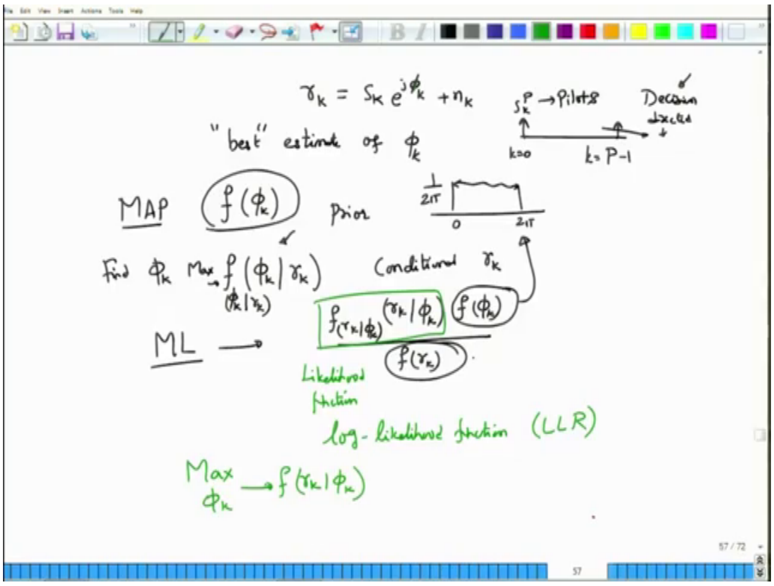
Fiber - Optics Communication Systems and Techniques
Prof. Pradeep Kumar K
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 58
DSP algorithms for Carrier phase estimation – II

Hello and welcome to NPTEL MOOC on Fiber Optic Communication Systems and Techniques. In this module we will look at couple of Algorithm widely use algorithm to estimate the phase phi k ok.

(Refer Slide Time: 00:28)



The problem we have already discussed. There a few samples are r k, which are given in these manner right. When there no noise and if you know S k, then it is easier for you to find phi k keeping in mind the phase unwrapping problem. If you do not know S k, then there are problems in (Refer Time: 00:44) this algorithm because, the best algorithm can be applied when this parameter f k or the symbol that you have transmitted is actually known to you.

In practice of course, you send beginning at k equal to 0 or whatever the time instant there going to send a certain number of pilot sequences. So, for k equal to 3 minus 1 all the symbol will be known So, I am going to write the known symbols as S k with P at the top here indicating that this is the pilot symbols.

So, this pilots are essentially training symbol which are meant for the receiver to use them to obtain the initial. So, starting at the receiver when it is blind you know it just started it does not know what to do with respect to the same. So, you send assist the receiver by sending this pilot, so that it can use them and build up the value of phi k and from then onwards it can actually going to what is called as decision directed mode ok.

So, in this mode you can actually track the phase by switching over to this mode and if your pilot symbols are well chosen if that channel is reasonable, alright. Then you will suffer very small change in performance when you go from pilot or training mode to the decision directed mode. Anyway these are something that we will not be getting detailed in this course, but you can look at other course from communication to understand what we mean by decision directed. So, far now we will assume that pilots are move towards and then we want to estimate phi k right.

Now, one of the standard ways of doing this estimation is to perform what is called as maximum a posteriori estimation. I do have some initial knowledge on the phase. If i do not have any knowledge, I can always think of this phase as any random variable between 0 to 2 pi which means it actually has a uniform distribution 1 by 2 pi.

So, this distribution that I have in mind there is a receiver that knows is called as a prior distribution because, before receiving any samples I know what this is; I mean I just I could before using any sample this is all the information that I can have about the phase phi k. So, that is why this is called the prior information, but as I start receiving the sample, I can update my knowledge of phi k right.

So, in fact, I can find out what is called as the posteriori distribution that is after I have received the samples, there will be a probability density function. This will be conditional and this conditional probability distribution or density function if conditional upon receiving r k. So, if you do not receive in a particular plot, if you do not receive r k. Obviously, you are not going to form or able to form this conditional probability density function but please remember this is conditional probability density function of phi k given that you are actually received r k.

And this map criteria tells you that if find this phi k such that, this function is maximized. So, you can find that phi k which maximizes this function which of course, puts us right into the problem in the sense that I do not know what value of phi k will actually match or maximized it right. So, what will it maximizes I do not know, so I need to know this one.

And clearly this is where the difference between actual practice and modeling comes in. All these functions that we are going to write are good for analytical theory, but in practice there only very good approximation, they are not exactly the same. So, your you do not get n k 2 b and IID sequence, you do not get this full Gaussian kind of a distribution all the time especially, when the sample sizes are small but nevertheless you need some handle to begin this algorithms right. And you need some sort of a performance comparison.

So, if these conditions are held what is the best estimate of phi k and in practice how you derived from that phi k is entirely dependent on the actual channel ok. If your algorithm is good and robust then the difference between what you get in theory, or what you compute in the numerical way would be more or less same as what you are going to get in the actual practice.

So, with that in mind; So we would like to find out what is the conditional priority density function of phi k and then maximized that one, but because I am assuming any value of phi k is possible right, I do not have to worry about this maximum a posteriori probability, but I can simply worry about what is called as maximum likelihood ok.

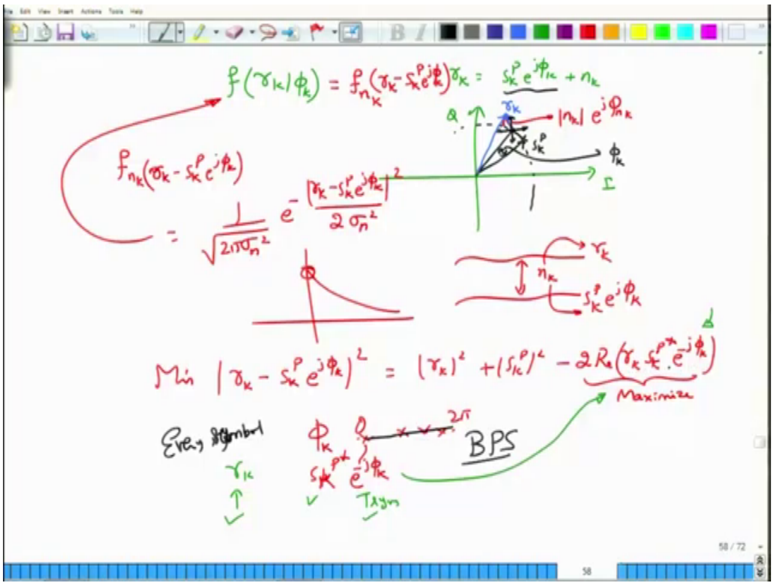
Where do I get this maximum likelihood? Maximum likelihood is essentially obtain by first going into or first using base theorem and transforming this conditional pdf of phi k given r k into r k given phi k ok, so I am transforming this pdf into this manner. Of course, need to multiply by phi k and then divide this entire thing by f of r k. This f of r k is the pdf of receiving the symbol and it does not really have any statistical nature do it; so it just receiving probability. So, you do not really worry about this one.

And because you said that f of phi k that is conditional probability density or rather the prior called distribution function or the density function in uniform any value of phi is allowed between the 0 to 2 pi; the thing that we are going to maximize is this particular probability. And this probability is called as likelihood function ok. So although, this is a pdf called as likelihood function frequently you will find out that people actually use what is called as log likelihood functions ok.

So, this is sometimes called as log likelihood ratio when phi k can take on only discrete values, but here phi k is not discrete value therefore, I would not want to use the reward ratio here, but you know this is the standard terminology. So, we may as well keep that in mind ok.

So, I am going to use this likelihood function and my rule of finding the best estimate of phi k is the one, which will actually maximize this probability distribution or density function. So, find the phi k such that, it maximizes this particular function ok, so this is called as maximum likelihood estimation and this is what we are going to do.

(Refer Slide Time: 07:01)



So, what is the conditional distribution function of r k given phi k. Well, if r k is given by S k e to the power j phi k plus n k, first at of think about this in terms of a vector ok. So, in the phase and quadrature space that I have this r k n k are all going to be vector. So, first let us assume that I have transmitted this symbol which is say S k p. So that is the vector here, and then because of the phase noise this vector has moved and become e to the power j phi k this is the phi k that you are actually estimating.

So, this difference is the phi k that you are estimating and please note that by multiplying S k with e power j phi k, you are only going to move it in a given circle. So, this is actually a circle and you are moved a certain distance depending on what is the actual value of phi k that have been realized at this particular sampling in trivial. Fair enough, so we have this at the point ok.

Now there is a noise. Noise could actually be added in any direction, meaning that it has both in phase and quadrature component and therefore, its amplitude could be some value which is a magnitude, but then the phase could be anywhere up there. So, after you know assuming that the noises as realised in this particular way, the width around this point is actually determined by the variance of the noise. This is what the actual received symbol here for you.

So, this is the received symbol r k after taking noise into account and this little red line here which I am redrawing is your noise vector n k. Please note that noise is also complex for you it can also have a magnitude and it also has a phase phi n k ok.

So, this is what the received vector is right, after (Refer Time: 08:49) phi k and n k are reasonably within their limits. Then I can actually look at what is r k find out what is S k, I could look at the distance, and then be reasonable assure that I am in the right decision space ok.

But in case the phase phi k is larger than this received vector r k would go into a different quadrant and then you may not be correctly estimating the phase. Because, if you estimate the phase to be a smaller and different then it will not bring it back to the actual constellation that you have transmitted. And therefore, you will start making error. Anyway, so this is the vector picture of this one, keep the vector picture in mind you are going to return to the vector picture shortly after we write this equation.

So, if you go back to the real kind of a case, if this is r k that you have received and this is the S k e power j phi k S k p e power j phi k that you have transmitted, the difference in the complex domain of this vector n k is the amount of noise that has added. And what is the probability of generating this noise if that noise probability dependent on r k? No is a noise probability dependent on S k p that is also no. So, this can be reconverted into unconditional probability density function of the noise alone of being certain amplitude which is given by r k S k p e power j phi k.

So, this is how you can actually get the likelihood function express purely in terms of the noise ok. And we know what is the Gaussian noise with an amplitude factor of r k minus S k p e power j phi k going to be right. This is equal to 1 by 2 pi sigma n square under root e to the power minus r k minus S k p e power j phi k magnitude square. Please remember this is the 0 mean process. Therefore, the amplitude magnitude square divided by 2 sigma n square is what will be the probability density function that I am interested in.

So, this probability density function is actually the likelihood function that you now have and you want to maximize this likelihood function. How can you maximize the likelihood function? Look at the exponential function, the exponential function will be maximum when this argument is actually 0. So, which means that if you can maximize this or rather minimise this difference right.

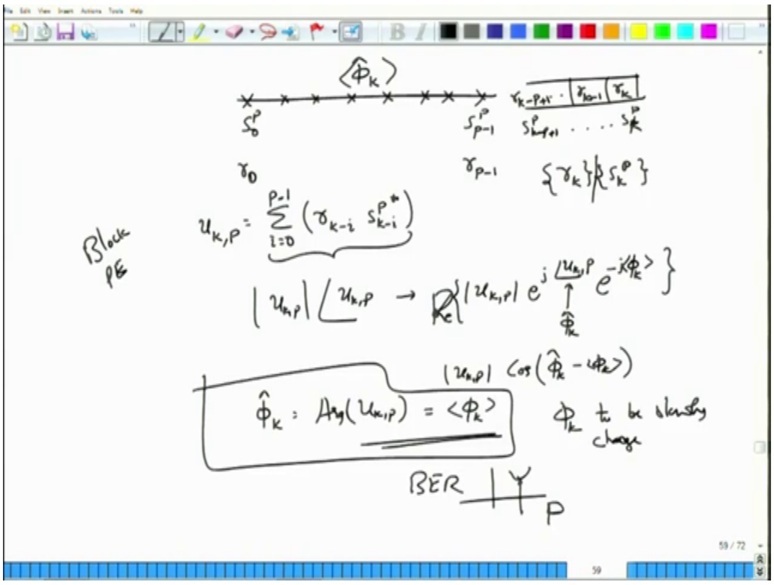
So, when you make this difference to be as small as possible then that would maximize this exponential function. Alternatively you could actually expand this out and then get r k square plus S k p magnitude square because, clearly e power j phi k magnitude will be this one. So, two times real part of r k S k p conjugate e power minus j phi k. And then you want to minimize this fellow right, minimising this means you want to maximize this quantity right.

So, you can actually come down to simple correlation of your received sample with the pilot sequence multiplied by an unknown e power minus j phi k and that would actually give you the total or estimated phase phi k ok. So, that is what you can actually obtain. So, that particular phi k, you may not know this, but you know what is this one right you know what is this S k p.

So, you try out various values of phi k. So, over 0 to 2 phi to try out some 10 samples up there and then when you multiply each of those samples by e power minus j phi k after conjugating the known sample S k. So, this is known to you, this is what you are actually trying, so therefore, this also known to you. r k is what you have received. Therefore, this is also known to you and then you take the product of these guys to obtain the correlation, take the real part of it and whatever that would maximize this product will be the estimate phi k ok.

So, this seems to be the method for going for a single symbol case.

(Refer Slide Time: 12:53)



But in case you actually receive let us say, over p training symbols as we called, so this is S 0 p to S p minus 1 p. So, these are the received symbol sequence that you have obtain and assuming that the phase is going to be constant right, then you also have a corresponding r 0 to r p minus one, then what you can do is you can take r k minus i S k minus i p conjugate and over this range the phase is actually constant; that is your actually looking for the average phase corresponding to this one.

So, this is the sorry this is the average phase phi k, so I have the average value which is independent of or which is constant over the entire symbol. And then you sum this over I equals 0 to p minus 1 and what you get is this particular quantity, which is the inner product of the received sample sequence with that of the training sequence right.

So, with that of the training sequence if you take the inner product of these two are the dot product of these two then, the argument of this, so if you actually look at this way, this is the inner product. So, I can express this inner product as its magnitude, I will call this sequence as say without this phi k, I will call this sequence as U k comma P, P standing for the length of the sequence K standing for the sequence at the K th time Of course, as K increases you have to keep moving this block of p symbol away from each other.

So, I have U K comma P and then argument of this UK comma P, which is the inner product of the received sample sequence with SKP. So, I can even write this in this usual manner UKP magnitude e power j angle of UK comma P multiplied by e power minus j phi k average is what you would have obtained as the exponential function of therefore, you take the real part of this and then call this one as the estimate phi k hat. So, clearly when you take the real part of it magnitude will be constant, so that would simply come this come around here and then you have cos phi k hat which is the estimate minus phi k average right.

So, the estimate phi k that you have obtained, which is basically the argument of UK comma P, will be the best estimate for the average phi k of the phase noise at that particular K th fabric time. So, if you now move on to say k plus 1, then you need to move this block. So, you actually have received symbols, you go back in time, for this is r k r k minus 1 and you go all the way up to r k minus p plus 1. So, that is the length over which were actually looking at the received sample. And then if you know these values SK I mean S P minus 1 of K or rather S; so same value, so it is actually going to be S K minus P plus 1, the pilot sequence if you know all these pilot sequences then you can take the inner product or generate the vector out there and then the argument of that inner product will be the average phase that you are actually looking for.

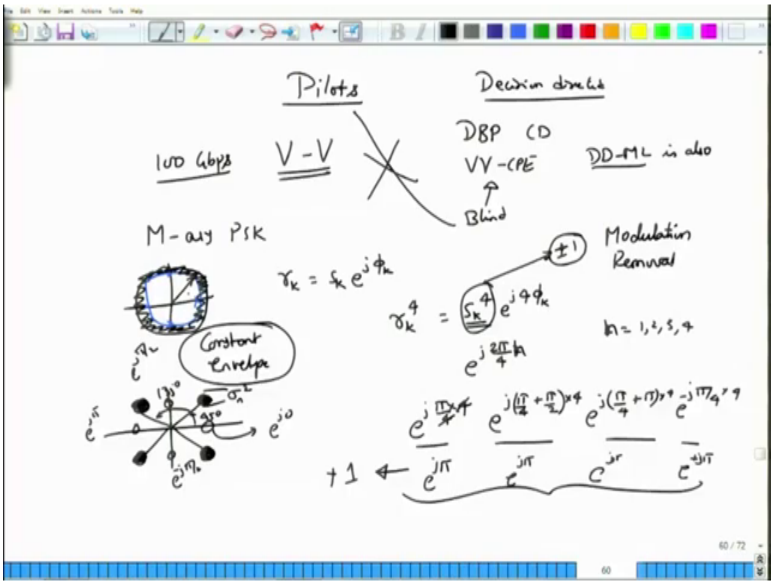
So, this is more implementable version than the previous one. The previous version was called as blind phase search because, in this case you did not know what phi k was, but you assume that phi k could be change in every time. Therefore, every symbol that you obtain you actually can do a blind phase search by stepping through all values of 0 to 2 pi in appropriate intervals that you have and then maximizing this particular inner product, but when you extend this into this form this is called as block phase estimation because, you are actually looking at a block of p symbols and when you know all the pilots in that particular block you take the inner product of the received symbols with the pilot symbol and then call this as your estimate.

Please note that, in this case we have assumed phi k to be slowly changing. So, which means that this is reasonably to apply to the case of you know high bitrates systems. The performance metric that you are looking at will be again the BER and that will actually depend on the sample or the block length P ok. So, there will be some optimum value sorry BER is actually going to decrease and increase. So, there is going to be some optimum value of P for which BER is kind of minimum and then you can work with that that would be the optimum block length.

Too large block length means you are actually assuming that many many symbols have the same phase which is absolutely not but too small link may not satisfy. Too small block length may not give you that I know statistical fluctuation that you are looking for noise will not be there to average it out. And therefore, inner performance will suffer. There is an optimum block length which lies somewhere between two small p and two large P which will give you best phase estimate phi k.

You can now see that there is been a problem in all our technique that we have discussed.

(Refer Slide Time: 18:00)



The problem is that I should know the pilots. Now if someone does not give me pilots, so after the pilots have exhausted, I can either do this decision directed mode or i have to look for other methods when pilots are not at all available ok. There is an important technique, which has become the benchmark technique called as wider V wider V algorithm, just as DBP is the benchmark technique for CD VV method VV CPE has become the method for phase noise.

Sometimes the decision directed maximum likelihood that I just describe previously is also used, but this one is much simpler and this is a method which we call as blind because, it does not require any pilot. The idea of this one is very simple. See first assume that there is no noise, so you have s k e power j phi k. Please note that I am not writing this p here because, I do not need to know the pilot here, I mean there is nothing I need to know.

Accept I need to have a constellation which is of the form M ary PSK, of which the simplest one is q p s k. You could either have symbols arrange in 45 and differences of 90 degree again that that way or you can actually have these symbols write the ones that are shown in blue but nevertheless all of this actually fall on the same circle, although I am not shown you the correct circle they all fall on the same circle, which means they all have the same magnitude on that one. So, these are also called as constant envelope signals.

Of course, noise will be present, so what would noise to noise would actually take this amplitude and make it lie anywhere in between these values right. So, it could actually make a bulge region around that one. And when you have finished looking at the phase and estimating the phase correctly, you are going to obtain this kind of a constellation. So, these are all going to be circle. So, the signals could lie anywhere in between, but because there reasonably separated the BER will be considerably less in this particular case.

And these circles actually are the variances are determine the determined by the variance of the noise. Anyway, so the point is that even ideal system the envelope of the q p s k signals will be constant and because q p s k is used currently in 100 Gbps system, this technique is very important for us to learn. The technique is very simple. Assume initially that there is no noise.

Now what I do is; I take this r k, raise it to the power 4 right. In general I will rise this 1 to the power m, but for now for the q p S k I will raise it to the power 4. Once I have raised it our raised this r k to the power of 4, what I have actually done. On the right hand side is average this one to the power of 4, e power j phi k to the power 4 is basically 4 phi k.

Now, what is S k power 4 right. If you look at this one, this is all given by e power j 2 pi by 4 right. So k, k will be n. So, n will be say 0, 1, 2 and 3 when you actually look for, so 0 that would be at 0, so n it is 1, no this is actually going to be 2 pi by I have to create this point right. So, this is 45 degrees and therefore, I need n equal to 1 to begin with and then when n is equal to 2.

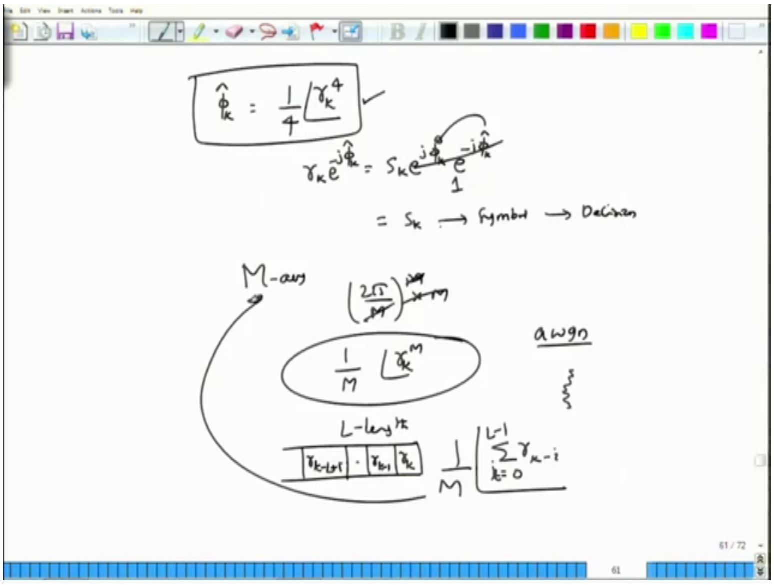
Anyway, so these are the 45 degrees and then 135 degrees and so on but when I take them to be the power 4, this is n equals 1, 2, 3 and 4 let us say these are fine, that when I take the first one is e power j pi by 4. The second one is e power j pi by 4 plus pi by 2 right and then e power j pi by 4 plus phi and then you have e power minus j pi by 4.

So, these are the four consolation points that I have and when I raised them to the power 4, I am going to multiply everywhere by 4 ok. So, when I multiply them by 4, 4 and 4 will cancel, 4 and 4 will cancel and then this will be 1 and this will be 2 pi for e power j two pi is 1. So, I will get this as e power j pi, this will be e power j pi, this will be e power minus j pi, this will also be e power j pi, minus j pi is same as plus j pi because, all these values after raising them to 4, will give me 1 or rather minus 1 ok.

And there is not really important, minus 1 or plus 1 is not an important. If you do not wanted to raise to the power 4, you raise it to the power 8 and you get rid of this minus sign and make it plus 1. But the important point is that, you got rid of this s k power 4, I can take this to be a constant or I can take this to be plus or minus 1 depending on what value I have raised it this particular thing and that depends on the kind of consolation. For this consolation raised to the power 8 and for this constellation that is on these axis, which are given by e power j 0, e power j pi by 2, e power j pi, e power j, minus pi by 2 you raised to the power 4 such that, everything will converge back to plus 1 ok.

So, what you have done is without knowing what one of the 4 symbols that I have transmitted, you converted that S k power 4 either into plus 1 or minus 1, but in both cases the effect of modulation has been removed right. So, we have removed modulation effect by simply raising it to the appropriate value of consolation. In this case 4 or in some cases 8, but these are done without knowing what has been transmitted. That is a big different because, if you do not know what you have received and yet you are able to remove the modulation, that is a very good point of that right, but you do not know what is the phase, which is there going to estimate it now.

(Refer Slide Time: 23:32)



How do I estimate it? I know what is r k power 4 that should be equal to e power j 4 phi k, right. Now if I take the argument of this or the angle of this one, then taking the argument essentially means I am going to have 4 phi k here right which is more than what I want, I want phi k not 4 phi k, so I can actually take this 1 by 4 onto this side and then get phi k right.

This is a very simple way of obtaining the phase estimate phi k. So, once I obtain the phase estimate, I am going to go back to the received signal r k, I multiplied by the phase estimate phi k. What is the result of doing this one? So, I had originally S k e power j phi k, which was unknown. When I multiplied this 1 by e power minus j phi k hat and if this estimate and phi k are essentially closed to each other then, the product will be equal to 1 and what you get here will be just S k, which is my symbol that I have and this can go to the decision device right. So, this is as simple as that.

So, this method can be applied for any M ary syste, which have you know which are divided by 2 pi by M angle apart. So, when you raise them to the power M, that is actually equalent of multiplying it to M and then removing this. So, modulation can be removed. And once you remove the modulation, you take 1 by M of the angle at any point of the angle at any sample point r k, which has been raised to the power m.

Now this was without noise. In that case everything works out fine, but what you do with noise? While noise is additive White Gaussian noise, so it has equally a property of being positive and negative. So, what you can do is to average this symbols. Again you can form a block ok. So, you have r k, r k minus 1. Let us say r k minus l plus 1, l length block you can take. Again you do not need to know any of your transmitted sequence; all that have to do is take the average of this. So, r k summation k equal or r k minus i let us say where, i is equal to 0 to l minus 1 right. So, that r k minus l plus 1 here, this is correct.

So, you just take the average then take the argument and then divided by M, M being the M ary phase system that you are looking at. So, this way you can reduce the effect of noise by simple average in process ok. And it turns out that this method which does not require any feedback because, in decision directed method you want a feedback, you want to know what has been detected and use that to estimate. So, this is called as detection plus estimation problem, whereas in this problem you are not looking at any detection part; I mean there is no information coming from the decision device back to the estimation algorithm. All that is coming back or all that is there happening is 1 way propagation or one way flow of information.

So, in that way, I start off with a symbol or a block of length r k, again that length l depends on you know application and again the b e r will be optimum for a given value of l. And then once you have that value over that range l, you simply average the received symbols, then take the average or take the argument of that one divided argument by M. You still have to deal with the phase unwrapping problem, but that anyway have to be dealt with every phase estimation algorithm.

And then once you have done that one, you can decide the symbol by the rotating r k right. So, this method is called as feed forward carrier phase estimation method and if very very popular for the case of simple q p s k type of system, but when you go to a quam system, when it still possible to use this method except now, you have to decide which ring your own right because, in a 16 quam system for example, you will find that there are amplitudes which are of 3 levels.

So, you have three different rings. So, there is usually a ring indicator or intensity indicator. Using that has an information, you can then decide the same kind of an algorithm or you can partition the signal set into those amplitude levels and adjusted slightly and then come up with what is an algorithm called as multi modulus algorithm, which is a generalization of this wider v wider v algorithm. Anyway these finishers are talk or are algorithms about carrier phase estimation. I encourage you to actually work out this algorithm choosing MATLAB or any of your preferred programming choice.

Thank you very much.