Time-resolved fluorescence measurements

Notes before starting

- Please also read the "**3 Steady-state FRET**" protocol for a basic understanding of FRET
- Please also refer to
    - The manual of the DataStation Software
    - The manual of the IBH TCSPC counting module "DataStation Hub"
    - The manual of the NanoLED light source
    - "The bh TCSPC handbook" published by Becker & Hickl, e.g. p. 61-65, 259-272
    - "Principles of Fluorescence Spectroscopy" by **Joseph R. Lakowicz**, chapter 1, 2.1, 4, 8, 10, 11, 13 & 14
    - The "TCSPC technote" from Picoquant
- **I have generated two training data sets, which will be used in the data analysis section**

Material & Methods

- FluoroMax4 equipped with the IBH time-correlated single photon counting module and the pulsed light source NanoLED
- Fluorescence cuvette (L-shaped excitation & emission configuration)
- Buffer for sample dilution
- Samples: Single donor labeled ("Donly", D0) and double donor-acceptor ("DA") labeled
- Ludox (colloidal silica solution)

Software for data analysis & fitting: **ChiSurf**

- Download for free at: http://www.fret.at/tutorial/chisurf/

Theoretical Background

Instrumentation

Time-resolved fluorescence intensities are measured via time-correlated single photon counting (TCSPC) using a pulsed excitation light source and a single-photon sensitive detector. The arrival time of the emitted fluorescence photon is recorded with picosecond accuracy with respect to the incoming excitation light (Figure 1).

Repeating the excitation-emission process many times will give a decay profile (Figure 2). Pulsed lasers or LEDs can be used as a source of excitation. The excitation light is passed through a polarizer before it hits the sample in the cuvette. Part of the light passes through the sample, the other to the electronics as "sync" signal. The light emitted by the sample molecule is passed first through a polarizer and then through a monochromator to select a specific wavelength (~range). The light is then detected and amplified by a photomultiplier tube (PMT). The emitted light signal as well as reference light signal is processed through a constant fraction discriminator (CFD) which eliminates timing jitter. After passing through the CFD, the reference pulse activates a time-to-amplitude converter (TAC) circuit. The TAC charges a capacitor which will hold the signal until the next electrical pulse. In reverse TAC mode the signal of "sync" stops the TAC. This data is then further processed by an analog to digital converter (ADC) and multi-channel analyzer (MCA) to get a data output. To make sure that the decay is not biased to early

arriving photons, the photon count rate is kept low (usually less than ~ 1.5 % of excitation rate (repetition rate of the light source).

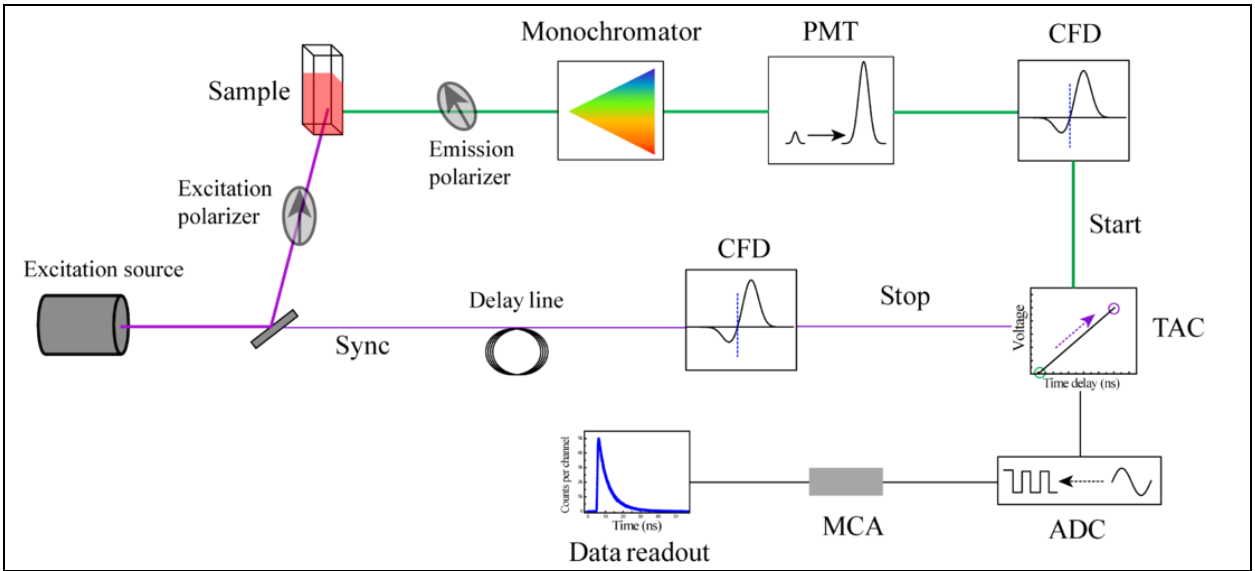


Figure 1. Schematic overview of a Time-correlated single photon counting setup *(Source: Wikipedia, https://en.wikipedia.org/wiki/Ultrafast_laser_spectroscopy#/media/File:TCSPC_schematic.png)*

The maximum – and usually used – repetition rate of the here available NanoLED is 1 MHz. Thus, the count rate on the detector should not exceed 15'000 counts/sec. Laser and LEDs can be operated at much higher repetition rates and are rather limited by the fluorescence lifetime of the samples.

The time interval between two laser pulses should be ~ 10x as long as the fluorescence lifetime time τ. For Alexa488, τ = 4 ns resulting in Δt >= 40 ns and maximum repetition rate of 1/40 ns = 25 MHz.

For the MCA, the longer a molecule takes to emit a photon, the higher the voltage of the resulting pulse. The central concept of this technique is that only a single photon is needed to discharge the capacitor. Thus, this experiment must be repeated many times to gather the full range of delays between excitation and emission of a photon (Figure 2). After each trial, a pre-calibrated computer converts the voltage sent out by the TAC into a time and records the event in a histogram of time since excitation.

This histogram has a limited number of bins (for our setup 1024), which are equally distributed between two subsequent light pulses. In our setup for a repetition rate of 1 MHz, the time between two pulses is 1/1 MHz = 100 ns.

Since the probability that no molecule will have relaxed decreases with time, a decay curve emerges that can then be analyzed to find out the decay rate of the event.

Since the emission of the photons is a stochastic process, it can be described by an exponential decay $F(t)$. The inverse of the characteristic decay time constant k_F is called the fluorescence lifetime τ of the respective fluorophore.

$$F(t) = F_0 \exp(-t \cdot k_F) = F_0 \exp(-t \cdot / \tau) \qquad \text{Eq. 1}$$

τ is defined as that time where the initial maximal intensity after excitation F_0 dropped to 1/e of its original value. The fluorescence lifetime of a fluorophore depends on many factors, including the surrounding environment e. g. the solvent or vicinity to certain amino acid residues when coupled to a protein, which can quench the fluorescence lifetime. As the environment of a fluorophore coupled to a biomolecule is heterogeneous, the fluorescence lifetime decay is in most cases no longer mono exponential but has to be described by more complex models.

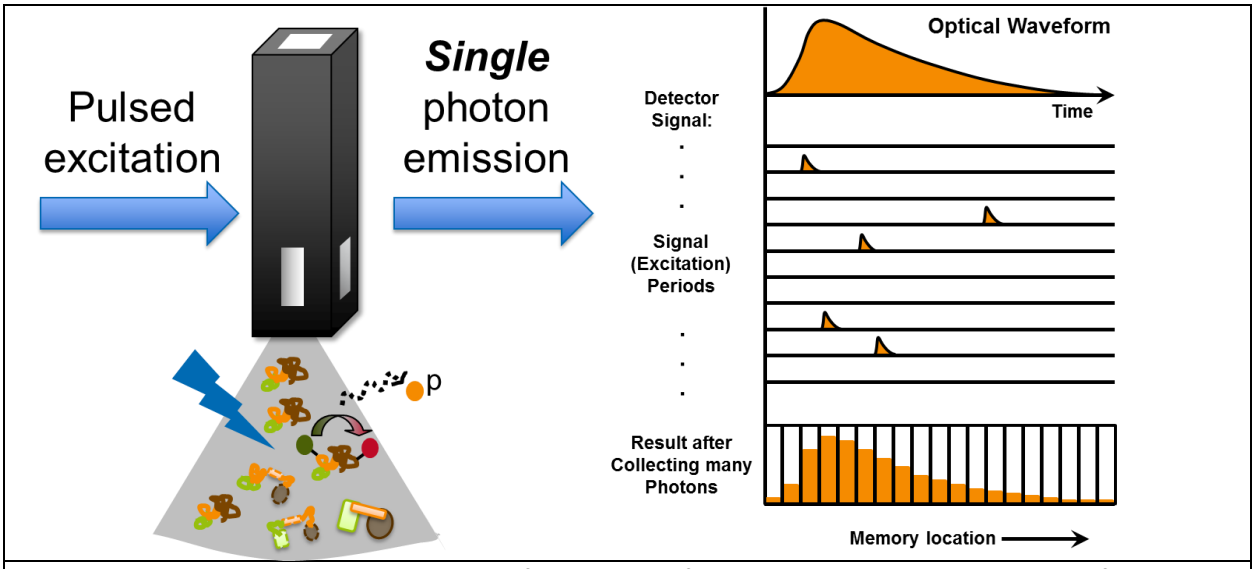


Figure 2. Repeated excitation & collection of the emitted fluorescence photons generated the fluorescence decay histogram.

Another complicating factor is that many decay processes involve multiple energy states, and thus multiple rate constants. Though non-linear least squared analysis can usually detect the different rate constants, determining the processes involved is often very difficult and requires the combination of multiple ultra-fast techniques.

Data analysis

For a quantitative description of the energy transfer E according to (Eq. 2), the lifetime of the donor in the presence and absence of the acceptor is required. In the presence of the acceptor, i.e. at FRET occurring, a quenching of donor fluorescence is observed.

$$E = 1 - \left(\tau_{D(A)} / \tau_{D(0)}\right) \qquad \text{Eq. 2}$$

$\tau_{D\,(A)}$ is the lifetime of the donor in the presence of acceptor and

$\tau_{D\,(0)}$ is the lifetime of the donor in the absence acceptor.

The determination of fluorescence lifetimes is realized by measuring the temporal fluorescence decrease. The fluorescence decay of the donor, $F_D(t)$, and the double-labeled DA-sample, $F_{DA}(t)$ are to be measured. For these applies:

$$F_D(t) = F_D(0) \cdot e^{-\frac{t}{\tau_D}} \qquad \text{Eq. 3}$$

$$F_{DA}(t) = F_{DA}(0) \cdot e^{-\frac{t}{\tau_{D(A)}}} = F_{DA}(0) \cdot e^{-t\left(\frac{1}{\tau_D} + k_{FRET}\right)} \qquad \text{Eq. 4}$$

As in the measured double-labeled sample a portion of the molecules is always single-labeled, the measured fluorescence in the donor channel ***$F_D(t)$* is composed of two parts**: fluorescence of **pure donor** in the absence of the acceptor, $F_D(t)$, **superimposed by (quenched) donor fluorescence** in the presence of the acceptor, $F_{DA}(t)$:

$$F_d(t) = \left[S_{d^+a^-}\right] \cdot F_D(t) + \left[S_{d^+a^+}\right] \cdot F_{DA}(t) = \left[S_{d^+a^-}\right] \cdot F_D(0) \cdot e^{-\frac{t}{\tau_D}} + \left[S_{d^+a^+}\right] \cdot F_{DA}(0) \cdot e^{-\frac{t}{\tau_{D(A)}}} \qquad \text{Eq. 5}$$

The time course of fluorescence of the acceptor, ***$F_A(t)$***, is due to the energy transfer, k_{FRET}, **given by** two terms: **an increase in time with τ_D(A) and a temporal decay with the fluorescence lifetime, τ_A,** of the acceptor, which would occur after a direct excitation of the acceptor (for $\Phi_D = \Phi_A = 1$):

$$F_A(t) = F_{DA}(0) \frac{k_{FRET} \cdot \tau_{D(0)} \cdot \tau_A^{-1}}{\tau_{D(A)}^{-1} - \tau_A^{-1}} \cdot \left(e^{-\frac{t}{\tau_A}} - e^{-\frac{t}{\tau_{D(A)}}} \right)$$

$$= F_{DA}(0) \cdot A \cdot \left(e^{-\frac{t}{\tau_A}} - e^{-\frac{t}{\tau_{D(A)}}} \right) \qquad \text{Eq. 6}$$

The latter corresponds formally to a mathematical convolution of the two acceptor fluorescence signals which result from the direct excitation of the acceptor and the FRET sensitized fluorescence. Thus:

$$e^{-k_1 t} \otimes e^{-k_2 t} = \int_0^t e^{-k_1 (t-t')} \times e^{-k_2 t'} \, dt' = \{e^{-k_1 t} - e^{-k_2 t}\} / (k_2 - k_1) \qquad \text{Eq. 7}$$

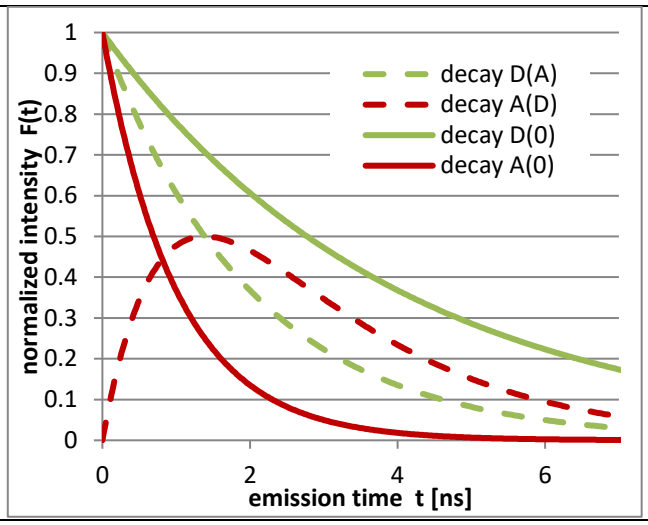


Figure 3 Fluorescence decay *of donor (green) and acceptor (red), excited directly without FRET (solid line) and via FRET (dashed).* $\tau_D = 4$ *ns,* $\tau_A = 1$ *ns,* $k_{FRET} = 0.25$ *ns*$^{-1}$.

When the measured sample exhibits a multi-exponential fluorescence decay of the acceptor; the acceptor fluorescence lifetime can be described by Eq. 8:

$$F_A(t) = \sum_i x_i e^{-t/\tau_{xi}} \qquad \text{Eq. 8}$$

x_i designates the amplitudes, which represent the fluorescent species concentration of the different components. The acceptor fluorescence in the presence of the donor and in case of FRET exhibits a multi-exponential fluorescence decay of the following formal shape:

$$F_A(t) = \sum_i \alpha_i e^{-t/\tau_{Ai}} - \sum_i \alpha_i e^{-t/\tau_{D(A)}} = \sum_i \alpha_i \left(e^{-t/\tau_{Ai}} - e^{-t/\tau_{D(A)}} \right) \quad \text{Eq. 9}$$

α_i denotes the general amplitudes. For a bi-exponential decay of the acceptor with the following relationships between the three amplitudes is obtained:

$$\begin{aligned}
F_A(t) &= F_{DA}(0)\left(x_1 \frac{k_{FRET} \cdot \tau_{D(0)} \cdot \tau_{A1}^{-1}}{\tau_{D(A)}^{-1} - \tau_{A1}^{-1}} \left(e^{-\frac{t}{\tau_{A1}}} - e^{-\frac{t}{\tau_{D(A)}}} \right) + x_2 \frac{k_{FRET} \cdot \tau_{D(0)} \cdot \tau_{A2}^{-1}}{\tau_{D(A)}^{-1} - \tau_{A2}^{-1}} \left(e^{-\frac{t}{\tau_{A2}}} - e^{-\frac{t}{\tau_{D(A)}}} \right) \right) \\
&= F_{DA}(0)\left(x_1 A_1 \left(e^{-\frac{t}{\tau_{A1}}} - e^{-\frac{t}{\tau_{D(A)}}} \right) + x_2 A_2 \left(e^{-\frac{t}{\tau_{A2}}} - e^{-\frac{t}{\tau_{D(A)}}} \right) \right) \\
&= F_{DA}(0)\left(x_1 A_1 e^{-\frac{t}{\tau_{A1}}} + x_2 A_2 e^{-\frac{t}{\tau_{A2}}} - (x_1 A_1 + x_2 A_2) e^{-\frac{t}{\tau_{D(A)}}} \right) \\
&= \alpha_1 e^{-\frac{t}{\tau_{A1}}} + \alpha_2 e^{-\frac{t}{\tau_{A2}}} - (\alpha_1 + \alpha_2) e^{-\frac{t}{\tau_{D(A)}}} \\
&= \alpha_1 e^{-\frac{t}{\tau_{A1}}} + \alpha_2 e^{-\frac{t}{\tau_{A2}}} - (\alpha_1 + \alpha_2) e^{-\frac{t}{\tau_D}} e^{-k_{FRET} \cdot t}
\end{aligned} \quad \text{Eq. 10}$$

In practice, the pre-exponential factors are often free fitting parameters, as for a calculation in addition to the fluorescence quantum efficiencies of the dyes the detection efficiencies of the detection channels would have to be known. A determination of these quantities is complicated in general.

Fit models

The fluorescence decay of the donor in the absence of FRET is multi-exponential, most likely, due to local quenching. To account for these effects, the donor only reference samples are fitted by a multi-exponential relaxation model.

$$F_{\mathrm{D}(0)}(t) = \sum_i x_{\mathrm{D}(0)}^{(i)} \exp(-t/\tau_{\mathrm{D}(0)}^{(i)}) \quad \text{Eq. 11}$$

Here $\tau_{\mathrm{D}(0)}^{(i)}$ is the donor fluorescence lifetime and $x_{\mathrm{D}(0)}^{(i)}$ are the pre-exponential factors.

These effects are considered in the analysis of the FRET samples by global analysis. Here, it is assumed that all donor species are quenched by the same FRET rate constant k_{RET}. This is true if quenching does not change the donor radiative lifetime and if FRET is uncorrelated with quenching of the donor by its local environment. Given these assumptions the donor fluorescence intensity decay in the presence of FRET $F_{D(A)}(t)$ can be factorized into the donor fluorescence decay in absence of FRET and the time-resolved FRET-induced donor quenching $\varepsilon_{D(A)}(t)$:

$$F_{D(A)}(t) = F_{D(0)}(t) \cdot \varepsilon_{D(A)}(t) \quad \text{Eq. 12}$$

The FRET-induced donor decay is related to the distribution of distances by the rate-constant of energy transfer as defined by Förster:

$$k_{RET} = k_F \cdot \kappa^2 \cdot \left(\frac{R_{0J}}{R_{DA}} \right)^6 \quad \text{Eq. 13}$$

Here, R_{0J} is a reduced Förster-radius, k_F is the radiative rate constant of fluorescence and κ^2 is the orientation-factor. The reduced Förster-radius is given by:

$$R_{0J} = \left[\frac{9(\ln 10)}{128\pi^5 \cdot N_A} \cdot \frac{J}{n^4} \cdot \right]^{\frac{1}{6}} = 0.2108 \cdot \text{Å} \cdot \left[\frac{1}{n^4} \cdot \left(\frac{J(\lambda)}{\text{mol}^{-1} \cdot \text{dm}^3 \cdot \text{cm}^{-1} \cdot \text{nm}^4} \right) \right]^{\frac{1}{6}} \quad \text{Eq. 14}$$

where N_A is the Avogadro's constant, n is the refractive index of the medium and $J = \int f_D(\lambda) \cdot \varepsilon_A(\lambda) \cdot \lambda^4 \cdot d\lambda$ is the overlap integral between $f_D(\lambda)$, the donor emission spectrum and $\varepsilon_A(\lambda)$, the acceptor absorption spectrum. This reduced Förster-radius stresses that the FRET-rate constant is independent of quenching and specific for the dye-pair under the condition that the spectral overlap is independent of dynamic quenching. Under the assumption that dynamic quenching is uncorrelated with FRET the FRET-induced donor decay relates to the distance distribution *p(R_{DA})* by:

$$\varepsilon_{D(A)}(t) = \int p(R_{DA}) \cdot \exp\left(-t \cdot \left\langle \kappa^2 \right\rangle \cdot k_F \left[1 + (R_{0J} / R_{DA})^6\right]\right) dR_{DA} \quad \text{Eq. 15}$$

Assuming the orientation factor can be approximated by an average $\langle\kappa^2\rangle \approx 2/3$, this time-dependent quantifier is directly related to the donor-acceptor distance distribution.

Discrete distance distributions are described by:

$$p(R_{DA}) = \sum_{i=1}^{N} x_{DA}^{(i)} \cdot \delta(R_{DA} - R_{DA}^{(i)}) \quad \text{Eq. 16}$$

where δ is the delta-function, *N* is the number of states with the apparent DA-distances $R_{DA}^{(i)}$ and $x_{DA}^{(i)}$ species fraction of the state.

The fluorophores coupled to the biomolecule are moving entities coupled to specific places via flexible linkers. Therefore, already for single protein conformations a distance distribution has to be considered. For simplicity, normal distributions are used to describe this dependency. If the donor and acceptor in average are separated by $\overline{R_{DA}}$ the corresponding distance distribution is approximated by

$$p(R_{DA}) = \sum_{i=1}^{N} x_{DA}^{(i)} \frac{1}{w_{DA}\sqrt{\pi/2}} \exp\left(-2 \left[\frac{R_{DA} - \left\langle R_{DA}^{(i)} \right\rangle}{w_{DA}} \right]^2 \right) \quad \text{Eq. 17}$$

Here, $\left\langle R_{DA}^{(i)} \right\rangle$ is the mean of the state *(i)* distance distribution with species fraction $x_{DA}^{(i)}$ and a width w_{DA} set to a physical meaningful value of 12 Å (flexible dye-linkers).

The final analysis model is obtained by substituting either eq. 16 or eq. 17 into eq. 15. Next, eq. 15 is inserted into eq. 12. Finally, the fluorescence intensity decays of the donor in presence and absence of FRET (eq. 18) are globally analyzed. By this procedure the photon counting statistics of both the reference- and fluorescence-decay in presence of FRET is preserved. Thus, the counting statistics are clearly defined (Poisson distribution). This allows for an analysis with proper error-estimates. By the global (joint) analysis of the reference sample and the FRET-sample the photophysical properties (dynamic quenching) are taken into account.

Given the model functions, the experimental fluorescence intensity decays are fitted using the iterative re-convolution approach. Here, the model-decay curves are convoluted with the experimental instrument response function (*IRF*). Furthermore, a constant offset *c* of the fluorescence intensity is considered. Given these corrections the experimental time-resolved fluorescence intensities of the FRET-sample and the donor reference sample are proportional to:

$$\begin{aligned} F_{FRET}(t) &= \left(N_0 \cdot \left[(1 - x_{DOnly}) F_{D(A)}(t) + x_{DOnly} F_{D(0)}(t)\right] \otimes IRF + sc \cdot IRF + c\right) \\ F_{Ref}(t) &= \left(N_0 \cdot F_{D(0)}(t) \otimes IRF + sc \cdot IRF + c\right) \end{aligned} \quad \text{Eq. 18}$$

Here, *sc* is due to scattered light of the sample. The scaling of the model fluorescence intensity histograms was performed given the experimental measured number of photons. The model function was scaled such that the number of photons corresponds to the experimental number of photons. This reduces the number of free fitting parameters as the initial amplitude N_0 is not fitted.

Experimental procedure

Software for data collection: DataStation

The FluoroMax allows either the FlourEssence OR the DataStation software to be open at a time.

When the DataStation software is not running, the emission of the NanoLED used for pulsed excitation is OFF.

Before starting, verify (I) that the correct NanoLED is connected (which fluorophores would you like to measure today?), and (II) that the connecting cables between FluoroMax & DataStation Hub are in the right modus. Here, measurements will be performed at 1 MHz repetition rate, i.e. the **reverse counting mode** is needed. The cables are on the back of the DataStation Hub (Figure 4).

	**<u>Reverse counting:</u>** Top **Blue cable** is connected to *TAC start* (from the detector of the Fluoromax). Bottom **black cable** is connected to *TAC Stop* (from the NanoLED).		**<u>Forward counting:</u>** Top **black cable** is connected to *TAC Start* (from the NanoLED). Bottom **Blue cable** is connected to *TAC stop* (from the detector of the Fluoromax).
Figure 4 Cable connection between FluoroMax & DataStation Hub			

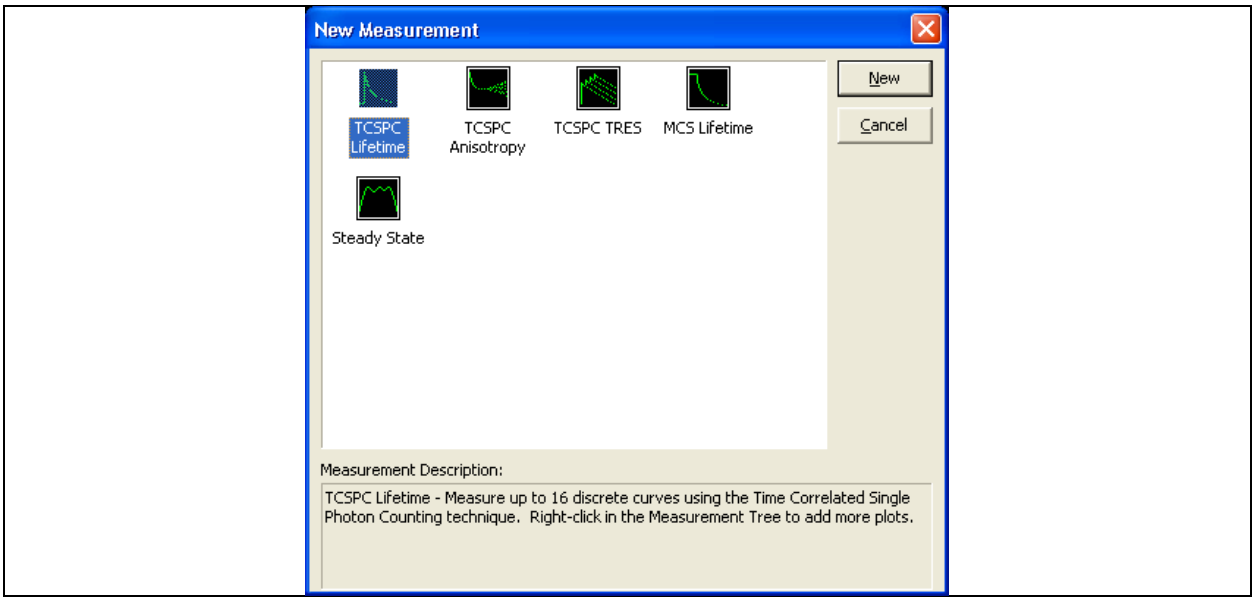
For the connection of NanoLEDs, it seems that their orientation is not relevant. Nevertheless, I prefer to connect them always in the same orientation such that the little red dot on the NanoLED drive is on top:

Figure 5 Connection of NanoLED to the FluoroMax

Now, start first the FluoroMax, then the Datastation Hub (TCSPC counting module) and finally the computer. Start the DataStation Software, it recognizes the FluoroMax and asks you whether you would like to use this device for measurement, select it. Then it asks you which polarizer to use, you select FluoroMax.

Figure 6 Starting the DataStation software

Select the experiment, which you would like to perform:



Turn on the NanoLED power in the Software by clicking in the NanoLED controls on the little red dot next to “Output: off”. The dot should turn green and change to “Output: on”. The repetition rate should be 1 MHz and no Sync delay. Let the NanoLED equilibrate for 30 min prior to the first measurement.

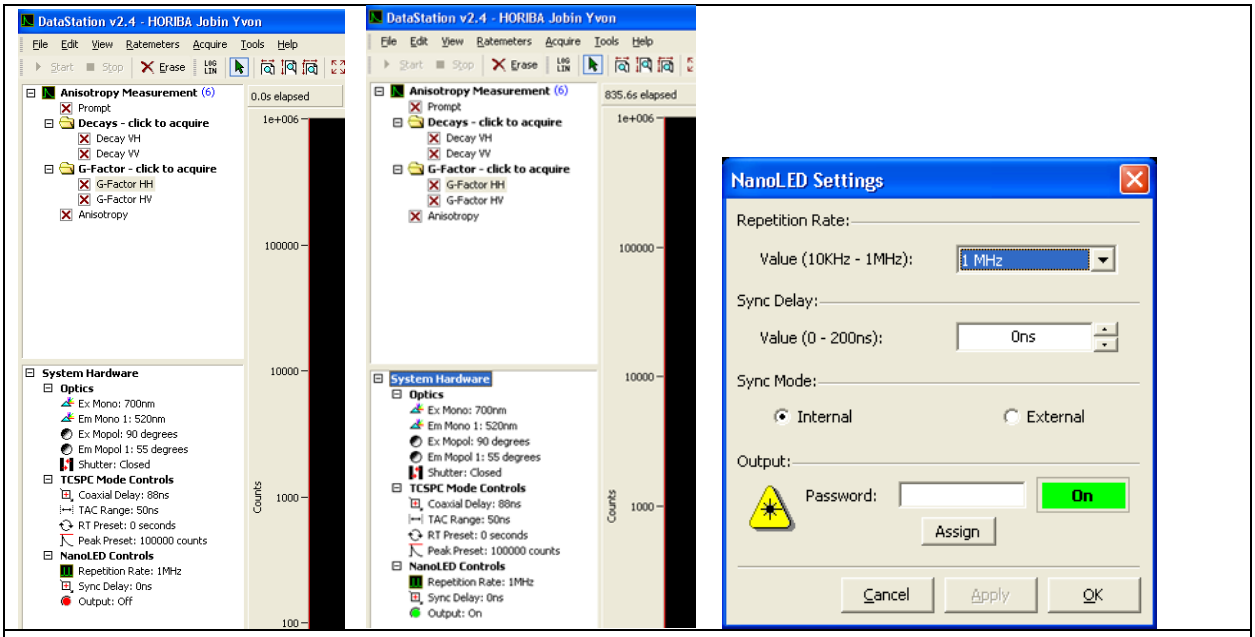
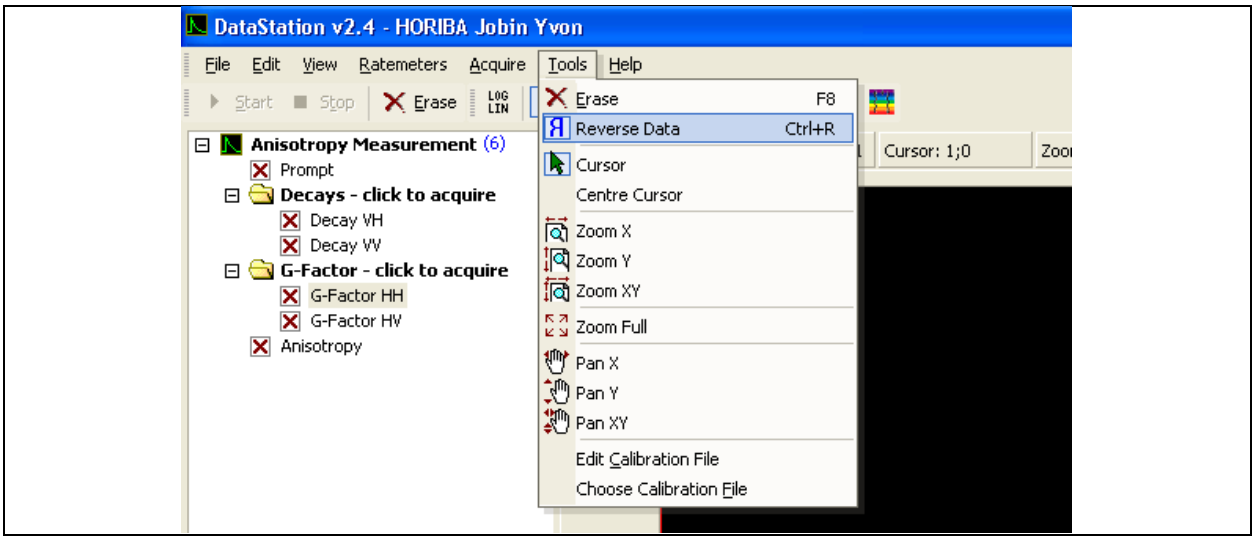


Figure 7 Switching on the NanoLED and NanoLED setting.

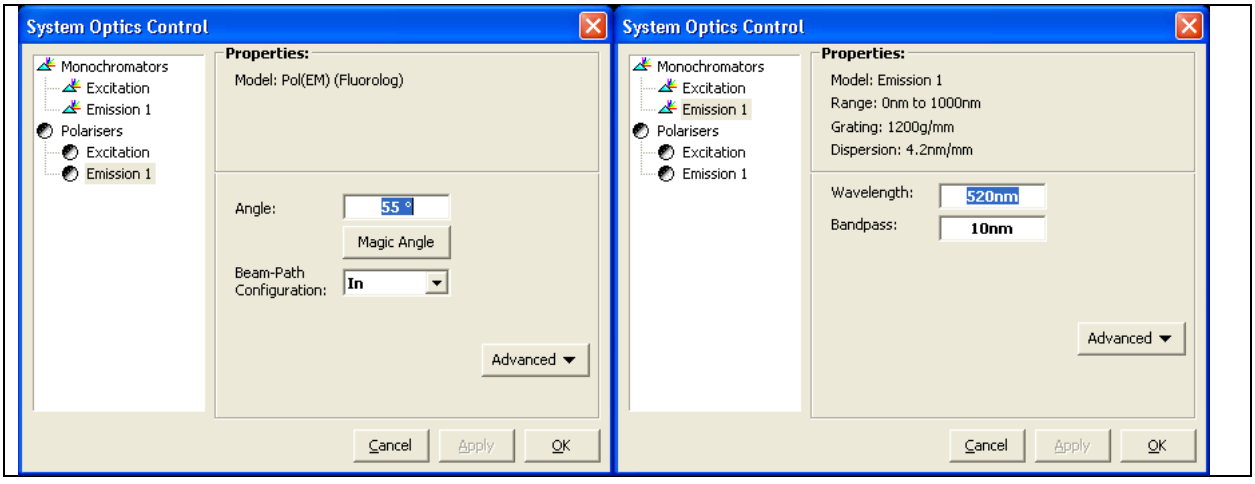
Select reverse counting option in the software:



Next, go to the TCSPC mode controls. For fluorescence lifetime measurements, the peak preset, i.e. number of photons in the histogram maximum, should be 100’000. For very low concentrated samples reduce down to 30’000 (the less photons, the noisier the data).

These settings result in 1024 time channels, each has a width of 56.3 ps, and total TAC window of 57.65 ns.

Finally, adjust the System Optics Control. For fluorescence lifetime experiments, the emission polarizer should be “in” and at “Magic Angle”. The detection wavelength depends on the fluorophore used, i.e. for Alexa488 it is 520 nm, for Alexa594 620 nm. The width should be adjusted depending upon count rate, but not exceed 10-15 nm.



Make sure the orientation of the **excitation polarizer**, which you have to adjust manually, is correct:

Excitation light:
0 °(V)

Emission light detected at:
54.7 ° => Magic Angle (VM)

or alternating at
0 °(V) => VV
90 °(V) => VH
for anisotropy measurements

Finally, start the first measurement, by clicking first on the measurements, which you would like to start (i.e. “Prompt” or “Decay”), then press start.

You need to measure:

1. **The donor fluorescence decays for the D0 and DA sample**
2. **The instrument response function (IRF) at excitation wavelength (diluted Ludox solution) under otherwise identical conditions than the samples.**

To save the data, right-click on the measurement, then select export as .txt file -> all channels (1024) and save under an appropriate name.

Data analysis using ChiSurf

ChiSurf is a Python-based program aimed for analysis of time-resolved fluorescence data, may they be collected in ensemble or single-molecule fluorescence experiments. It can fit time-resolved fluorescence data and fluorescence correlation curves, and can simulate fluorescence observables.

For this tutorial I will focus on how to use ChiSurf for the analysis of ensemble fluorescence lifetime experiments measured by a NanoLED equipped FluoroMax. For the purpose of training I generated two data sets (set I and set II).

The instrument response function

The characteristic of a complete TCSPC system that summarizes its overall timing precision is its Instrument Response Function (IRF or Prompt). The basic idea is that if the system is ideal, i.e., has an

infinitely sharp excitation pulse and infinitely accurate detectors and electronics, it should have an infinitely narrow IRF. Any deviation from this ideal results in a broadening of the IRF.

The weakest component in terms of timing resolution in TCSPC measurements will usually be the detector, here, in the case of NanoLED, it is the pulse-width of pulses generated by the NanoLED: the 494 nm NanoLED has a width of ~ 1.4 ns, the 591 nm NanoLED of ~ 1.6 ns.

For comparability, the full-width of half-maximum (FWHM) is calculated. For the IRF measurement, a diluted solution of LUDOX (colloidal silica) is used. **It should be diluted such that its count rate is comparable to the count rate of the samples and must be measured under identical conditions as the sample (i.e. polarization, slit width).** The only difference is its detection wavelength. As LUDOX is not fluorescent, and only scatters the light, it must be recorded at the excitation wavelength.

The measured decay of the samples is a convolution of IRF and sample fluorescence decay and must be deconvoluted in the fitting process to extract the "real" fluorescence decay.

Add picture IRF NanoLED + IRF PQ Fluotime (for comparison, how it is)

Multi-exponential fit of D0 sample

Start ChiSurf by double-clicking on the icon. This might take a while.

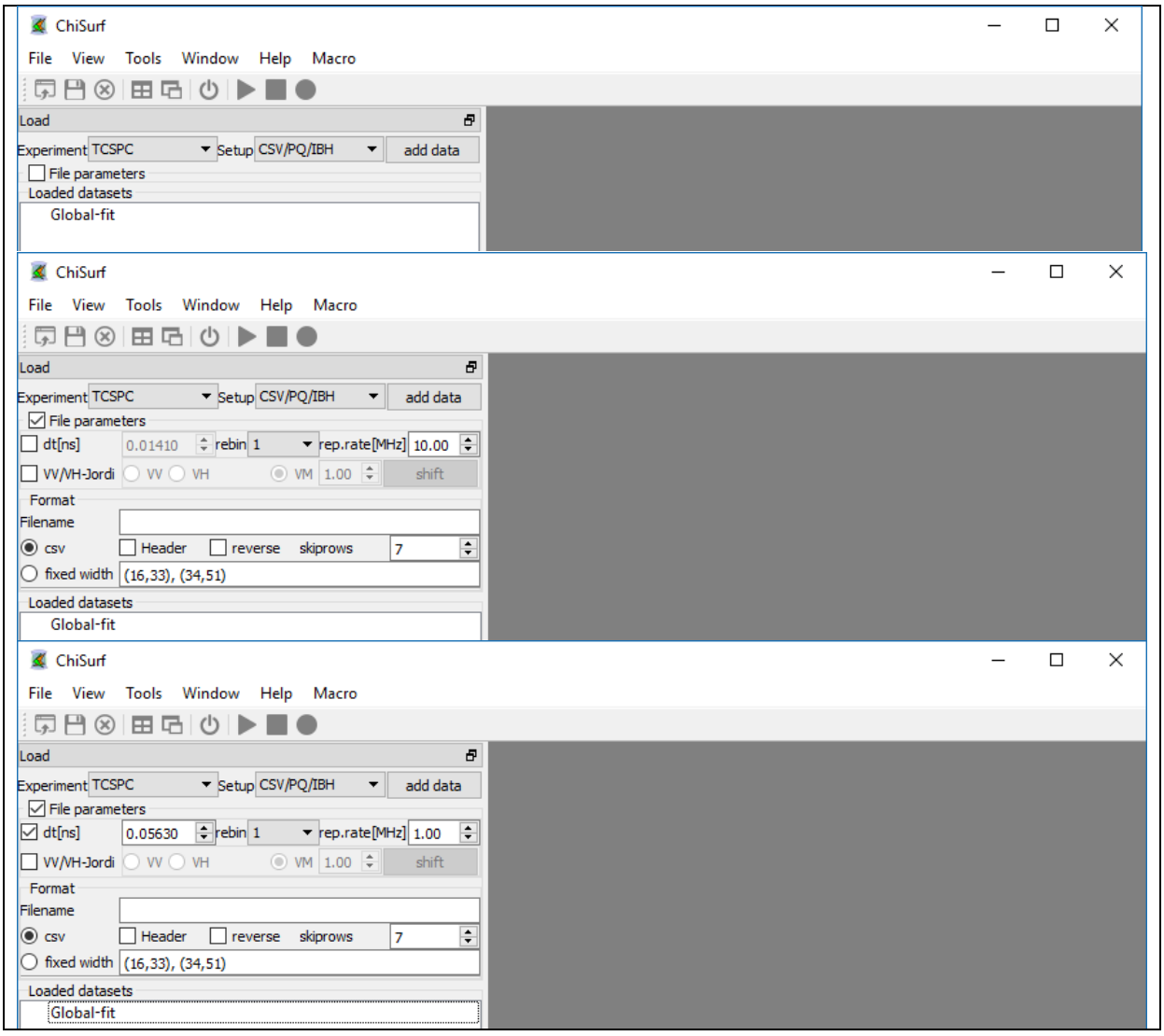
In the Start-Window click in the box next to "File parameters".

Change the default values: dt = 0.0563, rep. rate = 1 MHz.

Other options:

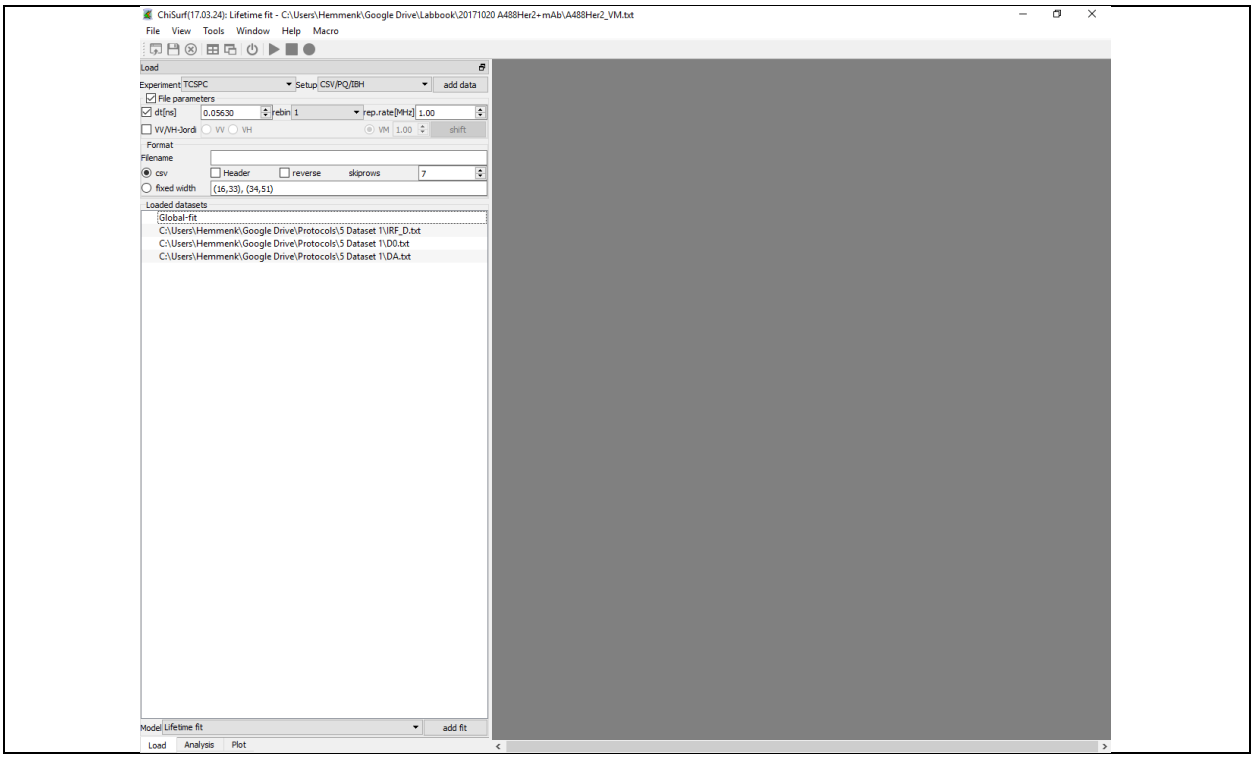
- Rebin: for high-resolution data, the histogram data can be rebinned, e.g. data collected at a dt of 4 ps can be changed to 16 ps binning by a 4fold rebinning.
- VV/VH-Jordi: mainly for single-molecule data, where one detector is placed parallel and the other is placed vertically to exciting laser beam of the microscope. Here, the data file contains first the decay in VV direction, followed by a blank line, and subsequently the VH data. The VM decay used for fitting is reconstructed in ChiSurf based on (I) the experimentally determined g-factor and (II) Japanese (lens) correction factor correcting for the un-ideality of the lenses/optical pieces in the beam path.
- Csv/ fixed width: In this section you can define your data format, if it is different than the IBH or PQ format

Concerning dt: Check your data files, what is **"x", the channel number (standard IBH)** or the time (PQ format). If x = time, leave dt unclicked!!! The software then knows, that the first column in your data is the time. If you click on dt and specify a value, it assumes that x = channel number and multiplies your x-values with the specified dt.

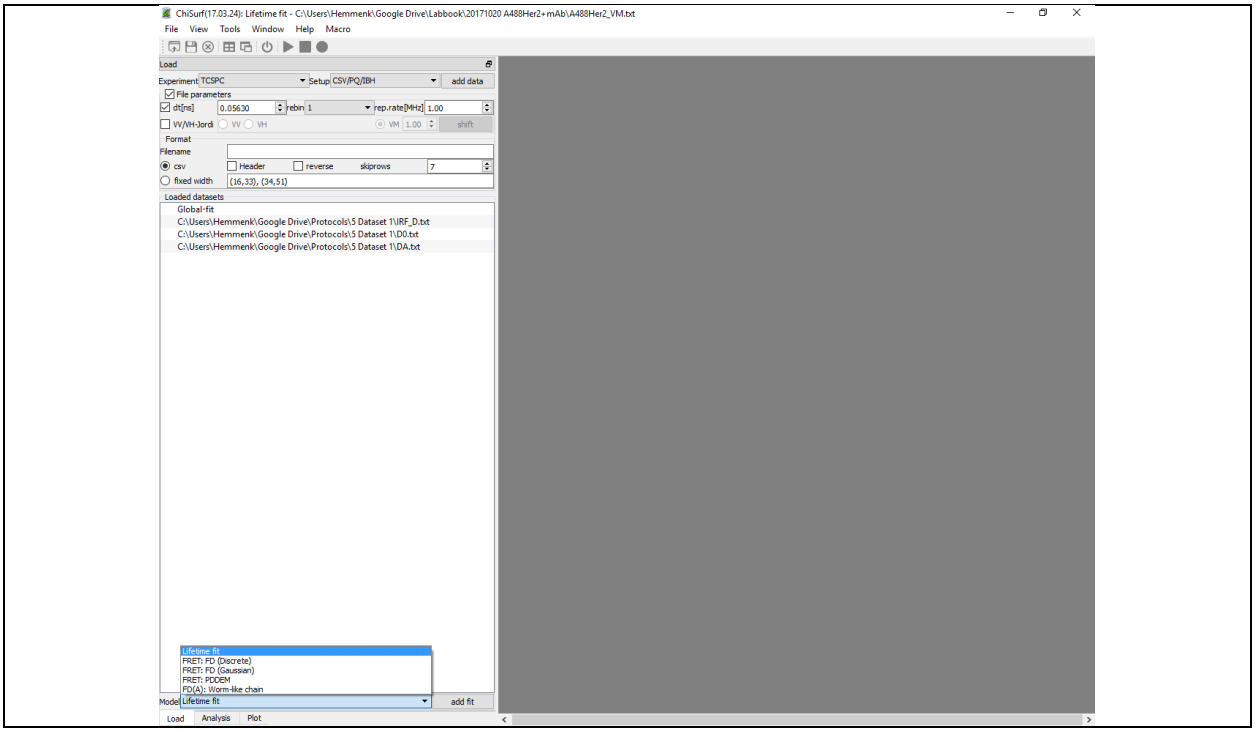


Load your data either by (I) Drag-and-drop or by clicking on “add data”. Generally, I prefer to load the IRF first, then D0, then DA. If your data path is quite long, you can adjust the width of the data/analysis pane, by dragging to the right.

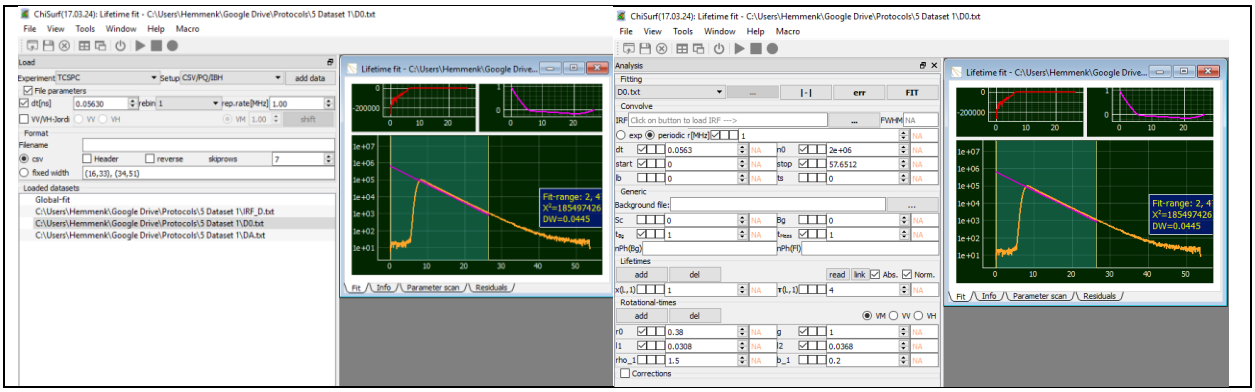
In the next steps, you will need three data files from **data set 1: “IRF_D”, “D0”, and “DA”**.



To start fitting, select data file (not IRF!), select the model (bottom line) and press add fit.



For the D0 sample, the lifetime fit is selected. Change from the “Load” tab to the “Analysis” tab. You can maximize the fit window, if you prefer larger plot view. However, as normally several data sets will be fitted, I prefer to keep them small, so that I can arrange them all next to each other in the gray pane.



Top to bottom explanation of parameters:

- Top line: Selected data set, "|-|" tries to determine optimal fitting range automatically, "err" runs uncertainty estimation, "Fit" fits dataset
- Convolve: click on "…" button to load the IRF, will also calculate FHWM of IRF
- Repetition rate: can be set to exponential or periodic, periodic is the better choice, especially for single-molecule data, where usually the TAC window is so short, that the decay is not finished before next excitation pulse comes (i.e. signal did not decay to background)
- Dt: channel width, start & stop: size of TAC window, n0: extrapolated number of photons if the decay is extrapolated to zero time => do not change any of these values!
- Lb: background/offset of IRF
- Ts: time shift between IRF and data set (some shifts occur either over time (morning vs evening measurement) or due to count rate dependent response of detector)

- Background file: only needed for single-molecule fitting via "Burst-integrated Fluorescence Lifetime Scatter" Method (BIFL-Scatter), then the values tBG and tMeas need to be defined (total measurement time of IRF and data in seconds) from where the number of photons stemming from the background and fluorescence sample in the single-molecule experiment are calculated automatically -> leave empty for ensemble data
- Sc: "fraction" of scattered light, must be positive (or zero)
- Bg: background/offset of data

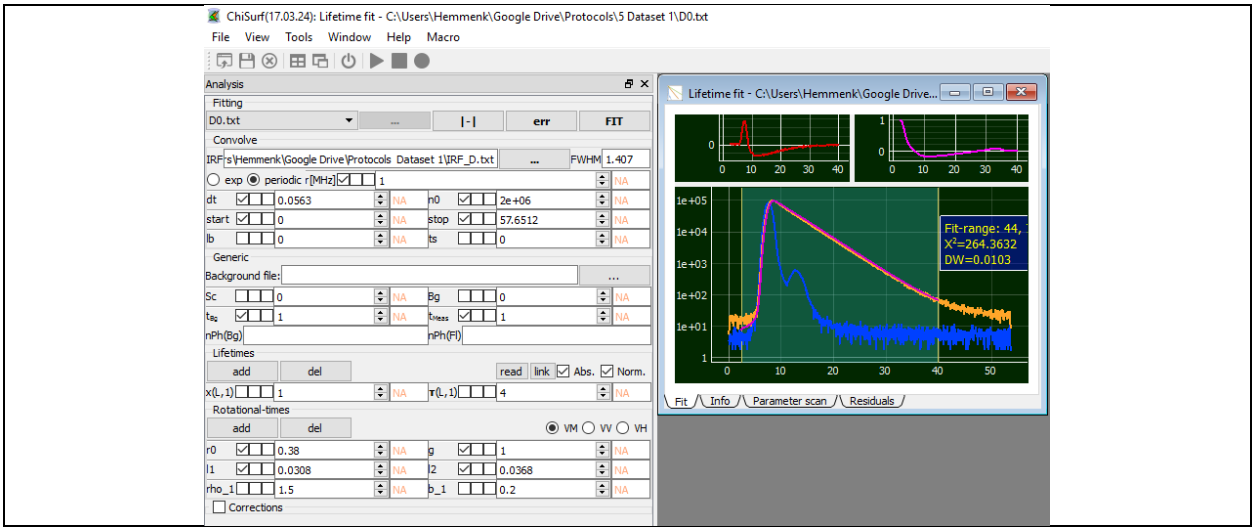
- Lifetimes: for multi-exponential decay, normally absolute and normalized fractions of the individual lifetime components are preferred (i.e. sum of fractions equals 1)
- x and tau are species fraction and lifetime, more or less components can be added by clicking on "Add" or "Del"

- Rotational times: **IGNORE, fitting of anisotropy does not yet work**
- Corrections: Only needed when setup shows differential non-linearities (i.e. when white light is collected and not a smooth line, but sth "wiggly" is obtained), uses then a linearization file and linearization procedure

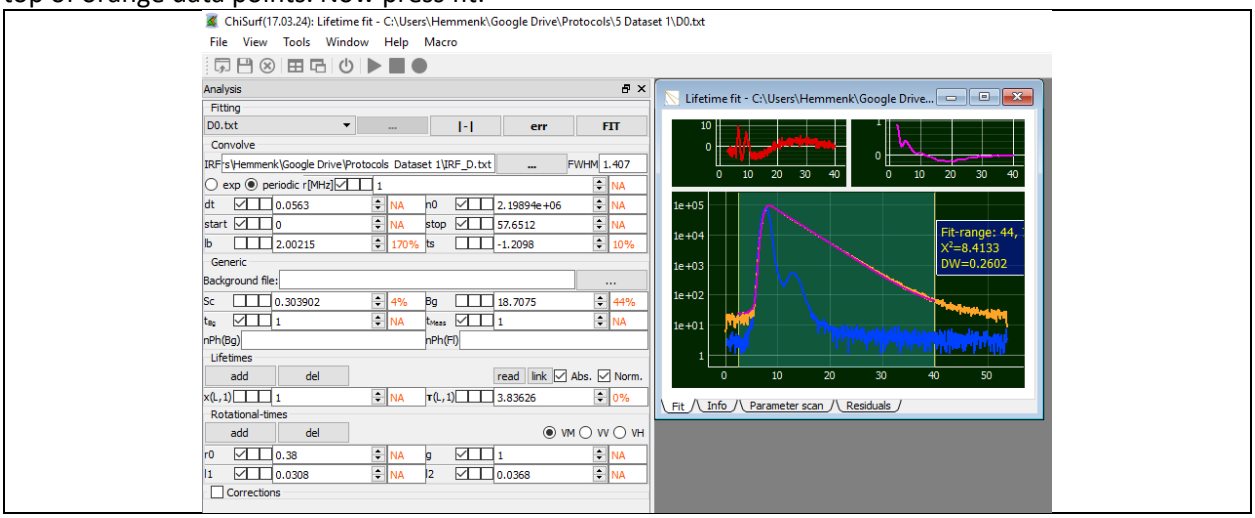
You might have noticed that next to each parameter are three checkboxes:

- Left: if checked, it fixes the parameter to displayed value
- Middle: enables bounds; if checked, two more boxes appear, in which you can give the minimal & maximal value the parameter can have, e.g. offset for IRF & data must be positive
- Right: used for linking of datasets and for global fitting

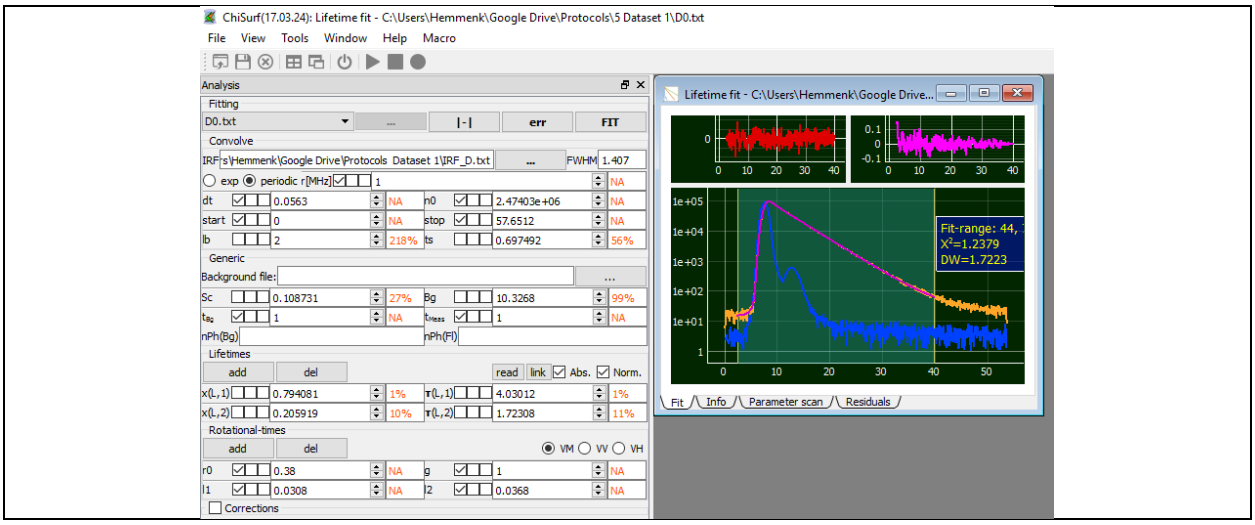
After going through all this parameter, let's now load the IRF and adjust the fit range by pulling the yellow borders in the plot such they encompass a part of the background (for offset determination) until ~ 40 ns. The last channel used for fitting should still have at least 10-100 more photons than the background/offset.



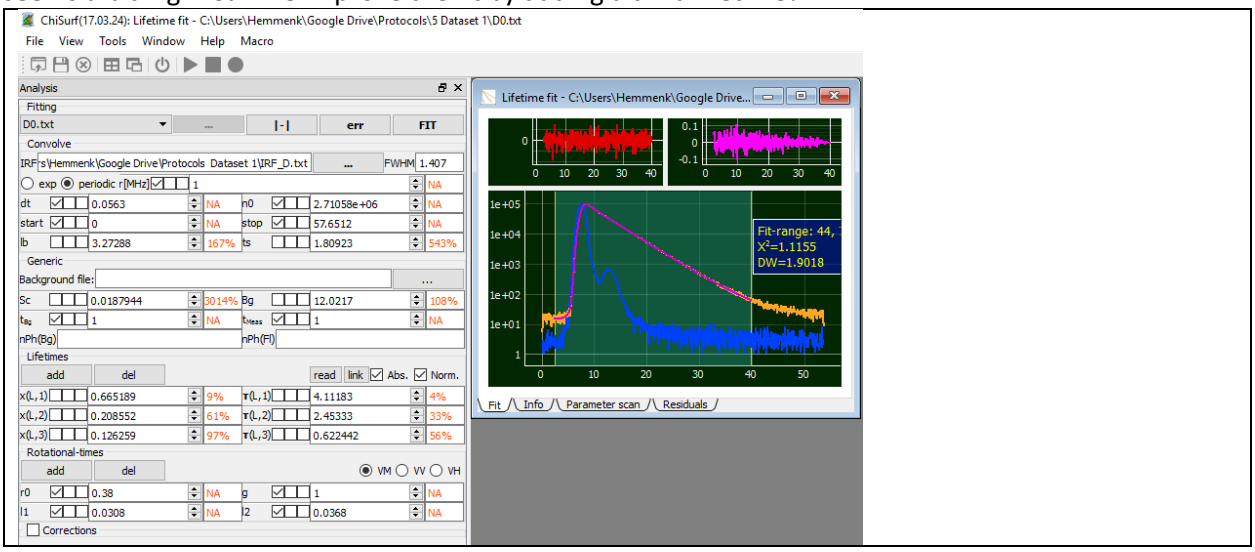
The IRF has been added to the plot window. Now, by looking at the beginning (baseline) of the two curves add the approximate "lb" and "bg" values. For the data set used, 5 and 25 seem to fit: Pink fit line lies on top of orange data points. Now press fit:



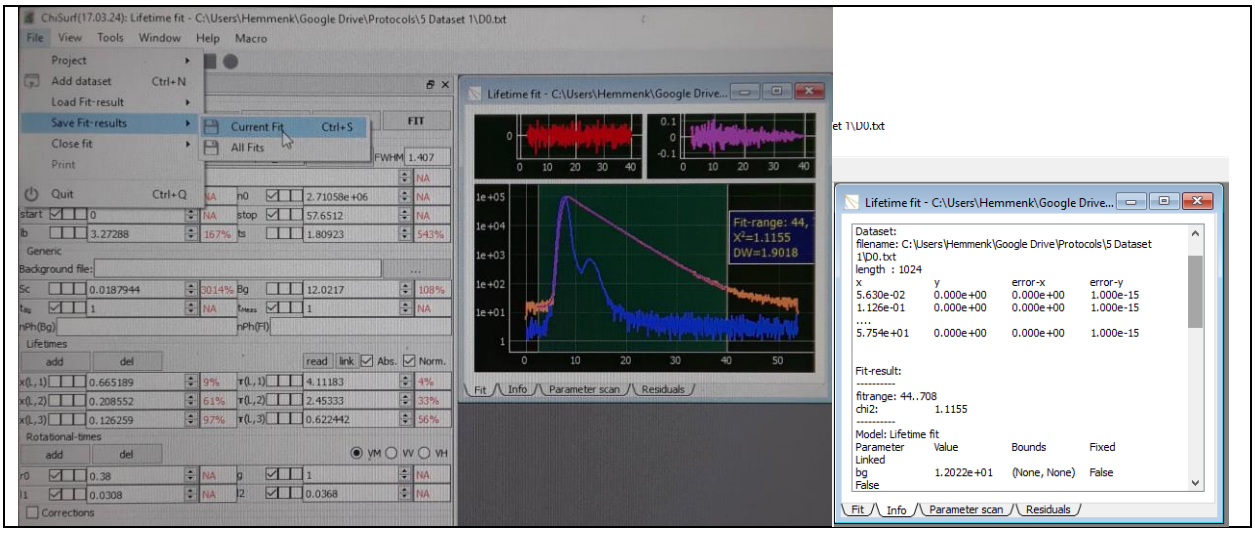
One lifetime is definitely not sufficient, so, we add a second one and repeat the fit. Based on experience, it works well, if for the new lifetime a value half as large as the found one is given as assumption, here ~ 1.9 ns.



Now the residuals (red) and autocorrelation function (pink) already look much better, but a chi2 of 1.2379 seems a bit high. Can we improve the fit by adding a third lifetime?



This fit looks fine for me and I would accept it. The fit results can be saved by either pressing “ctrl+S” or by going to file -> save fit results -> current fit



The following items are saved:

- Word-document summarizing the results including screenshot & data paths
- Screenshot of Fit and Model
- Data file
- Fit and weighted residual
- Info file, what is contained in the info file can be seen by clicking on the “info” tab below the plot

Next to all fit parameter also the species-averaged (<tau>x) and fluorescence-weighted (<tauf>F) averaged lifetime are reported:

$$\langle \tau \rangle_x = \sum_{i+1}^{n} x_i \cdot \tau_i \qquad \text{Eq. 19}$$

$$\langle \tau \rangle_f = \frac{1}{\langle \tau \rangle_x} \sum_{i+1}^{n} x_i \cdot \tau_i^2 \qquad \text{Eq. 20}$$

The saved fit results can be loaded any time again by clicking File-> Load result -> load result in current fit, but:

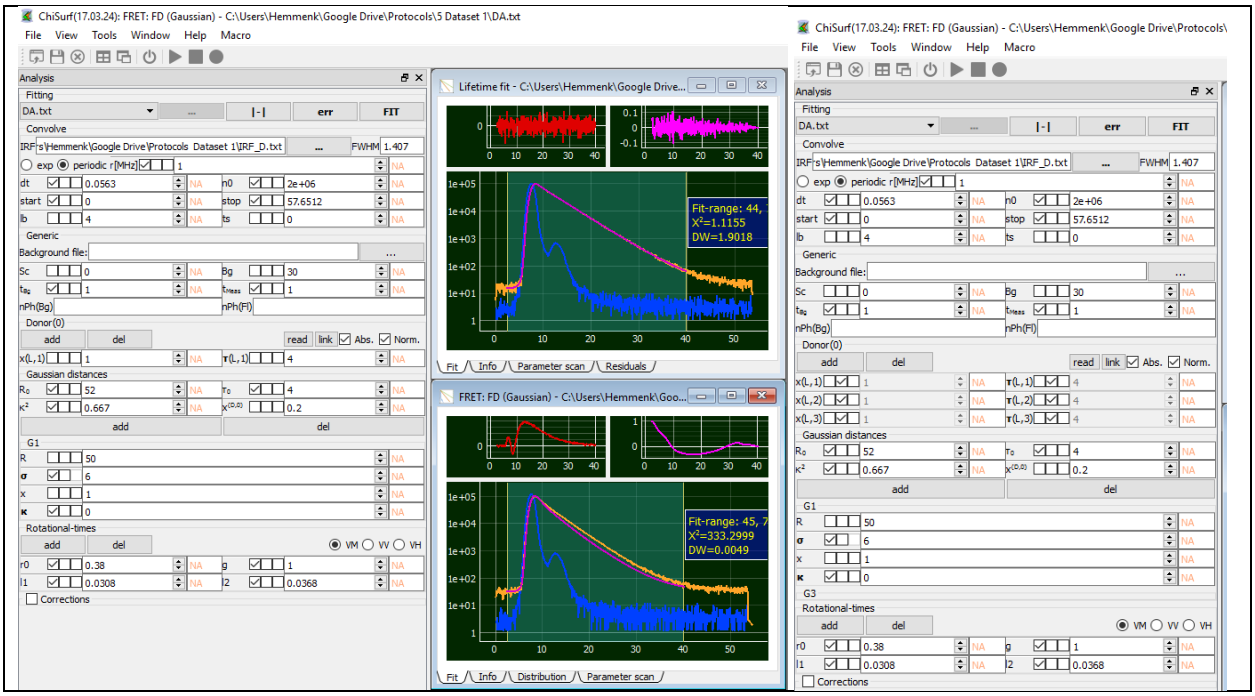
1. You need to load the data first
2. Make sure you give the same number of parameters as were in your fit (Don’t try to load a fit with three lifetimes, but having only two lifetimes in your model!)

Distance distribution fit of DA sample

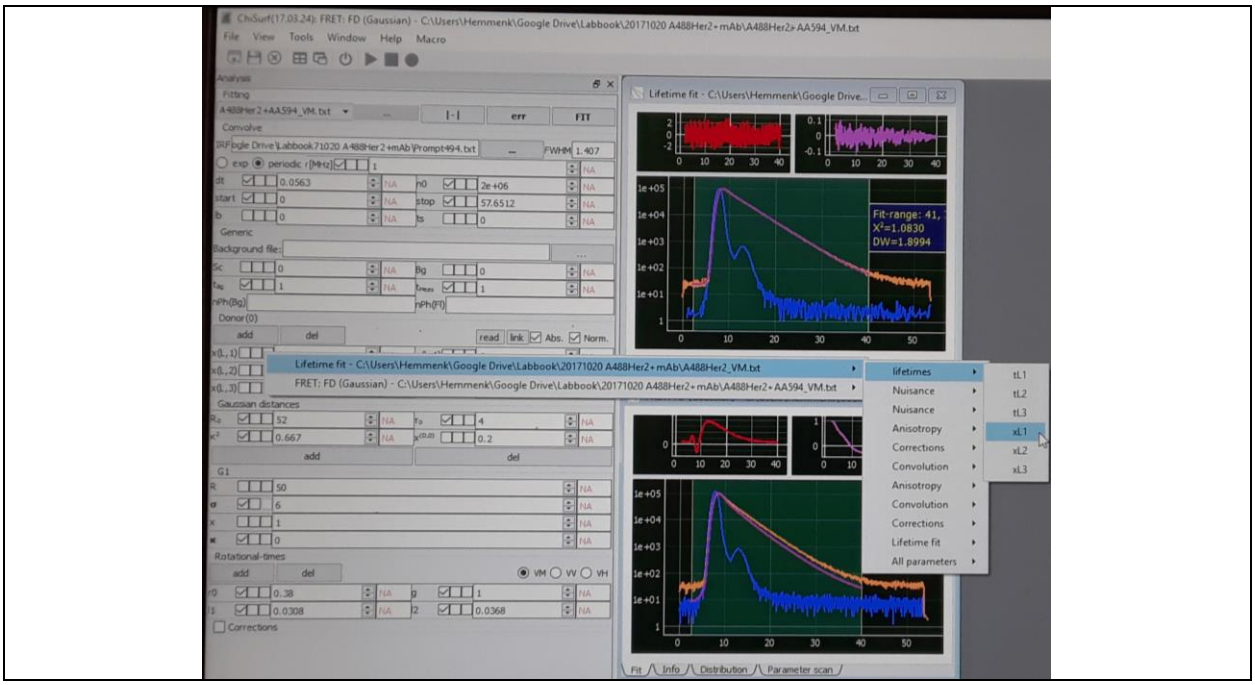
After we have identified the characteristic fluorescence lifetimes of the donor in absence of acceptor (only local quenching due to biomolecule environment), we can load the DA sample and try to determine the D-A distance distribution.

I highly recommend to fit the DA data also with a lifetime model first and compare the fluorescence weighted and species-averaged fluorescence lifetimes of the DA sample with the D0 sample. Only if a significant shortening of the DA lifetime with respect to the D0 lifetime is visible, it is worth trying to fit distance distributions.

The used fit model here is “FRET: FD(Gaussian)”. Add the fit, load the IRF, adjust the fit range and give the background estimates as for the D0 sample. Then we need to add the donor characteristic lifetimes by (I) adding manually (copy-paste) or by (II) linking the data sets.



For both cases, we first need to add more lifetimes in “Donor(0)” section by clicking on “add”. For linking, you right-click – while hovering with the mouse above the three checkboxes – and then select the fit from which you would like to use a parameter and the variable of the other fit which you would like to link to the variable in this fit.

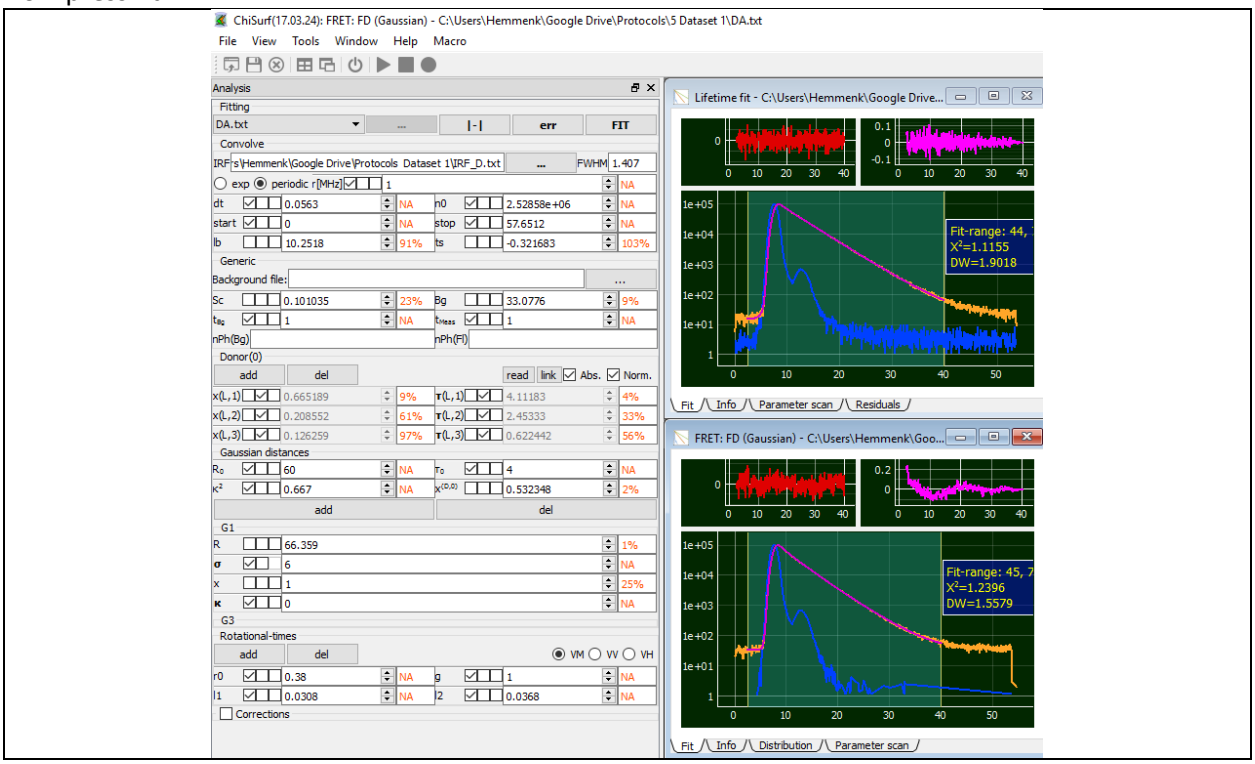


After linking, the number box will turn grayish, however, the value will – visibly - only update once the fit is performed.

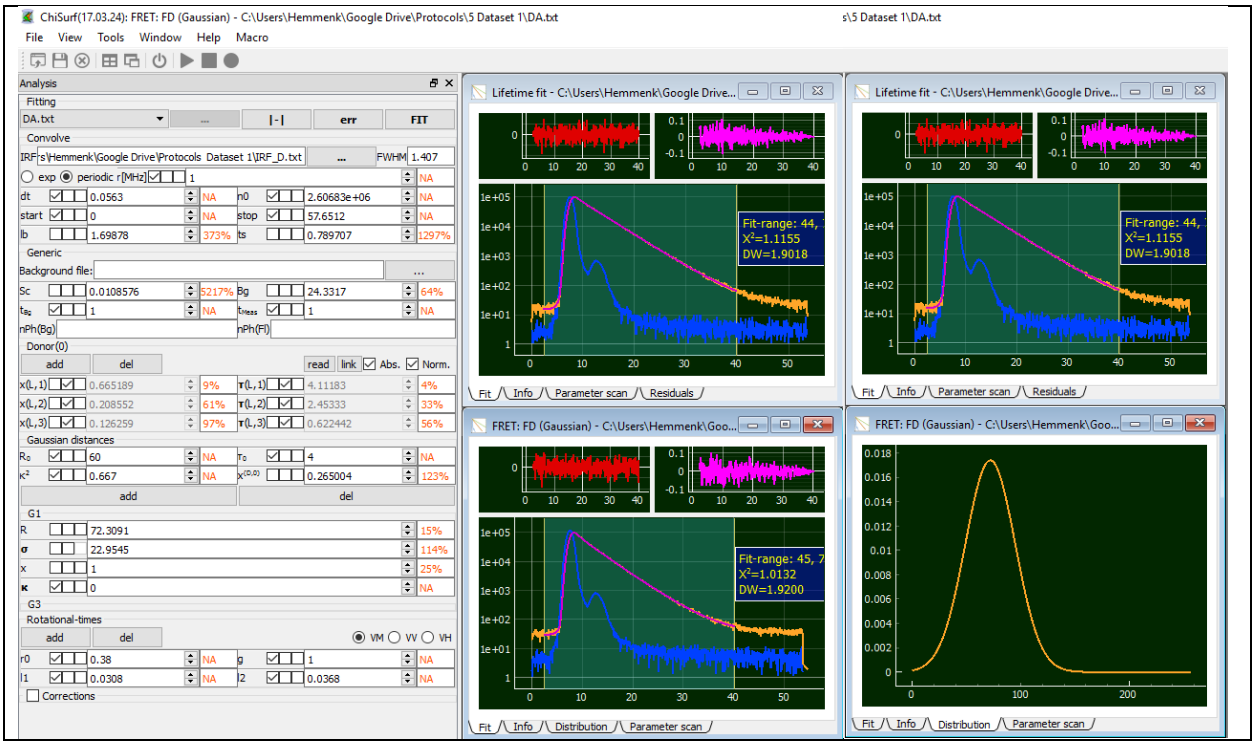
Next, adjust the parameter in the “Gaussian distances” section:

- R0 is the Förster radius, standard is 52 Å for the FRET pair Alexa488-Alexa647, the training data set is labeled with Alexa488-Alexa594, which has an R0 of **60 Å**
- Tau0 is the fluorescence lifetime of the free donor fluorophore (literature value, connected to the literature value of R0)
- Kappa2 is the orientation factor, equals 2/3 for freely moving fluorophores
- X(D,0): fraction ("contamination") of single donor labeled molecule in the DA sample
- G1 defines the first distance, Gaussian/Normal distribution with mean value R, half-width sigma, fraction x and skewness kappa (0 = symmetric, no tailing)
    - By standard, the half width is set to 6 Å, half-widths much larger indicate heterogeneity on the ns timescale
    - 6 Å is approximately the spatial width, which the dyes can explore based on AV Simulations

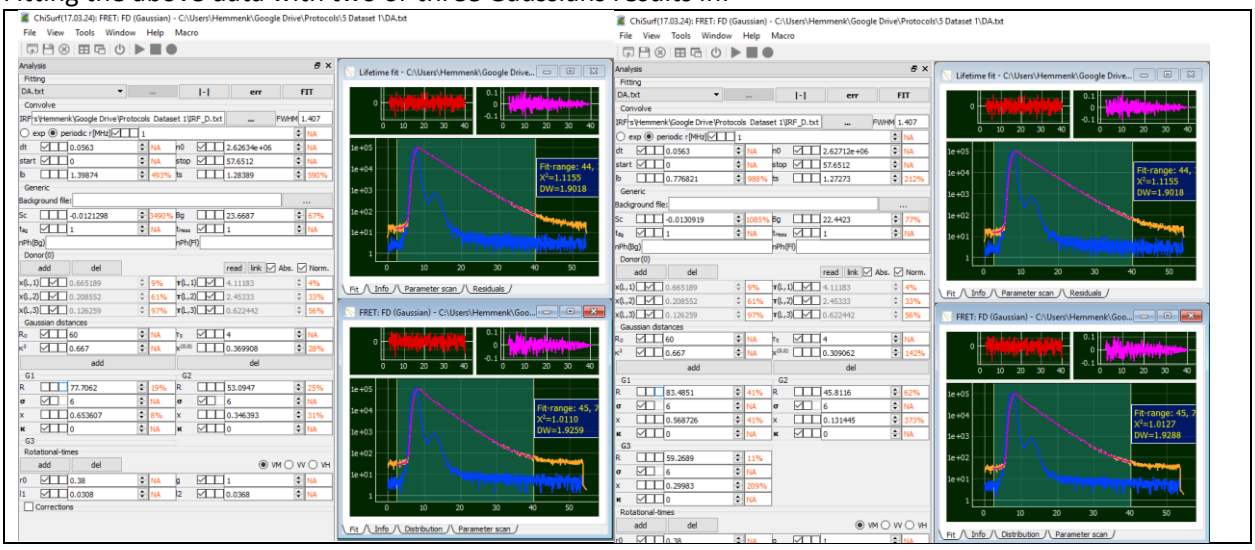
Now press fit:



As we fixed the half-width to 6 Å, the fit firstly defines the "center" of the distance distribution, in a second round, we also leave the half-width free for fitting:



Now the residuals look flat and the chi2 is fine, but the distribution is very broad. If a biomolecule samples discrete states and interconverts “slowly” between (i.e. the states are stable for several hundreds of ns and well separated), a single broad Gaussian is not sufficient for fitting. Instead, multiple Gaussians are needed to describe the data. Here, it is common to use the fixed half-width of 6 Å.
However, due to (I) the distance distribution already of a single state (flexibly connected dyes) and (II) the limited, accessible distance range of a single FRET pair (~ 0.5 – 1.5 R0), more than three Gaussians are difficult. By clicking on the tab “distribution” you can have a look at the fitted distance distribution.
Fitting the above data with two or three Gaussians results in:



Adding a third distance does not improve the chi2, however a second distance seems to improve the chi2 slightly. Now, which fit should we accept as final fit? This depends upon the knowledge of your

biomolecule. Is it expected to find two distances, e.g. a bound and an unbound state? Or is only one broad distribution anticipated, this would be e.g. the case in an unfolded or intrinsically disordered protein.

The following section will show you two possible ways to eliminate one or the other model. First, the next section will show you how to fit data from several data sets globally/jointly, and after that we will make use of the acceptor-sensitized/FRET decay (donor is excited, acceptor emission recorded) and will fit this decay jointly with the D0, A0, and DA decay.

Global fitting of several data sets

Global or joint fitting of several data sets reduces the number of degrees of freedom significantly and thus stabilizes the result. Global fitting can be performed for example when capturing the shift in conformational equilibrium upon ligand or substrate binding. Here, the distances of the "free" and "bound" state are fit jointly, while the species fraction reflect the binding equilibrium. Another application is given in a network of FRET distances, where the FRET pairs are distributed over the biomolecules surface to capture its spatiotemporal motions. In this case, the species fractions of the individual distances are fit jointly, while each FRET pair gets its own distances.

In the example given below, we will analyze how the distance distribution shifts upon ligand binding (data set 2).

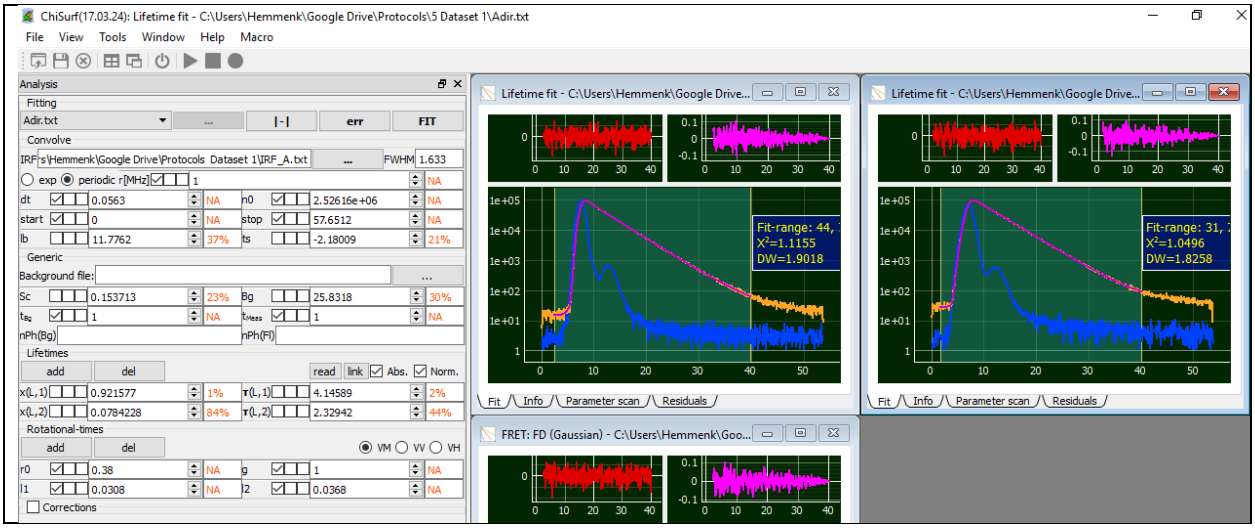
(Explain how to do linking for global fitting, how to add global fit and do global fitting)

Global fitting of Acceptor-sensitized decays (PDDEM)

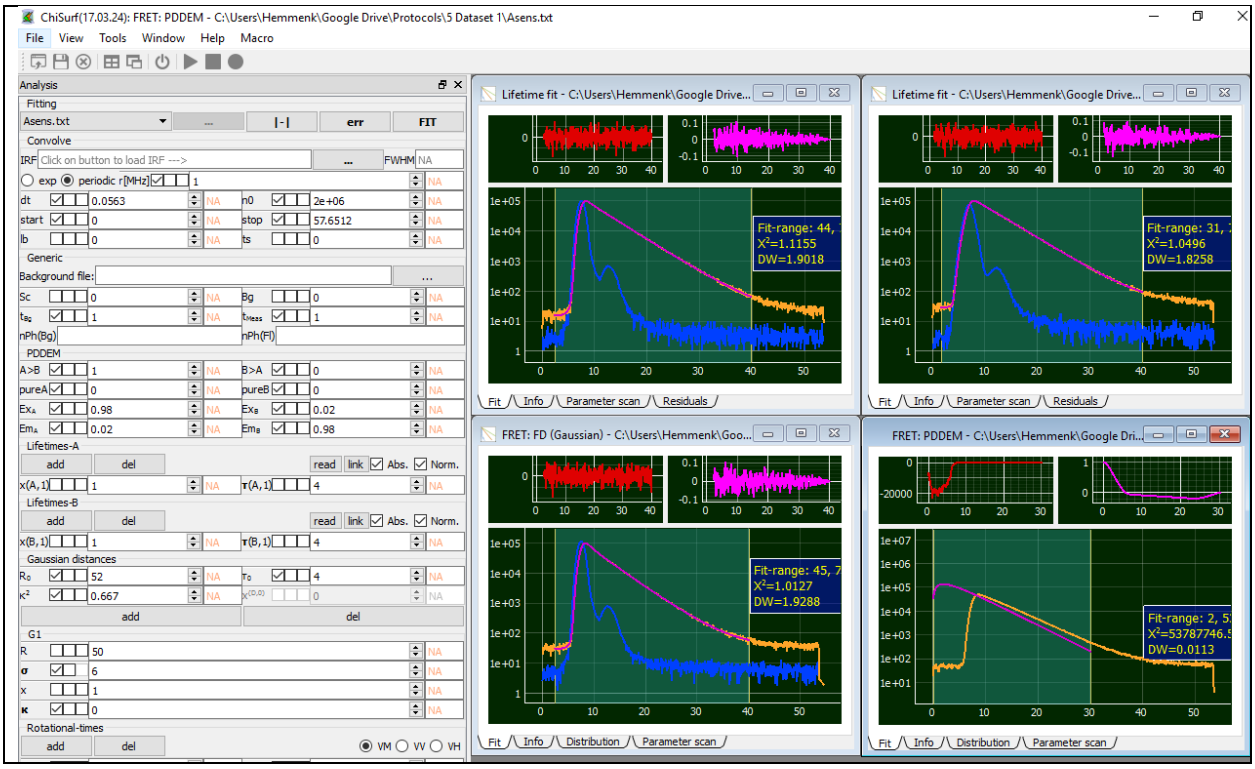
In this section, we will fit all four possible decays, which one can record from a single FRET pair globally: Donor-only and Acceptor-only decays, which characterize the local quenching of fluorophores, and from the DA sample (i) the decay of donor fluorescence after donor excitation "DA" and (ii) the decay of acceptor fluorescence after donor excitation "Asens" (acceptor-sensitized emission).

Load the three additional files of dataset 1: "IRF_A", "Adir" and "Asens".

Firstly, the fluorescence decay of the acceptor-only Adir sample has to be fitted with a multi-exponential lifetime model as described for the D0 sample above. Caution: As IRF of course the IRF of the acceptor NanoLED has to be taken: "IRF_A". Two lifetimes seem to be sufficient.



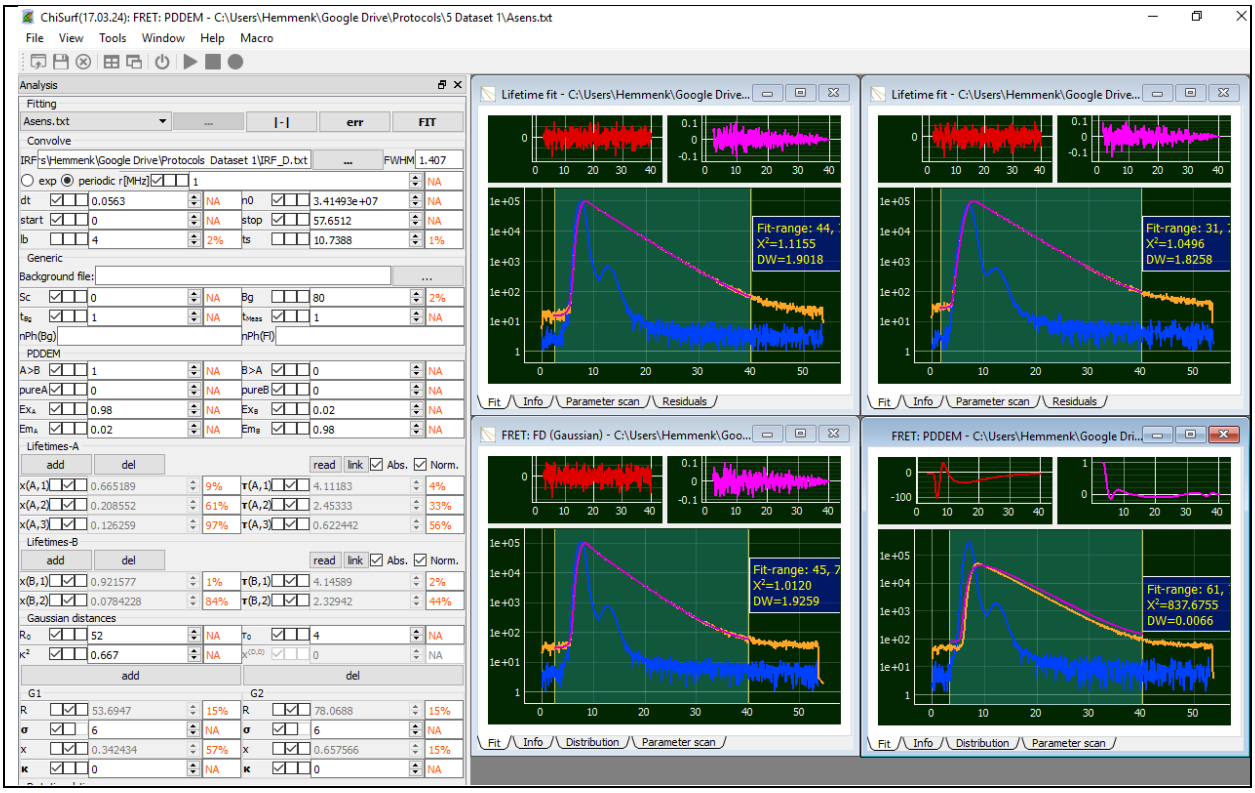
Next, select the “Asens” file and add the “FRET:PDDEM” model. As described in the introductory section, this model is quite complex and needs a lot of information from the other three fits. The convention in this model is that “A” is the donor fluorophore and “B” the acceptor fluorophore.



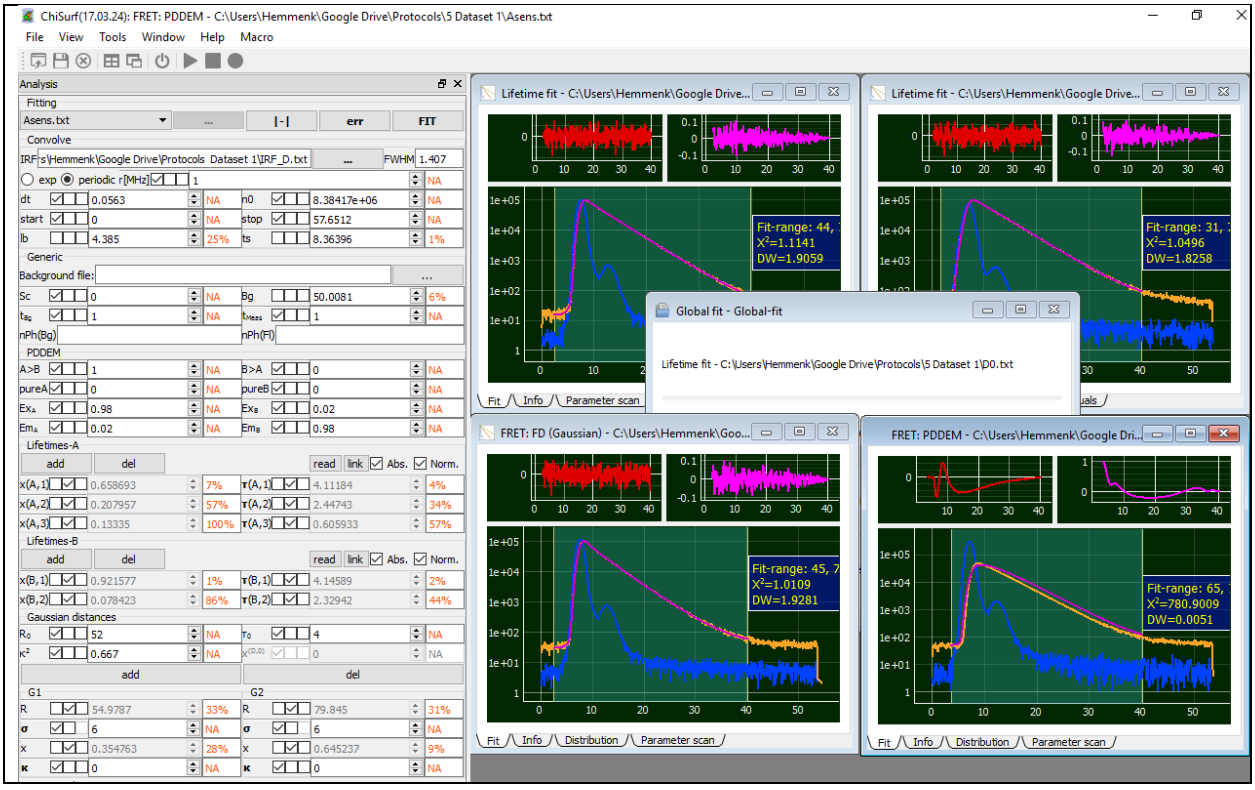
Top to bottom explanation of parameter:

- Convolve section: IRF is the IRF from donor channel (depends upon exciting light source!), otherwise identical to other fit model
- Generic: identical to other fit model, however, in general the lamp scatter “Sc” is 0
- PDDEM: defines how “clean” the decays are, i.e. whether there is spectral crosstalk from the donor to the acceptor channel or direct acceptor excitation by the NanoLED
    - “A>B”:
    - “B>A”:
    - “pureA/pureB”:
    - “ExA/ExB: excitation probability of donor/acceptor at the used NanoLED wavelength
    - “EmA/EmB”: emission probability of donor/acceptor at the used detection wavelength
- Lifetimes-A: In this section, the lifetimes of D0 sample will be linked to the PDDEM model
- Lifetimes-B: In this section, the lifetimes of Adir sample will be linked to the PDDEM model
- Gaussian distance: The variable here will be linked to the Gaussian distance section of DA sample

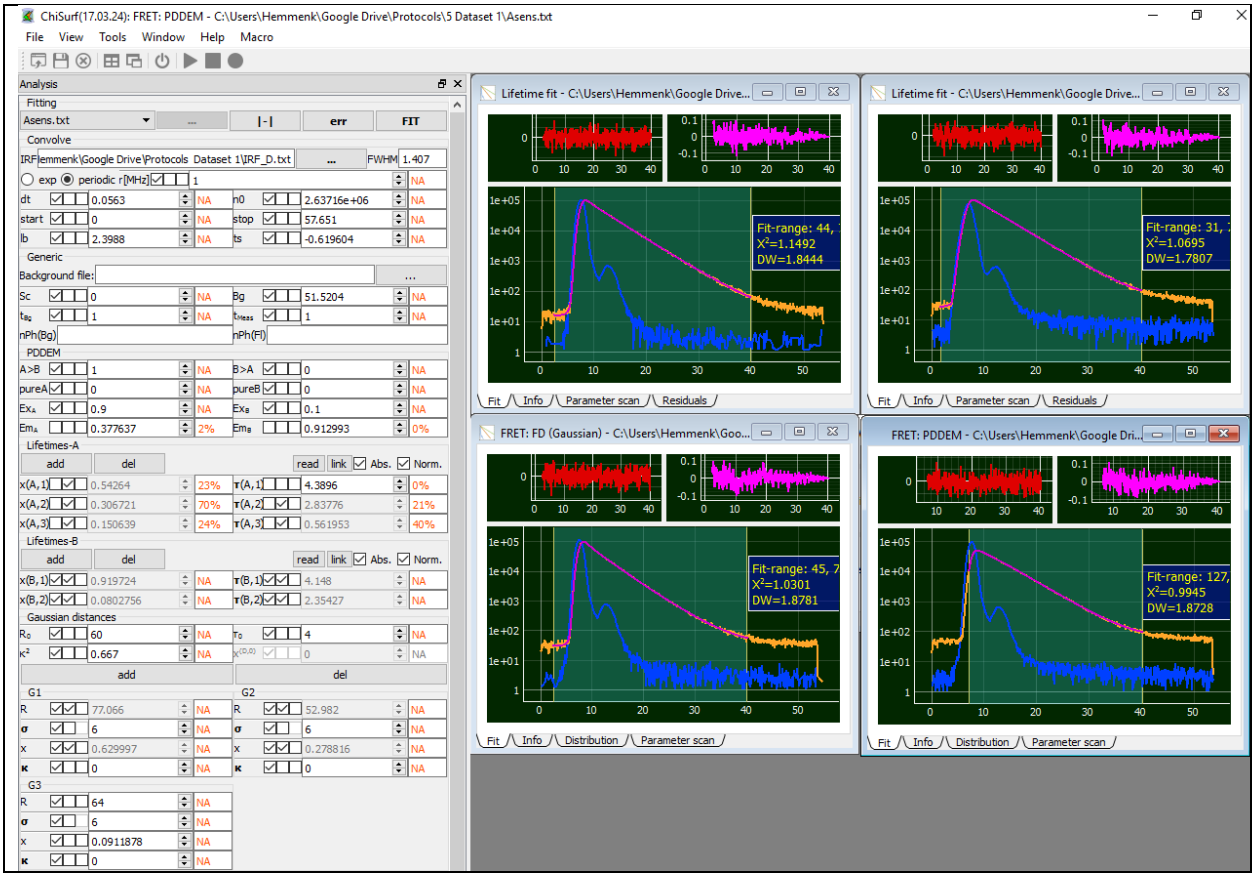
Now link the lifetimes from the D0 and Adir model to Asens and also link the distance and species fraction from DA model (2 distances fit) to Asens, press fit so that the linked parameter fields are updated and you can verify that really all variable are linked as they should be linked – most likely the fit will look horrible!



Now, add a global fit and fit and try to fit the data globally:



Fitting these four decays globally is not trivial. Basically, you go in rounds, optimize them singly, then start a global fit, then single again. It might be that for D0 or Adir you need more lifetimes in the global fit than in the single fit. Also, you might need more distances in the global than in the single fit. The Asens data also finds very long distances, which might be "overseen" in the DA decay and attributed to the D0 fraction. After many rounds and try-outs, you might end up with this fit:



(for your information, it took me 2 hours to get this final fit! Don't give up too fast!)

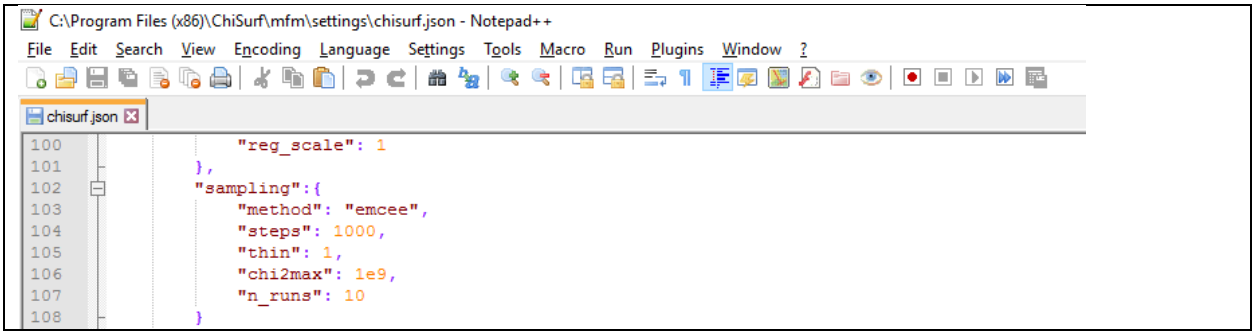
Uncertainty estimation using "err" option

The statistical errors of the DA-distances can be determined by sampling the parameter space and applying the *F*-distribution at a confidence level of 95% given the minimum determined χ^2. The maximum allowed $\chi^2_{r,\max}$ for a given confidence-level (*P;* e.g. for 2σ *P* = 0.95) was calculated by

$$\chi^2_{r,\max}(P) = \chi^2_{r,\min} \cdot \left[1 + n/\upsilon \cdot \mathrm{cdf}^{-1}(F(n,\upsilon,P))\right], \qquad \text{Eq. 21}$$

where $\mathrm{cdf}^{-1}(F(n,\nu,P))$ is the inverse of the cumulative distribution function of the *F*-distribution for *n* number of free parameters, and with ν degrees of freedom. $\chi^2_{r,\min}$ is the minimum determined χ_r^2. This option may also be used to conduct a support plane analysis, in which a parameter is systematically varied. (Literature: Goodman J & Weare J (2010) Ensemble Samplers with Affine Invariance. *Comm App Math Com Sc* 5(1):65-80., Foreman-Mackey D, Hogg DW, Lang D, & Goodman J (2013) emcee: The MCMC Hammer. *Publ Astron Soc Pac* 125(925):306-312., Lakowicz JR (2006) *Principles of Fluorescence Spectroscopy* (Springer, New York); Third Ed., Straume M, Frasier-Cadoret SG, & Johnson ML (1991) Topics in Fluorescence Spectroscopy, Principles. ed Lakowicz JR (Plenum Press), Vol 2, pp 177-240.)

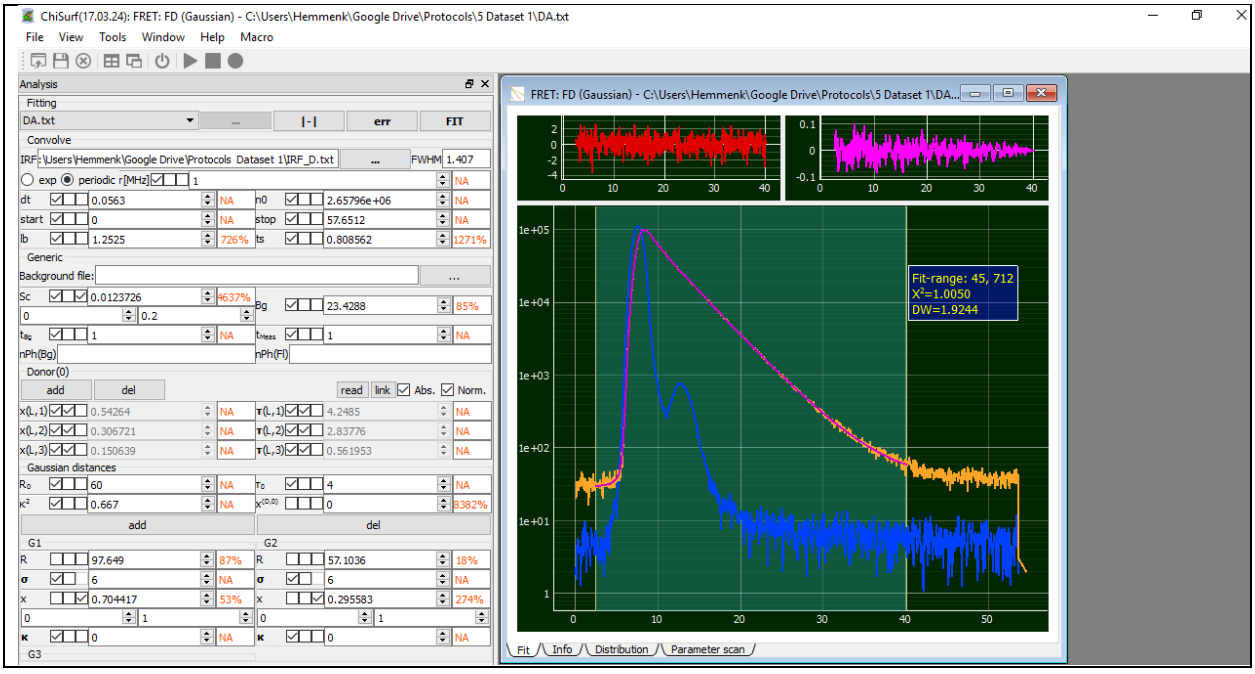
The standard setting for uncertainty estimation can be found in the following file:
C:\Program Files (x86)\ChiSurf\mfm\settings



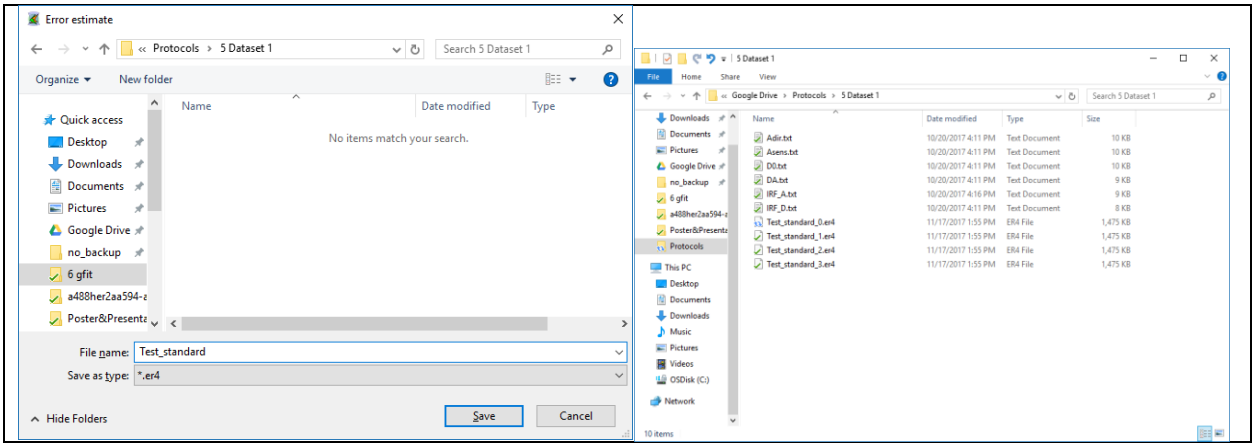
Generally, by standard 10 runs are started with 1'000 steps each. This is in most cases not sufficient for a complete sampling of the chi2 surface. I recommend to use 25 runs with 10'000 steps each.
Before starting the uncertainty estimation, think about what information it should deliver you (which variables are interesting to you? Are some parameters correlated?):

- Lamp scatter, time shift and short distance (and the short distance fraction) are highly correlated
- Long distance and D0-fraction are highly correlated
- Every free parameter more will increase the sampling, fix everything, which is not needed to be variable, e.g. background of IRF and data
- Species fraction: give boundaries so that they don't get negative or very large, of note: while sampling, the species fractions are not normalized, so far this has to be done manually

Exemplary, we will run now the uncertainty estimation of the two distances fit – once with the standard setting, once with the optimized settings.
We don't have a very short distance, so we can fix sc and ts, however we have a quite long distance, so we should not fix xD0:



After you pressed the button “err”, dialog box will open and ask where to save the runs and under which name. After you have specified the folder and given a name, the runs will start and you can follow – by checking in the folder how many runs are already completed.



If you run the uncertainty estimation according to default settings, it will be fast, while a more thorough sampling will take – dependent upon your PC and number of free parameter – 20 min or more.

For the second run change the parameter to 10’000 steps and 25 runs. For this, you have to close ChiSurf, make changes in the chisurf.json file and open ChiSurf again. Load your saved fit.

Why do you have to close it? ChiSurf reads its setting file when it opens, not while it is running. Alternatively, you can open the “console” and type:

XXXX

The generated files can be opened with **XXX**?