**Ultrafast Processes in Chemistry**
**Prof. Anindya Dutta**
**Department of Chemistry**
**Indian Institute of Technology-Bombay**

**Lecture No. 10**
**Data Fitting 2**

**(Refer Slide Time: 00:15)**

Data fitting models

Come back to our discussion of data fitting models. So far we have talked about single exponential decays, multi exponential decays and distribution of lifetimes. So essentially, our entire story can boil down to one theme, do our fluorophores experience a homogeneous environment or a heterogeneous environment. So another way, which I think is not as good as what we discussed in the last module. Another way of handling heterogeneous environment is by using a function that is called stretched exponential.

**(Refer Slide Time: 01:01)**

## Stretched exponential

Deviation from single exponential behavior

$$I(t) = I(0)\exp\left(-\left(\frac{t}{\tau}\right)^{\beta}\right)$$

Heterogeneous environment

I hope the function does not look very weird. In a sense it does because I have used too many brackets. But essentially what we have done is we have taken our good old single exponential function. And we have added an additional exponent and this exponential exponent of exponent. So for which value of beta is it going to a single exponential? Let us say beta can take several values depending on the situation. If beta is 1 what happens if beta is 1, isn't it a single exponential.

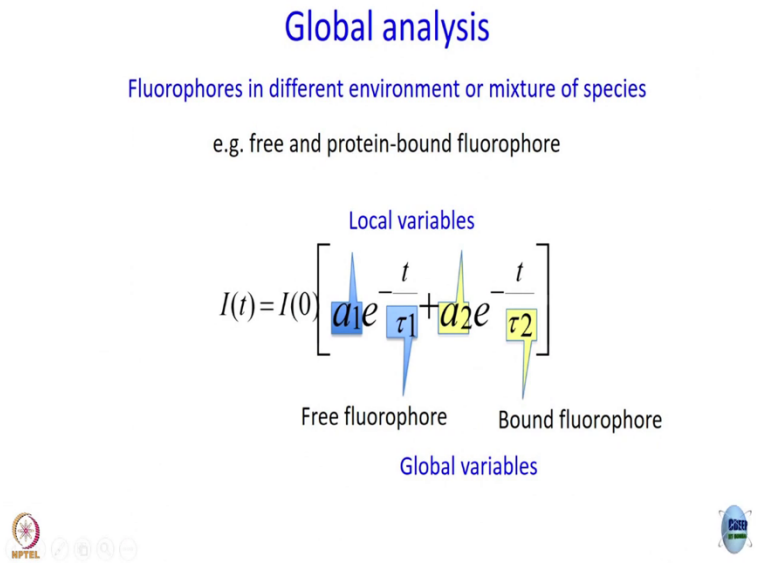
That means it is one homogeneous environment we can think. What if beta is anything other than one will with it means simply means it is not single exponential and some heterogeneity is there. So this is one rather simplistic way of describing heterogeneity to say that it is not homogeneous, stretched exponential. Now, of course, there are cases where people have staged the stretched exponential as well. And you will see papers in which they are fitted the data to a sum of 2 functions. First one of it is an exponential decay, second one is a stretched exponential.

So the idea is good, it means that you have one kind of environment, which is homogeneous. And then another kind of environment which is not homogeneous. And this might actually be a good model, not simply stretched exponential, linear sum of stretched exponential and exponential function. This can actually be a good model. When you have something like once again, since we use the example of Nano particle, let us go back to that. Let us see you nanoparticle where you bandedge emission and you have trapped emission. Bandage emission is more homogeneous we can expect that should be an exponential decay.

But traps that can be of many kinds. The traps are due to dangling bonds and stuff like that. There is nobody has said that all dangling bonds will dangle in exactly the same way. So, it is not very unreasonable to expect that in such a case, your decay will be is linear sum of an exponential function and a stretched exponential function. And once again, it is possible using the all the DAS software to fit your data to something like that the danger is will your program be able to handle it when you make a fitting function too complicated.

Sometimes it becomes too much for the program because your computer does not have eyes does not have ears does not have brain. It has to work on numbers, it has to work on algorithms to give you the correct answer, what it thinks is the correct answer. So if you make it too complicated for it, it might fail. So, but it is not completely impossible thing to do. So, if you want a very simplistic description of a heterogeneous system, then stretched exponential is a simple way to go.

**(Refer Slide Time: 04:47)**




Now, let us talk about something that is more common and very useful. And that is global analysis. We are talking about different environment, heterogeneity, and so on and so forth. In many cases, many such cases, this global analysis turns out to be a very useful tool. Let us think of an experiment where you have a fluorophore it binds to a protein and you are doing a titration. You have the fluorophore solution, you are adding a little bit of protein and you are recording lifetime

what will happen what should be the fitting model assuming that the bound fluorophore experiences a homogeneous environment of one kind.

And free fluorophore experiences a homogeneous environment of another kind? What do we expect what kind of decay. bi exponential is reasonable in this case. So, we expect something like this. bi exponential decay, but this special kind of bi exponential because let us say tau 1 is the characteristic lifetime of the free fluorophore for let us say tau 2 is the characteristic lifetime of bound fluorophore for now, when I do at titration what should happen for all these cases suppose I have 10 sets of data where concentration of fluorophore is same and concentration of protein it increases from zero to some value in all these sets.

Do I expect the tau 1's to be same or different do I expect the tau values to be same across the set or do I expect them to a different way expect them to be something in one case something else in the other. Remember, tau 1 is the free fluorophore lifetime. should it remains same across the set same should it change ideally should remain same? What about tau 2 lifetime of the bound fluorophore when we add protein, once again, it is a 2 state system. What we are saying is the fluorophore is either bound or free.

So, once again even for the bound fluorophore if there is no microheterogeneity then tau 2should remain the same across the set. So, for all these 10 or 20 sets that you have you should get all same values of tau 1, all same values of tau 2. So, these tau 1 and tau 2 these are called Global parameters or global variables which means they have the same value a tau 1is not equal to tau 2. Please do not get me wrong, but all the tau 1 values are the same across the set. All that tau 2 values are the same across the set.

What about the amplitudes will they be same when you did not have any protein there should be no tau 2 actually should be single exponential decay then the amplitude a1 should be 1, a2 it should be zero. And suppose you have a situation where all fluorophores are bound to the protein then a2 should be one a1 should be 0, once again second single exponential in any intermediate situation what will happen? a1 will be something between 1 and 0, a2 will be something between 0 and 1.

And as you increase the concentration of the protein, a1 will go from 1 to 0, it will go from 0 to 4, so, a1 in a 2 the amplitudes these are called local variables. And then the way you do it is that you do not take these decays separately. Of course, you do take them separately to start with, but then when you understand that it appears that this is a fit case for global analysis, then what you do is you take all the decays together and now we know about iterative deconvolution.

What I am saying is in each iteration, all the tau1 values are same all the tau2 values are same. In the next iteration, that one value can be varied, but even then across the set it will remain same. Just to illustrate a little bit, let us say to start with tau1 value of 1 nanosecond tau2 value of 10 nanosecond and I have 10 sets, I do a first round of fitting and then in the next round, I change 1 to 1.1 nanosecond, now it is 1.1 nanoseconds throughout and then I find the optimum value of tau 1 is 1.3 nanosecond.

Now, I started varying tau 2 instead of 10, I used 10.1 only 9.9 and finally, after doing a lot of iterations, I find that tau 2 value of 9.5 gives me the best result. So then, tau 1 is one point, what did I say? tau 1 is 1.5 tau 2 is 9.5. But that is finally all tau 1 values are same across the set all tau 2 values are same across the set. Important thing to understand here is that we are not holding them constant. We are holding tau1 values constant only for particular iterations.

In the next iteration it does get changed. So, we are optimizing them as well. It is just that we are optimizing them in a coordinated manner. But in every iteration, tau 1 values throughout the same for all the flow the set now, homework is it possible? Can you think of a situation where tau 1 and tau 2 are local parameters and a1 and a2 are global parameters, we can actually have that as well. Can you think of something like that? Let me give you one example. Let us say I have fluorophore that is bound to a protein again.

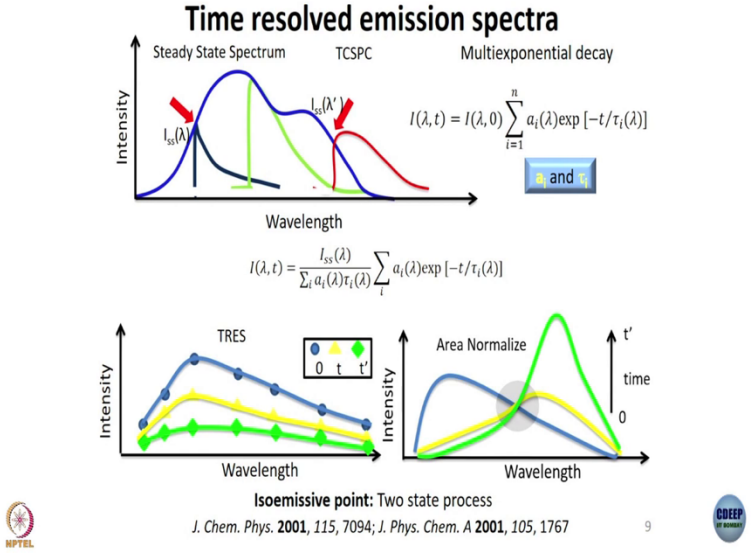
So you have today we are obsessed with nanoparticles and proteins. So, let's say your fluorophore that is partially bound to protein. And let us say I add iodide to the system. What will happen iodide is a good quencher, fluorophore will get quenched, but which what kind of fluorophore will get quenched? iodide does not get inside protein. So only free fluorophore will get quenched, so now suppose you have this situation where 50% of the fluorophore is bound. 50% is free. Now

you keep on adding iodide for all the different iodide concentrations what will happen unless it disturbs equilibrium 50% of fluorophore is still remain bound 50% will still remain free. So, that is what is given away amplitude. amplitudes will be actually fixed throughout tglobal parameters.

What will change lifetime of the bound fraction will also not change that is also double parameter, lifetime of the free fluorophore will keep on decreasing as you keep adding iodide. is that so, that will be a local parameter for different concentrations of iodide with lifetime of the free fluorophore for usually the shorter component will keep on decreasing lifetime of bound fluorophore before should not change amplitudes also should not change.

So, do not think that for all cases of global analysis, lifetimes are the global parameters and amplitudes a local parameters, not necessarily. So, global analysis something that is often very useful for us, if our system is like that.

**(Refer Slide Time: 13:10)**




Next, we move on to something which is a little different and perhaps that is why the title is in a different color time reserved emission spectra see you excite a molecule and then in the excited state it evolves into something else a different state, your locally excited state and then due to excited state process, it goes over to another new state emission spectrum should change. How do I see well of course in steady state, you might see a stoke shift and all or suppose I want to see the dynamics? I want to record the fluorescence spectrum at different times after excitation.

How do I do it? If you have an instrument called streak camera then you can see it in real time. Perhaps next day we are going to discuss streak camera briefly. But if you do not have a streak camera suppose you only we have TCSPC, can you still construct the time resolved emission spectrum and can you work out how it evolves in time? This is how you do it, let us say this is a steady state spectrum.

So, I have intentionally drawn 2 bands, because this higher energy band, lower wavelength that let us is the locally excited state. And this is some state that is formed as a result of some extensive process. Now, what I do is I record the fluorescence decays at different wavelengths across the spectrum, the more the better but with good enough bandwidth if you are going to open the slit of your monochromator to say 20 nanometer and then you are going to make you are going to record the case in 1 nanometer intervals, it makes no sense.

So if we are going to record decays at 5 nanometer interval, you should have a bandpass of 2 nanometer no more. So, how many decays you will be able to record across the spectrum actually depends on how strong the fluorescence is, how good a detector you have. And remember bandpass has to be such that you have good enough resolution so that whatever you see, after doing this analysis is believable. Poor band pass is going to mess up this kind of experiment completely. So you record these decays.

Now, I would like to remember something let us see I fit these to a good old multi exponential function. I hope you remember what this I at zero is, the only thing I have added here is I added lambda because we are recording decays at different emission wavelengths. But I hope you remember that there is a relationship between intensity at times zero, and intensity of steady state. So we put that because the problem is I do not know what I 0 is generally, when you do a time correlated single photon counting as you might have seen, when we did the lab session, you record up to 5000 or 10,000 counts or 20,000 counts.

So everything seems to have the same I at times zero, which is not really the correct case because you have recorded for different times. Even if you record for the same time you record all the case

for 5 minutes or 10 minutes or 1 hour, whatever even then, it is not possible to read off I 0 with any accuracy from the raw data because do not forget that what you see is convoluted data instrument function is convoluted with the decay and especially at initial times, whatever intensity you actually see is convoluted intensity until unless you deconvolute it makes no sense that is why it is better to record the steady state spectrum and use this expression and substitute I 0 by I ss divided by sum over I ai tau i. So, this way what do you get?

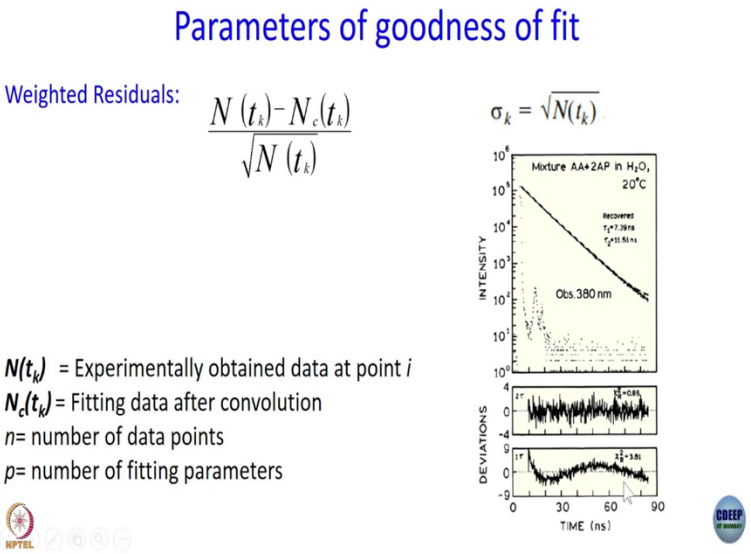
You get the intensity of fluorescence at any emission wavelength lambda at any time t after excitation. Now, what do you do? What for any wavelength let us say these other times zero intensities. It is then for another wavelength that is say these are the times zero time t and time dash intensities this for another wavelength and so on and so forth. So, now, if you join all the points at times 0 you know you get the time 0 spectrum, you join all the points at time t, you get the emission spectrum at time t, join all the points at time t dash you get the emission spectrum at time d dash of course from this figure.

I hope it is not difficult to understand that more points you get better it is and you can get more points only when you use a narrow enough bandpass how narrow bandpass you will be able to use depends completely on your system and your instrument. So these are the different factors that contribute and many times what we do is we want to area normalize it. So, when you do area normalization what are you normalizing to actually, what is the area under the graph of emission spectrum, area under the curve is the total number of photons emitted, but area is proportional to total number of photons emitted.

So, when you area normalize, you are really looking at spectra under the equal number of photons emitted condition. And then it is area normalization spectra also tell you some story which will leave for another day. Good thing of this approach is that it at if you are using the wrong model. Remember, in our earlier discussion, we have said that you have to break your head and work with the absolutely correct model. If you want information about non radiative rate constant and so on and so forth.

But in this case, all you care about is the correct value of time resolved fluorescence intensity. Even if your model is not accurate, it is fine as long as you get a good fit. So this is why sometimes it is better to work with time resolved emission spectrum. Because then first of all, you get to see how the emission spectrum is moving with time with evolving with time. And secondly, you really do not have to worry yourself about whether the model you are using the data fitting you are doing is at all correct. So these are the models of fitting. Now we come to the question, how does the computer know?

**(Refer Slide Time: 20:50)**




How does the computer know whether the fit is good or not? And as I said earlier, a computer cannot know anything, Only the number associated with it so, that is where we use parameters of goodness of fit those who have studied regression we perhaps know how to draw the correct line through experimentally observed points. In good old days when we used to draw graphs and graph paper, we had to do this calculation manually use a ruler and actually draw now, when you fit your data to a straight line or any function polynomial or whatever, your computer actually does this it tries to see how good the fit is by looking at some number and the easiest number you can think of is standard deviation.

You have some experimental points, you draw fitting curve, if standard deviation is small, then you have a good fit. The problem is how small is small how good is good fortunately, when you do photon counting, it becomes a little easier to answer that question. So, the parameter of

goodness of fit that is used is reduced chi square. Now, the expression that you see right now is not reduced chi square completely is on the way to reduce chi square when this chi squared.

Let us see what we have here. What is the denominator? sigma k squared sigma, the standard deviation, square or standard deviation variance and what is all this in a N tk and Nc tk in a N tk is the experimentally observed data at point tk actually I do not know why I have written I, it is not i it is tk and instead Nc tk is the fitting data after convolution at the same time, tk what do we have in the numerator?

The difference between the experimental value and the fit, and you have taken square of that what is the denominator sigma k square? What should they should be? Forget the summation for the moment just this fraction in N tk – Nc tk whole square divided by sigma k squared. What should it be for a good fit? Denominator, I am not said what the denominator is numerator the square root of N at tk where did that come from? That came from because of the noise model, where you do photon counting, the noise model is poisson and the noise is square root of count.

So this is well known this you can find even in say banwell's book chapter one and that is the theoretical limit to the noise and you cannot have less noise than this. You cannot have less variation this so, here in this expression of chi square in the denominator, your sigma k squared you can think that is the theoretical error best possible theoretical error in the numerator in a N tk – Nc tk whole square, you can see it is the actual experimental error.

So, taking a ratio of experimental error and theoretical error when a good fit what should this issue be multiple choice question 0.1, 1, 10, 100. One right it should be as close to the theoretical error as possible. But then so far we have neglected this summation altogether. It is not good enough to look at one point, is not it? You should look at all the points when you sum over all the points. What does this become for every point this ratio should be ideally one and you are summing by a 'n' number of points small n, what should you get? You should get n.

Then, but it is a little tedious to keep remembering in all the time. So, what you do is you divide by n - p. So, this is what it is chi square is some over k equal to 1 to n at N tk – Nc tk whole square

divided by N tk, where does this denominator in come from here sigma k equal to square root of N tk square of sigma k. So, this square root is gone. So, now, when you take this chi square and divide by n minus p, where small n is a number of data points and p is a number of floating parameters.

what is the meaning of floating parameters supposed you have bi exponential decay then what is a number of parameters a1 tau1 and tau2? I did not say a2 because a2 is just 1- a1. Now see what is the value of small n? typically in TCSPC experiment, what would the value of small n be how many points would you have at least 500 if I leave the choice to you 16,000 because you always work at 7 ps per channel.

And sometimes 60,000 might be required 500 at least and what will be the number of floating points unless you are fitting 200 exponential something that floating points will be small 2,3,4. So, what is say 1000 - 3 practically 1000 a little less than 1000? So, now, if you see this expression of chi r square, reduced chi square, it is chi square divided by small n - p, what is the value of chi square value of chi squared is about n for a good fit . denominator is also approximately n.

And so I should n by n be 1. So, for a good fit, reduce chi square should be, it is never equal to one should be close to one. So I would say anything within 1.1 is good, but 1 point that is important to remember and this is something that is not followed carefully by practitioners of this field is that it should be close to one, but it should be greater than one for 2 reasons. First of all, when you take this ratio, you have experimental error in the numerator, theoretical error in the denominator.

Experimental error can never be less than the theoretical limit of error. If that were the case, then you will be doing better than the best possible unphysical. Moreover, the denominator here when you work out reduce chi squared denominator is not exactly n, but it is n minus p numerator is close to n. So, numerator should be a little more than the denominator here also, that I mean as first approximation we can neglect but actually we cannot 1000 minus three is not 1000 it is 997. So, chi square would better be more than one.

There are plenty of reports in literature and I am going to show you one after this. And in fact, even an established textbook, it is often said that something between point 9 to 1.1 is it is not 1.1 is okay .99. or 98 is not, because if you write that, then you are saying that your error is less than the theoretical limit, see that you are wrong, or the theoretical limit is wrong. So if you get less than one you would like to fit your data once again.

So this is 1 parameter of goodness of fit reduced chi square there is one more. Well, there are plenty, but commonly used this one more. See, there is a problem with reduced chi square. The problem is you are taking a summation over all data points. So it is possible that error in one side is accidentally offset by error on the other. So a better thing to look at is weighted residual. And weighted residual is N at tk - Nc at tk divided by square root of N at tk.

Does this have anything to do? Does not any relationship with chi square I show you the expression for chi square once again some of our N tk – Nc tk whole square by N tk. So you can think N tk - Nc tk divided by square root of N tk everything squared. And now see what we have there? We have N tk - Nc tk divided by square root of N tk. Now understand what is going on here, if I say this is R at tk, this expression we have for residual if I say this is R value at tk, will you agree with me that this chi square here is sum over k r tk? R tk Square.

So residual and chi square are different ways of looking at the same thing. But the advantage of residual is that this residual value is defined point by point. So what you can do is you can plot the residual as a function of time. And here, of course, I am not happy with this because chi square is reported to be .85, reduce chi square. And I do not believe it, but if you neglect that for the moment and look at the residual, you see, the upper residual is actually good fit.

Because you have even distribution on both sides, and it should be within the limit of four. Too much of distribution is also not good. And in the lower one is definitely not a good fit. Chi square is bad 3.81 but more importantly, you can see where you are going wrong. And when you know where you are going wrong, especially if it is a multi-exponential fit, if you know that you are going wrong in the long time, you can try to play around with the long time.

If you know you are going around in the short time, you can try to play around with the short time constant. So this is why residuals are more helpful than reduced chi square. So reduce chi square may not be enough. Now, one thing I would like to draw your attention to is this. Look at the expression once again weighted residuals. What would happen if I did not have the denominator? I applaud that as well. What would it look like?

Look at the shape and then try to tell me what this residual would look like if I did not divide it by square root of N tk. It would look like a damped oscillation because see what is the denominator N at tk increases in goes down. Unless or something. So as n goes down then what is the error square root of n that also becomes smaller. So if you do not have the denominator, if you only take N at tk minus Nc at tk, you are going to get an damped oscillation you would not be able to work with that.

But if you actually divided by square root of N tk, good thing is, you get this kind of a plot where deviations are now weighted in such a way that you get the same kind of deviation throughout. So it is easier to judge goodness of fit using weighted residuals, rather than unweighted residual. We have discussed parameters of goodness of fit, and we have discussed the different fitting models. So we stop here today and next day, we start about another kind of experiment. It is called femtosecond upconversion or femtosecond optical gating.