r} + \mathbf{r}^H \underline{\mathbf{b}}_j - \underline{\mathbf{b}}_j^H \mathbf{R} \underline{\mathbf{b}}_j \right) \quad \text{ML Rx metric}$$

The slide also features an NPTEL logo in the bottom left corner and a small video inset of a speaker in the bottom right corner.

So, in a nutshell, the optimal ml receiver that we have talked about in the earlier lecture optimizes the probability maximizes the probability, that we have transmitted a vector $\hat{\mathbf{b}}$ and the received signal \mathbf{r} of t . Now \mathbf{r} of t is a continuous waveform. So, we are maximizing the probability that the received vector is \mathbf{r} of t , when the a vector $\hat{\mathbf{b}}$ was transmitted. So, in other words given \mathbf{r} of t we are trying to find that vector $\hat{\mathbf{b}}$ that maximizes this probability. So, using a series of a expressions and sets we have shown that this $\hat{\mathbf{b}}$ is nothing, but the vector \mathbf{b} that maximizes this expression. And this is the competition that we need to do for a maximum likelihood receiver and this is the other the basis of the formulation that that we have shown. So, this is the expression we have to try all combinations of the vector \mathbf{b} therefore, numbers of the combinations grow exponentially as the size of the vector grows or as the size of the constellation grows.

So, this is the maximum likelihood receiver metric, and you would optimize it over all possible \mathbf{b} ones and choose the one that gives us the best solution. So, this was the optimal receiver or the maximum likelihood receiver.

(Refer Slide Time: 05:39)

Lec 51
4

Decorrelating Rx - How it works

$$\underline{R} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \quad \underline{R}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$

$r_1 = \sqrt{E_s} b_1 + \sqrt{E_s} \rho b_2 + \eta_1$
 $r_2 = \rho \sqrt{E_s} b_1 + \sqrt{E_s} b_2 + \eta_2$

$\underline{R}^{-1} \underline{r} = \begin{bmatrix} \sqrt{E_s} b_1 \\ \sqrt{E_s} b_2 \end{bmatrix} + \begin{bmatrix} \frac{\eta_1 - \rho \eta_2}{1-\rho^2} \\ \frac{\eta_2 - \rho \eta_1}{1-\rho^2} \end{bmatrix}$
potential for noise enhancement

Outputs K correlators
Decorrelating Receiver
 $\underline{r} = \underline{R} \underline{b} + \underline{\eta}$
 $\underline{\hat{b}} = \underline{R}^{-1} \underline{r}$

Example

$$\underline{R}^{-1} \underline{r} = \begin{bmatrix} g_1'(t) \\ g_2'(t) \end{bmatrix} = \begin{bmatrix} g_1(t) - \rho g_2(t) \\ g_2(t) - \rho g_1(t) \end{bmatrix} \frac{1}{1-\rho^2}$$

$r(t) \rightarrow \text{sum} \rightarrow \hat{x}$
 \uparrow
 $g_1^*(t)$

NPTEL

Now on the other hand the decorrelating receiver was formulated differently. So, if the received vector not the received continuous time wave form, but this is the vector of outputs this is the outputs of the k correlators, and if you write down the expression for that the output of the k correlators comes out to be \mathbf{r} which is the correlation matrix which takes into account the correlation between the different spreading waveforms.

The vector of which transmitted by each of the k uses plus eta the noise component that is present in the measurement of each of the antennas. So, if this is our underlying equation then we should we showed that the decorrelating receiver competition is the following. It says that the best that what we have is the best estimate of the transmitted vector is nothing, but \mathbf{R} inverse times the output of the k correlators. So, that would be our expression for the decorrelating receiver, and again this is something that we spent time deriving.

Now, just as at the close of the last lecture we mentioned that it is good for us to understand the why a decorrelating receiver works and what are the underlying mechanism. So, if you were to write down the expression just for a two by two cases easy for us to visualize, we will see that the \mathbf{R} matrix can be written in terms of the

correlation between the two spreading wave forms we call it as ρ , R inverse we can write down the expression and we write down the expressions for the output of the correlators, each of the correlators will produce the following expressions and R inverse times r gives us the expression. So, this is the expression that we saw in the last lecture this is where we saw that the decorrelating receiver, completely separate the two users and therefore, we do not have any issues with the near far problem or the interference from one user signal leaking and affecting the decision of the other users.

So, we see that there is a clean separation of the 2, we did make an observation that there is a potential for noise enhancement because of the way the decorrelating receiver works we have in the noise term $1 - \rho^2$ in the denominator, ρ is the number that is less than 1. So, therefore, the denominator will become smaller and therefore the expect the noise variance is likely to increase. Another very interesting way to visualize why or how the decorrelating receiver works, is to re look at the expression and this is what we mentioned towards the end of the last lecture.

Now, if you were to think of the received signal r of t , now normally you would multiplied with g_1 of t , and you would then follow it up with the integrator. Now instead of that if you multiplied it with g_1 prime of t , where g_1 prime of t is an expression that is given in the following manner. g_1 of t minus ρ times g_2 of t into $1 - \rho^2$ if you think of this is a modified de spreading waveform and then you apply it is in this form r of t multiplied it with g_1 of t multiplied by g_1 prime of t followed by the integrator you would find that the expression that we obtained is exactly the expression that we got for r universe r .

So, in other words the multiplication by the inverse of the correlation matrix the correlation of the spreading waveforms, can be equivalently represented as a modified spreadings de spreading sequence which takes into account the correlation between the two the spreading users. Again this is an interesting interpretation that helps us understand how a decorrelating receiver works.

(Refer Slide Time: 10:04)

The screenshot shows a presentation slide with the following content:

Lec 51
5

$$\left. \begin{aligned} r_1'(t) &= \sqrt{E_1} b_1 + \frac{\eta_1 - \rho \eta_2}{1 - \rho^2} \\ r_2'(t) &= \sqrt{E_2} b_2 + \frac{\eta_2 - \rho \eta_1}{1 - \rho^2} \end{aligned} \right\} \equiv \underline{R}^{-1} \underline{r}$$

NPTEL

A small video inset shows a man in a light blue shirt speaking.

So, the r_1' of t which is what you would obtain by using this expression by using the r of t multiplied by g_1 prime star of t followed by the integrator, if you write down the expressions you would get the expressions given on this graph and in this page which is exactly the expressions that we got for when you multiply it by R inverse r .

So, this set of equations is identically equal to what we got by multiplying R inverse times r . So, both are equal and one is multiplying by matrix inverse the inverse of the correlation matrix the other one is doing the receiver using a modified spreading sequence and both of us both of them give us an interesting perspective.

(Refer Slide Time: 10:51)

Lec 51
6

Ex 2-user synchronous

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -0.08 \\ -0.47 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & 0.33 \\ 0.33 & 1 \end{bmatrix} \quad \sqrt{E_1} \cdot \sqrt{E_2} = 1$$

ML: $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right\}$ $\max b^H r + r^H b - b^H R b$

$$= [-3.76, -0.56, -2.12, -1.56]$$

Decorrelating Rx: $R^{-1} r = R^{-1} \begin{bmatrix} -0.08 \\ -0.47 \end{bmatrix} = \begin{bmatrix} 0.084 \\ -0.498 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

So, we have looked at the following example I would like us to quickly look at it once more it is a two users synchronous CDMA, at the receiver the values of the output of the first correlator and the second correlator are given by minus 0.08 and minus 0.47.

Now, given that this is a binary decision whether you want to declare it as a plus 1 or a minus 1 with the threshold being at 0 both of these would map to minus 1, because both of them have a value that is negative. We are given that the correlation matrix is 1 0.33 0.33 and one and using this information we first of all estimate the maximum likelihood solution, maximum likelihood requires us to try all combinations would be four possible combinations of the transmitted vector, for each of those we have computed the metric $b^H r + r^H b - b^H R b$, and with that we are looking at the following different values of the metric, to find out which is the maximum all of these are negative numbers. So, the maximum would be the least or the one with the and maximum of these numbers is minus 0.56, and that corresponds to the second sector 1 comma minus 1. So, this was what we obtained as the vect.

Now, if for the same problem if we had done the following, we had done the decorrelating receiver which is $R^{-1} r$. So, this would be the same as R^{-1} times the matrix the vector that we have received minus 0.08 minus 0.47. Please do a compute the R^{-1} matrix and you can find out that this corresponds to the following expression $R^{-1} r$ comes out to be 0.084, and this one corresponds to minus 0.498

and if you were now were to quantize this information into plus and minus 1, this would lead to the following quantization.

The first bit would be since its positive would be mapped to a plus 1 then the second bit would be to a minus 1 and we find that this is the same as the solution that the maximum likelihood receiver also obtained. And this is different from what you would have obtained had you applied the quantization directly to the output of the correlators. So, there you have made a mistake it would have been minus 1 comma minus 1, where as the correct answer is minus 1 plus 1 comma minus 1, which is consistent with the result obtained in the maximum likelihood case, and also with the with the decorrelating case. So, with this as the as the basis we would now like to move into the next section where we would like to understand and implement the performance of a maximum MMSE receiver, and this is the starting point for the second sub optimal receiver that we have we would like to study in today's lecture and this is the formulation ok.

(Refer Slide Time: 14:21)

MMSE Detector

Decorrelating Rx

Let $\hat{b} = \underline{A} \underline{r}$

MMSE $|\hat{b} - \underline{b}|^2$

$$J(\underline{b}) = E \left[(\underline{b} - \underline{A} \underline{x})^H (\underline{b} - \underline{A} \underline{x}) \right]$$

Principle of Orthogonality

$$E \left[(\underline{b} - \underline{A} \underline{x}) \underline{x}^H \right] = 0$$

$$= E \left[\underline{b} \underline{x}^H - \underline{A} \underline{x} \underline{x}^H \right] = 0 \quad \text{①}$$

$\underline{r} = \underline{R} \underline{b} + \eta$

$\hat{b} = \underline{R}^{-1} \underline{r}$

$$= \underline{R}^{-1} \left[\underline{R} \underline{b} + \eta \right]$$

$$= \underline{b} + \underline{R}^{-1} \eta$$

potential for noise enhancement

Lec 51
7

So, the decorrelating receiver. So, just for reference, the decorrelator receiver does the following computation which says we are given the expression that the output of the correlators is R times b plus eta and the decision vector b hat is given by R inverse times r that is the expression for a decorrelation receiver and of course, if you were to substitute for r this would be R inverse times R times b plus eta, and expanding upon this equation this would give us b plus R inverse times eta. And what we find is that the

decorrelating receiver is focused on obtaining an expression for the transmitted vector while it does not pay much attention to what happens to the noise, and this is where there could be a potential for noise enhancement.

So, there is a potential for noise enhancement in a decorrelating receiver, which is something that we have mentioned a multiple times because of its importance potential for noise enhancement. Now how would you formulate another receiver which takes into account the noise enhancement and this is where we would like to present the MMSE receiver. Now we make the assumption that in an MMSE receiver that our best solution \hat{b} is the vector of outputs of the k correlators, multiplied by a linear transformation basically a linear transformation of those vectors which is denoted by a constant matrix a .

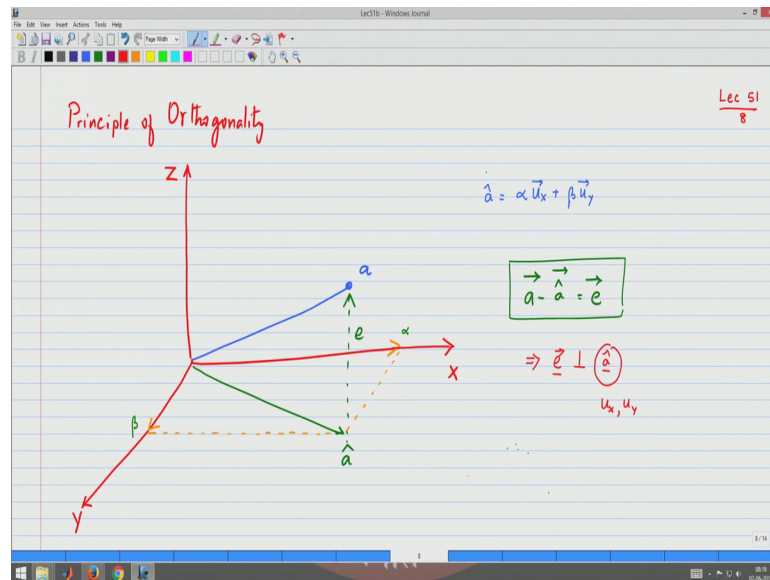
So, if you were to assume that the outputs are obtained as some linear operation on the output of the correlators k correlators, that would be given by \hat{b} is equal to A times r and this would be our expression for the linear representation or in terms of the out of the bits transmitted with respect to the output of the correlation. So, the MMSE criterion says that we would like to minimize the error, and what is the error that we are trying to minimize what does the vector that was transmitted and vector that was estimated and we would like to minimize the square of this with the expectation in front. So, the metric that we are looking at is given by the expected value of b minus \hat{b} , and b minus \hat{b} is nothing, but b minus A times r , b times where \hat{b} represented as A times r this is our expression for the output of the an MMSE detector.

So, this Hermitian times b minus A r . So, the MMSE formulation is can be stated in the following way very simply, we would like to find an expression for b which is a linear combination of the output of the k correlators. So, that is assumption the first step one, the second one where we find the optimum \hat{b} is by the finding the \hat{b} that minimizes the mean squared error between the transmitted vector and the vector that was at the obtained as a decision.

So, the metric can be written in the following form, where it is a expected value of the magnitude squared which is given by b minus A Hermitian b minus A r . Now one of the very powerful results in optimal filtering theory is something known as the principle of orthogonality and what we would like to do is give a intuitive expression for principle of

orthogonality, and use that in our derivation of the MMSE receiver. So, here is a quick overview of the principle of orthogonality. So, the principle of orthogonality says the following, very it is a geometric interpretation.

(Refer Slide Time: 19:28)



So, think of a three dimensional plane with x y z, x in the in the horizontal direction and y and z, and the point a that is in the three dimensional space and it has got coordinates in x y and z. Now we here is the principle of a orthogonality, now I would like to find an approximation to a, but an approximation that lies in the xy plane. So, xy plane; if you were to think of it as of in the following manner, that the approximation a hat will be alpha times the unit vector in the x direction U_x , plus beta times the unit vector in the y direction. So, if these are my unit vectors alpha and beta are scalar.

So, basically I am trying to find a approximation to this three dimensional vector, but in a two dimensional space. So, the question that arises is which would be the best possible approximation to a in the two dimensional space, in the two dimensional space I have represented it as a hat. So, in a most students of geometry would immediately point out, that point of closest approach which would be the error between the vector that you are trying to approximate and our approximation would be that which lies perpendicularly below it in the xy plane.

So, basically if we were to drop a perpendicular from a to the xy plane, and you take its coordinates that would turn out to be the best approximation to that. So, basically this

point would become your alpha this would become your beta, and then you would get the best possible approximation. I hope you are comfortable with the explanation a is a point in three dimensional space I am trying to approximate it in the two dimensional space in the xy plane and I find that the point that is perpendicularly below it is the one that gives me the best approximation.

So, if you were now to where to think of the error between the original vector and the approximation in the xy plane, and call it as e the vector e. So, basically $\mathbf{a} - \hat{\mathbf{a}}$ is the vector e or in other words $\hat{\mathbf{a}} + \mathbf{e} = \mathbf{a}$ which is the equation that is given here. Now given this expression given the geometric interpretation, it is easy for us to see the following result and this is a very very key observation. We find that the vector e is perpendicular to the vector $\hat{\mathbf{a}}$ because it is perpendicular to the plane because $\hat{\mathbf{a}}$ is perpendicularly below \mathbf{a} in the in the xy plane.

Now, this is also means that you will be perpendicular to the components of \mathbf{a} . So, U_x and U_y are the components of the unit vectors and the error has to be orthogonal to the to the components of $\hat{\mathbf{a}}$. In other words the error vector is orthogonal to the plane that is represented by $\hat{\mathbf{a}}$. So, this is what we referred to as the principle of orthogonality, the best possible approximation $\hat{\mathbf{a}}$ is that which produces an error which is orthogonal to $\hat{\mathbf{a}}$, the error must be orthogonal to the input vector that with which is the approximations to the to the given vector.

So, given this understanding of the principle of orthogonality let it go back and rewrite the principle of orthogonality. We are trying to find that solution that minimizes the mean squared error and from the principle of orthogonality we also know the following, the principle of orthogonality states that the error must be orthogonal to the vector. So, the expected value of the error $\mathbf{b} - \mathbf{A}\mathbf{r}$ this is the error and this will be orthogonal to the input vector \mathbf{r} . So, the orthogonality condition is given by error orthogonal to \mathbf{r} and where \mathbf{r} is the can you think of \mathbf{r} as the components of the best approximation and $\mathbf{A}\mathbf{r}$ being the best approximation of the received vector.

So, this is the principle of orthogonality and this must be equal to a 0 vector. So, this is a principle. So, the minimization of the squared error can be translated or simplified as a condition, which is given by the principle of orthogonality. So, now, the now that we have this form we would now like to simplify it and see what is the result that we obtain

from this from this expression. So, the expansion of this expression would be expected value of b times r Hermitian minus A times r Hermitian this has to be equal to 0.

Now, what we would like to do is a simplify it further and derive a expression for the results that that would be helpful for us. So, we call this as equation 1, now we would like to go ahead and expand it term by term. So, first let us take the first term in the approximation in the expression the first time in the expression we let us say is expected value of b times r hermitian.

(Refer Slide Time: 26:01)

The handwritten derivation on the whiteboard shows the following steps:

$$E[\underline{b} \underline{b}^H] = E[\underline{b} (\underline{R} \underline{b} + \eta)^H] = E[\underline{b} \underline{b}^H \underline{R}^H] + E[\underline{b} \eta^H]$$

Since $\underline{R} = \underline{R}^H$, this simplifies to:

$$E[\underline{b} \underline{b}^H] = E[\underline{b} \underline{b}^H \underline{R}] + E[\underline{b} \eta^H]$$

The second term $E[\underline{b} \eta^H]$ is shown to be 0. The first term is expanded using the definition of \underline{R} as a diagonal matrix of eigenvalues λ_i and eigenvectors \underline{e}_i :

$$\underline{R} = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_k \end{bmatrix} \quad \underline{b} = \begin{bmatrix} \sqrt{E_1} b_1 \\ \sqrt{E_2} b_2 \\ \vdots \\ \sqrt{E_k} b_k \end{bmatrix}$$

The expected value of the first term is then shown to be:

$$E[\underline{b} \underline{b}^H] = \begin{bmatrix} E_1 & & \\ & E_2 & \\ & & \ddots \\ & & & E_k \end{bmatrix} \underline{R} = \underline{D} \underline{R}$$

where $\underline{D} = \text{diag}\{E_1, \dots, E_k\}$.

Now, please write down or substitute the value for r, this would be expected value of b times R b plus eta Hermitian and this would be equal to if you were to expand it, it would be equal to b times b Hermitian times R Hermitian, but we know from the property of the auto correlation matrix that R is a has satisfies the property that R is equal to R Hermitian. So, we could equivalently write this as instead of R Hermitian we could write it as R.

Now, the second term in this expression. So, this would be the first term, the second term would be expected value of b times eta Hermitian. Now notice that one of these components is the input vector which is a binary vector equally probable plus or minus ones and the other one is the noise vector which again has got zero mean. So, the second term can be ignored because the expected value will go to zero what we have is the first term only.

So, even in the first term R is a constant matrix. So, what we are interested is in b b Hermitian and when we look at the expressions for b and b Hermitian, if we were to write it down this is $b_1 b_2$ up to b_k the number of users times $b_1 b_2$ up to b_k notice that the diagonal terms will give you $b_1^2 b_2^2$ all the way up to b_k^2 . So, when you take the expected value; expected value of a plus 1 squared is equal to 1 minus 1 square is equal to 1. So, you get ones along the diagonal, but keep in mind that the vector b itself actually incorporates the energies of the transmit signals.

So, it would actually be correct to call it as $\sqrt{E_1}$ times b_1 $\sqrt{E_2}$ times b_2 all the way to $\sqrt{E_k}$ times b_k . So, this would be the correct expression. So, what we should do is introduce the values $\sqrt{E_1}$ $\sqrt{E_2}$ $\sqrt{E_k}$ as the and here again we would introduce the $\sqrt{E_1}$ $\sqrt{E_2}$ and $\sqrt{E_k}$. So, that would b b hermitian.

So, when we look at the expected value this basically becomes a diagonal matrix where it becomes diagonal value $E_1 E_2 E_k$ of diagonal terms are a product of b_1 and b_i and b_j both are independent and each of them is equally zero mean. So, therefore, the half diagonal terms become 0. So, the this becomes the b b hermitian. So, this is expected value of b b Hermitian is given in this form.

So, this equation now becomes expected value of b times r Hermitian is nothing, but if you were to call this matrix has D . Where D is a diagonal matrix whose entries are E_1 through E_k then this becomes equal to D times R . So, the first express the term in the expression is a D times R . So, this is equation two.

So, going back to the expression we are trying to write down a expression for a based on the principle of orthogonality, the first time was b times r Hermitian, the second term was a times r Hermitian we are trying to get expressions for both we have now completed getting the expression for b times r Hermitian in terms of D times R .

(Refer Slide Time: 30:43)

The slide shows the following mathematical derivation:

$$E[A \underline{y} \underline{y}^H] = E[A (\underline{R} \underline{b} + \underline{\eta}) (\underline{R} \underline{b} + \underline{\eta})^H]$$

$$= E[A \underline{R} \underline{b} \underline{b}^H \underline{R} + 2 \text{ terms } \underline{b} \text{ or } \underline{b}^H + A \underline{\eta} \underline{\eta}^H]$$

$$E[A \underline{y} \underline{y}^H] = \underline{A} \underline{R} \underline{D} \underline{R}^H + \underline{A} \underline{N}_0 \underline{R} \quad (3)$$

Eqn (1) \Rightarrow $\underline{D} \underline{R} - \underline{A} \underline{R} \underline{D} \underline{R}^H - \underline{A} \underline{N}_0 \underline{R} = 0$ $\underline{R}^H = \underline{R}$

Subs in (1) from (2) & (3) $\underline{D} \underline{R} - \underline{A} \underline{R} \underline{D} \underline{R} - \underline{A} \underline{N}_0 \underline{D} \underline{D} \underline{R} = 0$

$$(\underline{I} - \underline{A} \underline{R} - \underline{A} \underline{N}_0 \underline{D}^{-1}) \underline{D} \underline{R} = 0$$

$$\underline{I} - \underline{A} (\underline{R} + \underline{N}_0 \underline{D}^{-1}) = 0$$

$$\underline{A} = (\underline{R} + \underline{N}_0 \underline{D}^{-1})^{-1} \quad \hat{\underline{b}} = \underline{A} \underline{y}$$

Now, let us move on to the next term in our expression the next term in the expression is expected value of a times r r Hermitian. So, this would be expected value of A times write down the expressions for R, the expressions for R is R times b plus eta, the second term would be r times the plus eta R Hermitian. So, this again if were to expand this product we would get 4 terms and what we would like to do is quickly look at what those 4 terms are.

So, the first term would be expected value of A times R times b b Hermitian times R. You should get R Hermitian, but that would be same as R that would be the first term, the second term will be there will be 2 terms which contain expressions for b or b hat. Now expected value of this would be zero because the individual components are zero mean and then the last time will be expression with the two noise terms it should be A times eta eta hermitian. So, these are the expressions that we have.

So, if you were to take the value of expression expectation inside, we can show that this becomes A times R times remember expected value of b b Hermitian, we got it to be equal to D and then we got it the write down this as A as R Hermitian. Now if were to write down the expression for expected value of eta eta hat. So, if you were to if you write down the expressions for that, we have already shown that this is equal to expected value of eta eta hat we have shown is equal to N naught times the N naught is a the noise

variance times the correlation matrix, where R stand represents the correlation between the at different spreading sequences.

Now, if you were to write this down this would become plus A times N naught times R and this is expected value of A times r r Hermitian this is right take this down as equation 3. So, equation 1 gave us the expression to go for based on the principle of orthogonality, equation 2 gave us an expression for the first term equation three gives us an expression for the second term.

So, if you now were to go back and substitute the values where d . So, then what we would get is equation 1 can be equivalently written as D times R from equation 2 minus A times R D R Hermitian minus A times N naught time R from the from equation 3 and this has to be equal to 0 we are basically substituting in equation 1 from 2 and 3 that is our basic operation which is what gives us this expression and we know also know that R Hermitian is equal to R . So, we can simplify this.

So, now we write a we do this following do the following expression. So, think of the following from where we write this as D times R minus A times R times D times R because R Hermitian is equal to R minus A times N naught I am going to write D inverse D basically that would be give me the identity matrix and basically substituting this D inverse D into this between N naught and R multiply it by R equal to 0. Now you may wonder can is it is D inverse does it exist yes it exist because D is a diagonal element matrix with the energies E_1 E_2 .

So, it is basically diagonal matrix with non zero entries. So, therefore, the inverse exist. So, there is no issue with writing D universe times D . So, now, notice that I can factor out D times R from this expression. So, then it becomes I minus A times R minus A times N naught D inverse times D times R equal to 0.

Now, from the situation we know that D is non zero R is non zero, R is the correlation matrix between the spreading wave forms. So, D and R cannot be 0. So, from the observations this term must be equal to 0, which basically tells us that I minus if you factor out the A matrix R minus N naught D inverse, R plus naught D inverse sorry must be equal to 0, which also tells us that A has to be equal to basically if you take the one term to the other side A it has to be equal to R plus N naught D inverse.

Then in other words A times R plus N naught D inverse z has to be equal to identity matrix. So, therefore, a has to be the inverse of that matrix. So, this is a very very important result. Now we started off by saying that MMSE solution, will give us an up estimate of the transmitted vector where b is equal to b hat is equal to A times r and the steps that we were doing is actually to find out the expression for the matrix a , which will give us the best possible solution b hat and through is using the principle of orthogonality writing down expressions for each of the terms given by the principle of orthogonality we have been able to derive the following expression.

Now what you would like to do is understand this expression a little bit more. So, the MMSE receiver that we have derived so far has the following properties.

(Refer Slide Time: 38:01)

MMSE $\hat{b} = A z$

$E \left\| \underline{b} - \hat{\underline{b}} \right\|^2$

$\Rightarrow \underline{A} = \left(\underline{R} + N_0 \underline{D}^{-1} \right)^{-1}$ MMSE Receiver $\hat{b} = \underline{A} z$ (output of K correlators $\begin{bmatrix} z_1 \\ \vdots \\ z_K \end{bmatrix}$)

Sp1 Case1 N_0 is small $\underline{A} \approx \underline{R}^{-1}$ MMSE same as Decoupling Rx

Sp1 Case2 N_0 is large $\underline{A} = \begin{bmatrix} E_1/N_0 & & \\ & \ddots & \\ & & E_K/N_0 \end{bmatrix}$ $\hat{b} = \underline{A} z$

We say that b hat will be equal to A times r , and we wanted to minimize b minus b hat magnitude squared expected value, and this resulted in us obtaining the matrix A in the following manner which is R the correlation between the spreading waveforms plus N naught that is the noise variance of each of the antennas multiplied by D inverse and the whole inverse.

Now, what I would be very helpful for us is to look at two special cases, the first special case that we would like to look at is. So, this is case one special case and we would like to look at the case where the noise variance is small. So, in other words the additive noise from each of the antennas the thermal noise is small compared to the other forms

of interference the multi user interface. So, N_0 is small. So, if N_0 is small compared to the in the expression then the expression here can be approximated as $R + N_0 D$ because N_0 is we ignore $N_0 D^{-1}$. So, this becomes R^{-1} notice that what this is telling us is the MMSE solution is the same as in this particular case MMSE same as the decorrelation receiver.

Same as the decorrelation receiver or order correlating receiver and it is a very very interesting result because we what did we say was the drawback of the decorrelating receiver, the decorrelating receiver we said may result in a noise enhancement. Now if the noise enhancement is noise itself is small or negligible to begin with then we can ignore the effect of noise enhancement and say that the decorrelating receiver is the best possible solution and that and the MMSE receiver more or less confirms that that is the best way to do that.

Now, of course, we take the other extreme where the special case number two; special case number two says that N_0 is large. So, this is the case where the decorrelating receiver will not do a good job because it will end up further magnifying the noise. Now notice that if you were to have this expression N_0 becomes large. So, therefore, we between r and $N_0 D^{-1}$ the second term dominates. So, under this condition we find that the expression for A can be written in the following manner it is given by $E^{-1} N_0$, $E^{-2} N_0$ after D^{-1} .

So, basically notice that there are two inverses given and finally, it is $E^{-k} N_0$ notice what is this telling us remember our best estimate for the transmitted vector is $b = A^{-1} r$. So, what the MMSE receiver is telling us is that when N_0 is a substantial when N_0 is significant noise component is present, then the best that we can do is make a decision directly on the output of the correlators directly on the output of the. Because A is nothing, but A diagonal scale in matrix, it says b_1 based on the output of r_1 b_2 based on the output of r_2 because any attempt to do the suppression of the multi user interference ends up further degrading the noise and not giving us any additional benefit.

So, basically this is both the special cases one where N_0 is small and the one where N_0 is large both of us both of them give us a good explanation of why the MMSE receiver works in both these environments. So, to summarize MMSE receiver is

one where we write down the expression, that you want the output or the decision to be a linear combination of the output of the correlators and the expression for the matrix A is given by the following expression, and this is the MMSE receiver for multi user detection. An the decision statistic itself is given by A times r where r is the output of the k correlators.

So, we get a received correlators and we get a set of a received get a received signal passes through the edge of the k correlators match filter to each of the case spreading waveforms, you get the vectors r_1 through r_k uppercase k you are able to pre compute the matrix the matrix a do a times r and then do the decisions based on that. So, to go back to the figure that we drew, so, CDMA multiuser detection it is a very vast problem widely studied.

What we have try to do is to give a flavour for what are some of the aspects the first one is the optimal solution requires you to do a exhaustive search of all combinations of the transmitted vector. The sub optimal solution the one that we saw first was a decorrelating receiver. A decorrelating receiver is one that computes the correlation matrix inverse of the correlation matrix, and then makes a decision based on that. So, this would be R inverse times r is the decision made by a decorrelating receiver.

On the other hand the MMSE receiver says we would it reduce the matrix r it also uses the an estimate of the noise in the system. Noise times the D inverse, D representing a diagonal element matrix with the energies of the different transmitted signals inverse. So, notice that when N naught is large N naught is small it convergence to the decorrelating receiver, and when N naught is large gives you a solution which is different from the decorrelating the receiver and which turns out to be the best in. So, anything in between where you do its somewhere which is neither the very in low noise or the very high noise anything in between we find that the MMSE receiver is a good solution for us and we get the decision statistic \hat{b} to be equal to A times where a is given by this matrix. So, that summarizes for us the CDMA multiuser detection and also concludes our discussion on the corner on the chapter of CDMA.

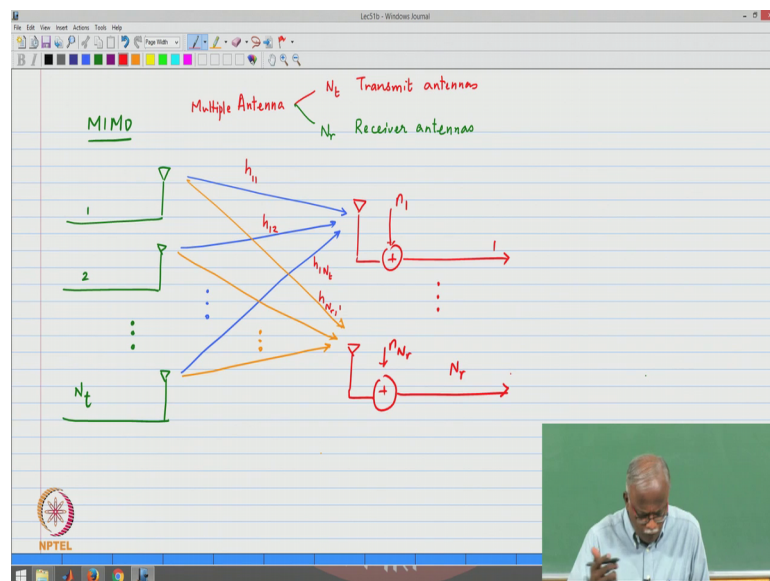
I hope you will have a chance to read the corresponding text there the material both from each of the books the coverage that is presented in each of the books adds a unique flavor, definitely I would encourage you to read the corresponding chapters in reports

book also in Haykens book Molisch and in goldsmith. Each of those adds a different perspective and of course, the multiuser detection part is very nicely covered in throw it is. So, general CDMA the concepts of CDMA in from each of these books and the multi user detection very specifically the decorrelating receiver the MMSE solution would the best a good reference would be a Proakis digital communications.

So, I hope you will be able to give get a good understanding and confidence of that, because CDMA is the backbone of our 3G systems and our 3.5G systems. So, whether we are talking about CDMA 2000 or talking about wideband CDMA or any of the enhancements such as HSPA these are all based on these and our understanding of CDMA, and how we achieve capacity how does a base station do multiuser detection and all of the benefits and the and the complexities of a CDMA system. So, that would be a good summary of that.

Now, having completed the discussion on the multiuser detection, we would know like to move into the next big aspect of the wireless systems where we have the benefit of multiple antennas both at the transmitter and receiver. Now we find that in the fifth fourth generation and in the first generation as we presented in the introduction the use of multiple antennas has become a very significant and a very prominent feature, and that is something that we want to gain a very good understanding of from the point of view of the course.

(Refer Slide Time: 48:24)



So, with that in mind we would now like to look at the formulation of the multiple antenna environment. So, multiple antennas at the multiple antennas and transmitter receiver which it is been characterized in terms of there are N_t transmit antennas. So, N_t antennas at the transmit and transmitter and correspondingly we have N_r antennas at the receiver. So, N_r receiver antennas.

Now, there is no assumption being made as to which is a smaller or larger N_t can be greater than N_r it can be less than N_r or can be equal to N_r . All three combinations are possible and we will mention the various scenarios that are likely to happen. So, in the most general framework what we find is that we can look at the transmit side as a set of N_t transmitters, which is transmitting to a set of N_r received antennas. Now for a moment take the case where you just have one received antenna.

If you have only one received antenna what you will find is the transmitted signal from antenna one which goes on the first branch, let us call this gain as h_{11} . So, I would like to call this particular gain as h_{11} . So, that is at the signal gain from at received antenna 1 from transmit antenna 1, similarly if you look at the. So, the receive antenna 1 from transmit antenna 2 call it as h_{12} and h_{1, N_t} . So, basically there are three keep in mind we are looking at the case where there is only one receive antenna.

So, we get three different signals that are received by antenna one each of them having a complex gain denoted by h_{11} , h_{12} , h_{1, N_t} . Now if you were to write down the from the other side from let us consider only one transmit antenna and N_r receive antennas.

So, the first antenna transmits its picked up by antenna one that would be h_{11} , now the signal picked up by the antenna N_r would be h times N_r comma 1. So, that would be the h the receive. So, like that each of those receive antennas would be picking up antenna from the signal from antenna 1. So, this is a important framework of characterization. So, when we have the general case of N_t antennas at the transmitter N_r antennas at the receiver.

Then we can write the write the expression for the MIMO system in the following way.

(Refer Slide Time: 51:40)

$$y_i[n] = h_{i1}[n]x_1[n] + h_{i2}[n]x_2[n] + \dots + h_{i,N_t}[n]x_{N_t}[n] + n_i[n]$$

$$\begin{bmatrix} y_1[n] \\ y_2[n] \\ \vdots \\ y_{N_r}[n] \end{bmatrix} = \begin{bmatrix} h_{11}[n] & h_{12}[n] & \dots & h_{1,N_t}[n] \\ h_{N_r,1}[n] & h_{N_r,2}[n] & \dots & h_{N_r,N_t}[n] \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] \\ \vdots \\ x_{N_t}[n] \end{bmatrix} + \begin{bmatrix} n_1[n] \\ n_2[n] \\ \vdots \\ n_{N_r}[n] \end{bmatrix}$$

$N_r \times 1$ $N_r \times N_t$ $N_t \times 1$ $N_r \times 1$

$y = Hx + n$ Given Y known H Estimate X

So, if you were to write down the output of the at the receive antennas y_1 of n , which is the first receive antenna this would be h_{11} of n times x of $n \times 1$ of n the first transmit signal from the first transmit antenna, then h_{12} from the h_{12} of n n basically shows that these are all time varying quantities the channel can be a time varying quantity.

So, h_{12} of n times x_2 of n plus dot dot dot h_{1, N_t} of n times x of N_t of n . Notice that this would be the expression for the signal received by antenna 1 and as in all of the communication systems each antenna adds the presence of thermal noise is reflected by a noise component which is given by n_1 of n . So, if you were to capture this in matrix form it would be y_1 of n , y_2 of n dot dot dot y_{N_r} of n notice the index is a N_r because there are N_r receive antennas.

So, this would be N_r cross one vector, the vector in the middle the channel a matrix would be h_{11} of n h_{12} of n , all the way to h_{1, N_t} of n , the last term will be $h_{N_r, 1}$ of n $h_{N_r, 2}$ of n dot dot dot last term will be h_{N_r, N_t} of n and this multiplied by the signal transmitted by each of the transmit antennas that would be x_1 of n , x_2 of n all the way to x_{N_t} of n and each of the receive antennas has its noise component represented by n_1 of n , n_2 of n all the way to n_{N_r} of n .

Let us quickly write down the dimensionalities of each of the vectors the transfer matrix is N_r cross N_t the transmit vector has got N_t components. So, this is N_t cross 1 and the last the noise sources is N_r cross one notice that it will be dimensionally consistent and

we can write it in the following form, y is equal to h times x plus n in compact form. So, this is the problem or the formulation of a MIMO system there are N_t transmit antennas each of them transmitting a signal N_r received antennas picking up the transmissions from each of the N_t transmitters and of course, the noise terms being added in each of the receive antennas.

So, in a compact form we can write it as y is equal to $h x$ plus n . So, the MIMO problem statement is that, we are given the observation y ; given the vector y we assume that the channel matrix is known H . The task before us is to estimate x that is the MIMO problem statement and that is something that we will pick up in the next lecture and build on. So, that we can get a good understanding of how MIMO systems work, and what are the advantages and the benefits of using MIMO in a wireless system, and very specifically in a cellular type environment.

Thank you very much.