Time Resolved Fluorescence Spectroscopy: Analysis

Patrik Callis

Appendix B: Advanced Data Analysis: Convolution


Objective: Obtain accurate decay parameters by trial and error variation of the theoretical function until its convolution with the instrument response function (IRF, i.e., laser light + electronic distortion only) matches your experimental results.

Background: “Determining the fluorescence decay rate from dye solutions becomes more difficult at high concentrations. We are using a photodetector with approximately 0.8 ns time resolution, and as the effective fluorescence decay time we are trying to measure approaches this value, the real fluorescence data becomes masked by the detector. The experimental observable is determined both by the detector time-resolution, and the time scale of the fluorescence relaxation. More generally, an experimental measurement is an integral over the product of the instrumental Response function (IRF), $R(t)$ (the laser pulse intensity as modified by the detector electronics time response) and the molecular response Signal from the sample $S(t)$ (which is what would be the fluorescence decay if there were no distortion from the electronics). The observed signal (the time-resolved fluorescence intensity signal on the oscilloscope) is termed a convolution integral:"

\[ I(t) = \int R(t') S(t - t') dt' \]

This is a physically understandable statement, if you think of $R(t')$ as a continuous burst of infinitesimally short pulses of light, each arriving at a different time $t'$ and amplitude $R(t')$, each causing a fluorescence $S(t-t')$. For example, $S$ might be an exponential decay, $\exp(-k(t-t'))$ if $t'$ is less than $t$, but zero otherwise. This is a sum of exponentials of different amplitude and each shifted by the different starting time, $t'$.

While this weighted linear combination of exponentials starting a different times, $t'$, is easily conceptualized, it is awkward (but not impossible) to simply calculate it with a spreadsheet. It is the overlap integral over starting times, but with the molecule response reversed in time. (This is because $t'$ appears with minus sign in $S(t-t')$.)
Because there is only one IRF curve, but several molecular S curves corresponding to different Nile Blue concentrations, it is more efficient to reverse the IRF curve. This can be done because mathematically it does not matter which is reversed. Thus, this equivalent form (See http://en.wikipedia.org/wiki/Convolution) is:

\[ I(t) = \int R(t - t')S(t')dt' \]

In words, this says the overlap integral (scalar product) of a single decay function starting at different times with only that part of the laser/response function that happens before time \( t \).

Now it is the laser/response function, \( R \) that must be reversed in time. This is easier to calculate on a spreadsheet because of the nature of Excel, and is what we will use here. Again using the exponential example, now in discrete form, this becomes:

\[ I_j = \sum_i R_{j-i} S_i = \sum_i R_{j-i} e^{-i\Delta t} \]

, where \( \Delta t = \) the time increment between points. This is saying that at time \( j\Delta t \) after some arbitrary time before the laser pulse, the signal is given by the scalar product of a column vector \( R \) dotted into column vector \( S \), whose components are \( S_i \), where \( i \) goes from 1 to a value such that \( S_i \) is negligible.

Written this way, the \( R_{j-i} \) are the time reversed \( R(t') \) (again because we are summing on \( i \) which has which has the minus sign in the subscript. Basically, the observed fluorescence intensity at time \( j \) is the sum of all dot products of the IRF (the trace obtained for laser light only) with the theoretical fluorescence decay function.

We can visualize 2 extreme limits:

1. The **perfect instrument**: This is if the instrument response is infinitely fast and light pulse is infinitely short in duration, then what we call the instrument response function, \( R(t) \) approaches a “delta” function (The Dirac delta function = 1 at some point, and = 0 otherwise, but the area under the curve = 1).

In this case, the molecular response Signal tells exactly what the molecules are doing (no distortion from the laser pulse width or the instrument electronics).

2. When the molecular fluorescence decay time is much shorter than the width of the light pulse and the instrument electronic response.
In this case, \( S(t) \) just follows exactly the instrument response. An example of this extreme is light scattering, for which the photon comes back out of the molecule in about 1 femtosecond (1 fs). This is why we obtained the IRF by measuring the signal with no dye in the solvent.

Perhaps somewhat surprisingly, when the two exist on similar time scales, we can still extract information about the molecular response \( S \) by comparing our fluorescence data with the convolution of the IRF with a theoretical molecular response \( S \), by varying the theoretical function until the convolution agrees well with the measured signal, \( I(t) \).

**Remarks:** All integrals are simply scalar products (dot products) in a continuous space. We can approximate them quite accurately in a spreadsheet by performing a discrete dot product between two columns (thinking of them as vectors).

The convolution integral above is simpler than it may look: The form \( R(t-t') \) means the instrument response function is in reversed time order (the minus sign on the \( t' \) means as we sum over \( t' \), \( R \) is going backward in time.

So, the observed signal at time \( t \), \( I(t) \), is just a series of dot products of the time reversed \( R \) column with the \( S \) column in which the time reversed \( R \) column successively overlaps later time segments of \( S \). Each successive dot product represents the predicted oscilloscope trace amplitude at a different time \( t \). This may be compared to the experimental curves to infer the correct parameters by the following strategy:

1. For the pure R6G sample, we determine R6G lifetime \( \tau_0 \) by manually varying its value in the spreadsheet until the calculated signal matches the observed signal within eyeball accuracy.

2. Using this value of \( \tau_0 \) for each of the R6G + NB solutions (where the concentration of the Acceptor \( n_A \) is known) we need only manually vary \( R_0 \) until the calculated curve superimposes the observed curve. Obviously, accurate value(s) can only come from those concentrations for which the FRET makes an observable difference in the intensity profile.

**Deconvolution**

While it is easy to create the convolution, our goal is to find the unknown molecular behavior, \( S(t') \). Because the \( R(t') \) is not a nice analytical function, there is no simple way to extract the \( S \) from the \( I \).

What most people mean by “deconvolution” is guessing the form of \( S \) that makes the convolution fit the observed result. That is what we will do.

**Using the Spreadsheet to guess the form of \( S \).**
Overview

The procedure consists of 7 distinct phases

Phase I: Conditioning the observed IRF signal.
   This involves:
   1. truncating the IRF to the useable length;
   2. making copies of the time and intensity columns
   3. changing the time units to ns,
   4. zeroing the baseline, and
   5. reversing the time ordering of the IRF

Phase II: (6.) Set up some cells that will contain the variable that will be manually varied.

Phase III: (7.) Enter the theoretical formula $S(t-t')$ such that it contains references to the
   cells set up in phase II.

Phase IV: (8.,9.) put time in col. F and the convolution formula in col. G

Phase V: 10. Plot the convolution to see if it looks correct. Experiment with the
   parameters in E1-E6

Phase VI: (11., 12.) Obtain $\tau_D$ from the R6G solution with no NB

Phase VII (13.-15.) Obtain R0 from the NB data curves by varying R0 which is in E4)

There are about 650 points (~130 ns spaced at 0.2 ns) in the laser response curve starting several
ns before it rises and ending several ns after it decays to what appears to be zero. *(We don’t need
all of this, but we will use it all for simplicity).* We are going to reverse the order so that the
bottom of the column represents the first photons from the laser to hit the dye and those at the
top are the last.

To the right of the time reversed laser column we put the theoretical molecular response function
starting at row 410 (or whatever the length of the IRF data is).

The convolution integral = a series of dot product of these two columns; the first tells us the
fluorescence intensity produced by the first photons arriving at the dye that are emitted instantly.
If the time reversed IRF (laser) column is dotted into a vector starting 1 row lower in the S
column, we now get the signal predicted in the second time interval, which is that generated by
the first photons but emitted after a delay of 1 time interval plus signal generated by photons
from the laser arriving at the dye in the 2nd interval and emitted instantly.
Below you will see that we will make a column of dot products in column G, starting at G10, that will be predicted fluorescence curve:

The first dot product will be in cell G10 and will be predicted fluorescence curve: 

\[
\text{G10} = \text{SUMPRODUCT( D$10:D$410, E10:E410 )}
\]

By copying that formula down the column G8 will = SUMPRODUCT( D$10:D$410, E11:E411 ),

\[
\text{G9} = \text{SUMPRODUCT( D$10:D$410, E12:E412 )}
\]

and so on. The $ sign causes what is next to not change as the formula is copied down the column. Thus, the D column stays put, but overlaps successively pieces of the E column that are farther down the column E (later times).

Before going on, pause to realize that for any a single molecule that is excited, the probability that it will emit a photon in the next 0.2 ns is a constant dependent only on its transition dipole moment. Molecules emit randomly, like flipping a coin. (Even if by a quirk of luck you have gotten heads 5 times in a row, the probability of a tails on the next toss is still 50%)

We will use an exponential decay as the S function for the unquenched R6G. This is describing the rate of emission from the entire ensemble of molecules. There are fewer molecules left as time goes on because once the excited molecule has emitted a photon, it cannot emit again until the it gets excited again during the next pulse.

Continuing the example, after shifting the laser column down by 10 increments, the dot product will give signal emitted at time 10 by 10 ways: the 10 ways ranging from that generated from laser light arriving at time 10 and emitted instantly to signal generated from light arriving at time 0 and emitted after a delay of 10. And so on. Here is an example of how to do this.

What does it mean shifting the laser column down?

**Detailed Procedure**

**PHASE I**

1. Paste the laser light (instrument response function = IRF) curve into cols A (time) and B (intensity). This should be only the range -1 x 10^-8 to 7 x 10^8 -10 ns to + 70 ns

   For convenience for what comes later, start at row 10.

2. using a) and b) below: make copies of the time and IRF so that col C is time in ns and col D is a positive IRF signal that is zero at the start and ending times.

   If the baseline is not within about 0.1 % of the maximum, the calculation will give a distorted curve in which the non-zero baseline creates simulated fluorescence intensity
before the pulse. If the baseline is much more than 0.1 % of maximum, we will subtract a constant from the IRF that will reduce it to ~0 on average.

a) make column C values = column A values * 1e9 (this makes the units ns).
   i.e., write =A1*1e9 in the top cell and copy the formula down the column
b) make Column D = B1 + or - (whatever number will zero the baseline)

c) Choose “number” instead from the number menu instead of scientific
d) make a quick plot to check how things look: select the C and D data
   insert, scatter plot, straight lines
   should look like this

![Chart Title](chart.png)

3. copy columns C and D, and from the Edit pulldown paste special (values) back into A and B to get rid of the formula. Now the time and IRF are just numbers and may be manipulated without complications associated with moving formulas.

4. copy A and B back into C and D. We will work with C and D, and A and B are the backups in case C and D get butchered.

5. Reverse the order of D by: select C and D and Data Sort using Col C descending. Then copy A to C again so that time in C is again going forward and extend the times in column C to 70 ns, i.e., to a time for which the signal curves for all samples have decayed to zero.
6. To ease the trial and error process, insert 6 rows at the top, and **enter 6 constants at the top of col E.**

In cell **E1** put the trial decay time ($\tau_D$) for the pure R6G. (start with 8 ns)

In cell **E2** put a scale factor to adjust the height of the trial function. (start with 1)

In cell **E3** we may later enter a shift in time to facilitate comparing with experimental curves.

In cell **E4** is the trial $R_0$, the distance in Å at which the probability = 50% for energy transfer for a molecule that has been excited. Use 100 Angstroms to start.

In cell **E5** put the molar concentration of Nile Blue. Put 0. to start

In cell **E6** put $=E5*6.02e23/10^27$ which is $n_A = $ concentration of Nile Blue in molecules/Å$^{-3}$.

---

7. Enter the theoretical S curve, Equation 8, into **Col E**, **starting at the cell beside which the reverse R ends**, *i.e.*, beside the earliest laser point in time (which will be about row 411).

$$I_D(t) = \exp \left[ -t/\tau_D - \frac{192\pi^2 (\kappa^2) \tau}{18\tau_D} n_A R_0^3 \right]$$

(8)

* **Note that $n_A R_0^3$ is not under the radical**

Make sure the units are correct: t and $\tau_D$ must have the same units and $n_A$, the concentration of quencher in **molecules per unit volume**, requires that $R_0^3$ must have reciprocal units. e.g., make $n_A$ in units of molecules per cubic Å if $R_0$ is in Å (expect on the order of 50 Angstroms.)
Put a 1.0 in cell E2; we will vary this number later to adjust the height.

Something like the following should be in E410:

$$=\$E$$2*\exp(-(C410-70)/\$E$1-(192*\pi)^3*0.667*(C410)/(18*\$E$1))^{0.5*\$E$6*\$E$4^3}$$

where 70 is the value of time in cell C410, (just so that the value of the exponential will =1.0 arbitrarily).

Copy this cell down for the same number of cells in the IRF data in Col D.

8. To facilitate quick plotting of the calculated curve in column G, put a copy of column C in column F.

9. Starting at row 10, enter into col G the dot product of col D with col E. Excel calls this the SUMPRODUCT. In cell G10 you put:

```excel
=SUMPRODUCT(D$10:D$410, E10:E410)
```

This is the convolution of the IRF with the theoretical Signal function.

You may have a bit more or less points. 410 here stands for the last IRF cell (bottom) of D data.

Copy the cell G10 to fill G down to the bottom of the E data.

Note that the two arrays D and E must be exactly the same length. The $ signs means that D array indices stay fixed while the E array indices increase down the Col F. This means that the D array will progressively overlap with the E array at later and later times, thereby creating the entire simulated decay (convolution) curve in G.

**IMPORTANT:** The time in Col C MUST be extended down to 140 ns (because we started the S function at time=70 and extended it down another 70 ns.)

The G result is the trial calculated oscillograph trace.

10. Plot the convolution G by selecting F (the time) and G (the convolution) and make a xy scatter plot.

**If the plot does not look correct:** It is easy to make mistakes getting the spread sheet set up correctly. It is strongly suggested that you temporarily replace the Col D data with all zeros except for a 1 in the last cell (cell 410 in this example). This should create the pure theoretical molecular curve (S) you entered in Col E, starting at the time associated with cell F10. Then change the location of the 1 to a later time (upwards in Col D) and add some more 1's (or other numbers) spaced a few ns apart. This should show you a series of overlapping exponentials in Col G weighted by the numbers you put in column D. Once you have control and feel you get what you expect, put the reversed laser intensity data back into D. Then vary the $D$, scale, R0, and concentration values to understand the effect they have.

11. **Compare to the R6G dye only experimental curve:** copy the observed decay curve for R6G without quencher into column J, starting at J10. Copy the time series of col C into col I.
Now make columns K and L to be adjustable versions of columns I and J: In col K10 write =I10+$K$1 and copy down. In L1 write =J10+$L$1 and copy down. Now K and L are the same as the original experimental time and intensity, but time will be shifted by what is in K1 and the baseline of L will be shifted up or down by what is in L1.

The height of the trial curve G is adjusted with what is in E2.

Add the K,L plot to the graph by adding a series by right clicking on the plot, selecting Select Data, Add, and highlight the data for x and for y.

Procedures for extracting $\tau_D$ and $R_0$:

12. First use the $~1 \times 10^{-4}$ M R6G only as the observed decay curve. Set $n_A$ or $R_0 = 0$, so that the Equation 8 is simply a pure exponential. By adjusting the decay time $\tau_D$, scale (E2), and shift constants(K1,L1), match the calculated and observed spectra until they superimpose as exactly as possible. This is possible to do so well that, by eye, the curves are perfectly superimposed. (See Figure 4.)

From this you find one result, the natural decay time, $\tau_D$. The scale and shift have no physical significance. They are only so you can tell if the calculated and observed curves are exactly the same shape.

Do 13., 14., and 15. below for each Nile Blue solution:

13. Insert a new WorkSheet from the Insert pull down. Copy the entire worksheet with the R6G convolution into the new worksheet, and replace only the observed spectrum (cols I,J) with one from a solution with the Nile Blue acceptor.

14. Use the $\tau_D$ value found for R6G, and make sure the correct concentration is in the col E formula (value in E5), and vary $R_0$, scale, and shift again until a near perfect match is found. Each concentration should give the same $R_0$, but experimental precision will probably lead to some variation, especially for low concentrations of the quencher.

15. Report $\tau_D$ for R6G and the $R_0$ values found for the solutions with quencher, and the average. Include plots of the superimposed curves that lead to the $R_0$ values. (see example on last page)

Make a rough estimate of the possible error by noting the size of changes in the parameters required to create an obvious mismatch.

Compare the observed $R_0$ with that which you computed from Eq. 6. (If asked to do this.)
**Figure 4.** Decay curve from rhodamine 6-G fluorescence quenched by nile blue obtained with the actual apparatus described in Figure 1. The solid magenta curve is experimental data, and the dashed blue curve is an example of the fit using trial and error convolution of the theoretical response of **Equation 8** with the instrument response function found from the scattering of the excitation pulse only. This is the type of match expected (if not better).