

# A Short Tutorial on DeerAnalysis2006

G. Jeschke

Max Planck Institute for Polymer Research, Postfach 3148, 55021 Mainz, Germany

March 9, 2006

## 1 Preface

This tutorial describes analysis of distance measurements by pulsed ELDOR [1, 2], specifically by the four-pulse DEER experiment [3, 4, 5], with the program DeerAnalysis2006. The program is based on some algorithms discussed previously [6, 7, 8] and on some new algorithms that will be described in a forthcoming paper.

## 2 Worked example: A single well-defined distance

### 2.1 Basic Tikhonov processing

- Load file `deer_3.6nm_50K` DeerAnalysis automatically corrects the phase and tries to determine zero time.
- In the present example, zero time should be 96 ns, but the program finds 112ns. Correct the value to 96 ns, using the green - button in the **Original data** panel.
- As data were measured until almost complete decay, the **Dipolar evolution** function is rather noisy. Cut the data by repeatedly clicking on the brown ! button in the **Original data** panel.
- Optimize the range for the background fit by clicking on the blue ! button in the **Original data** panel. This will take some time. You can follow progress in the **Status** line (left bottom).
- In the **Distance analysis** panel, select the Tikhonov **reg.** radio button. Click on the **Fit** button right from the radiobutton. Fitting will also take some time, as this data set is very long.

The plot panels should now look as shown in Figure 1.

They provide the following information:

#### Original data

- *Black trace*: real part of the data after phase correction.
- *Magenta trace*: imaginary part of the data after phase correction.
- *Green cursor*: current setting of zero time.
- *Light blue cursor*: start of the range for background fitting.
- *Brown cursor*: cutoff of the data at the end and end of the range for background fitting.
- *Red trace*: fitted background function, solid line within the fit range, dotted line outside the fit range (extrapolation).

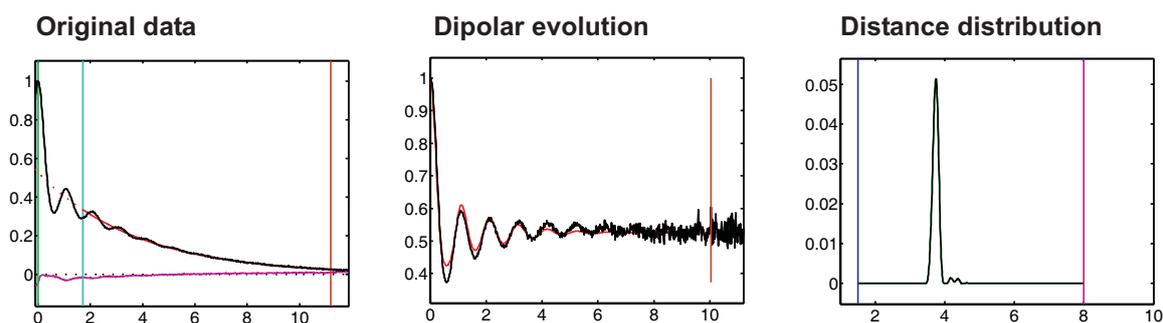


Figure 1: Plots in DeerAnalysis2006 after basic Tikhonov processing of file `deer_3_6nm_50K`.

## Dipolar evolution

- *Black trace*: dipolar evolution function after background correction (divided by background function and renormalized).
- *Red trace*: fit of the dipolar evolution function corresponding to the displayed distance distribution.
- *Brown cursor*: suggested cutoff time.

Note that the suggested cutoff time is different from the cutoff time using in Tikhonov regularization. This is because Tikhonov regularization provides a better fit to the data than the original APT (approximate Pake transformation) result. Thus, the noise level can be estimated more precisely and it can be judged more precisely, at which dipolar evolution time experimental noise becomes too large.

## Distance distribution

- *Black trace*: distance distribution computed by the currently selected technique for data analysis (here Tikhonov regularization).
- *Blue cursor*: lower end of the fit range or range for data analysis.
- *Magenta cursor*: upper end of the fit range or range for data analysis.

The mean distance  $\langle r \rangle$  (3.75 nm) and the width of the distance distribution, given as standard deviation  $s(r)$  (0.12 nm), are displayed in the panel **Distance analysis**.

## 2.2 Inspecting data

The plots can be changed to allow for closer detection. For instance, the **Original data** (real part) are somewhat better seen when switching off the imaginary part (click on the `imaginary` checkbox). The other two plots can be adapted more extensively.

## Dipolar evolution

- *Closer inspection of the initial part*: Click on the `<>` button in the **Zoom** subpanel. Each click expands the initial part of the evolution function by a factor of  $\sqrt{2}$ . The `><` button reduces zoom in the same way. The `!` button resets zoom to 1.0.

- *Dipolar spectrum*: Select the **Spectrum** radiobutton. The Fourier transform of the baseline-corrected dipolar evolution function (black trace) and its fit corresponding to the computed distance distribution (red trace) are shown. The **Zoom** button  $\langle \rangle$  expands the range around zero frequency (Figure 2). The zoom factor can also be input directly in the edit field.

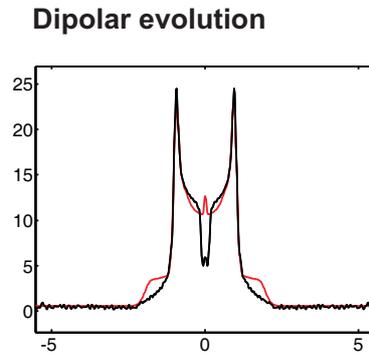


Figure 2: Dipolar spectrum of data set `deer_3_6nm_50K` at zoom factor 5.7.

### Distance distribution

- *Expand*: To inspect the peak more closely, click on the **Expand** checkbox. Only the range between the cursors is now displayed. Select 3 nm in the edit field **Start** and 5 nm in the edit field **End** (Figure 3).

The main peak is situated between 3.5 and 3.9 nm. It is slightly asymmetric (steeper slope towards long distances). There are some further peaks between 4 and 4.5 nm. To exclude them from analysis, select and 4.1 nm in the edit field **End**. The values values in the panel **Distance analysis** change to  $\langle r \rangle = 3.74$  nm and  $s(r) = 0.08$  nm.

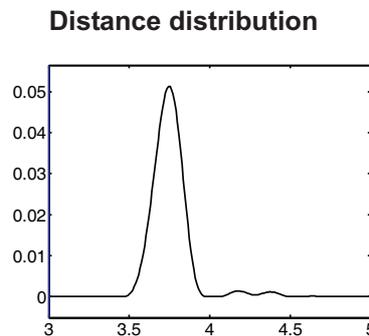


Figure 3: Range of interest from 3.0 to 5.0 nm in the distance distribution for data set `deer_3_6nm_50K`.

- *Testing for the relevance of peaks*: Uncheck the checkbox **expand**. The whole distribution is shown again. Select 4.05 nm in the edit field **Start** and 4.5 nm in the edit field **End**.

Now click on the green **Suppress** button. Select the **Time domain** radiobutton in the panel **Dipolar evolution** and set the zoom to 1.0 (! button) (Figure 4).

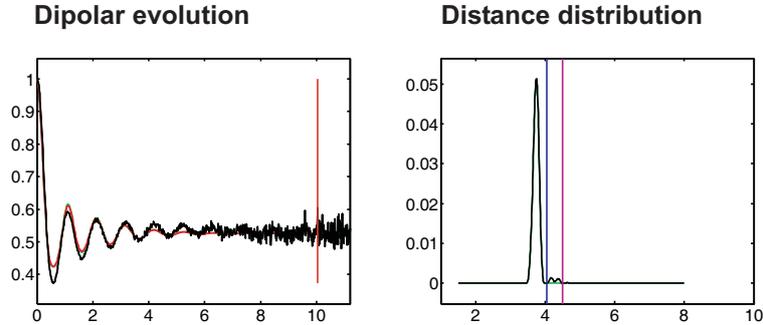


Figure 4: Test for the relevance of small contributions to the distance distribution. The green traces correspond to suppression of the peaks between the blue and magenta cursor.

In the case at hand, suppression of the small peaks does not significantly change the quality of the fit of the **Dipolar evolution** function. The green trace hardly deviates from the red one. These small peaks are thus noise-related.

### 2.3 Using the L curve for optimized Tikhonov regularization

Before attempting a new fit, you need to reset the **Distance distribution** cursors to their default values. This can be done by clicking on the blue and magenta ! buttons in that panel. The values should change to 1.5 and 8.0 nm. In the case at hand, also reoptimize the cutoff time by repeatedly clicking on the brown ! button in the **Original data** panel.

- *Computing the L curve:* Confirm that the radiobutton **Tikhonov reg.** in the **Distance analysis** panel is selected. Select the checkbox **Compute L curve** and click on the **Fit** button. Tikhonov regularization is performed for regularization parameters  $\alpha = 0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000, \text{ and } 100000$ . The status line shows the currently computed data set. This takes several minutes. Finally, the **L curve** checkbox in the **Distance distribution** panel is automatically selected and the L curve displayed. A red dot marks the point where the program locates the corner of the L curve. (Figure 5a)
- *Navigating the L curve:* To select a data set corresponding to a different point in the L curve (different regularization parameter), use the black **Reg. par. +** and **-** buttons in the **Distance distribution** panel. The value of the regularization parameter, the fit in the **Dipolar evolution** plot and the r.m.s. value,  $\langle r \rangle$ , and  $s(r)$  in the **Distance analysis** panel update.

In the case at hand, regularization parameters of 1.0 or larger correspond to overdamping of the **Dipolar evolution** function, i.e., to oversmoothing of the distance distribution. A regularization parameter  $\alpha = 0.1$  is the optimum choice (Figure 5c,d). At even smaller regularization parameters, the fit and r.m.s. value improve only insignificantly. The distance distributions corresponding to all computed values of  $\alpha$  can be inspected by unchecking the **L curve** checkbox in the **Distance distribution** panel and using the black **Reg. par. +** and **-** buttons for selection.

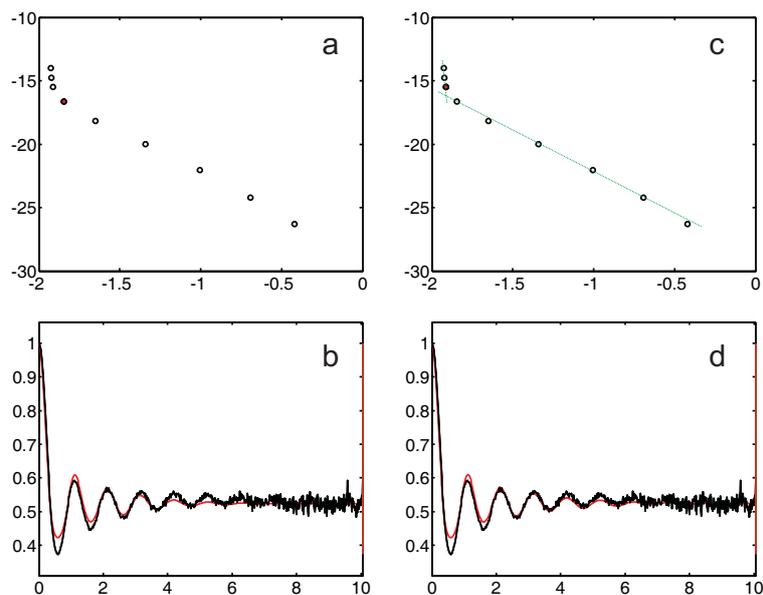


Figure 5: L curve for the data set `deer_3_6nm_50K`. a) Corner suggested by the program (red dot). b) Fit corresponding to selection by the program. c) Corner selected by visual inspection (red dot). The green dotted lines are guides to the eye, added in a drawing program. d) Fit corresponding to visual selection.

### 3 Problem set: Two peaks

#### 3.1 Problem

Perform an analysis of the data set `DEER_mix_50K`, using Tikhonov regularization with L curve computation (Figure 6). In contrast to the example above, do not change the automatically determined zero time. Give

1. the optimum regularization parameter
2. distance and width (standard deviation) for the first peak
3. distance and width (standard deviation) for the second peak

#### 3.2 Answer

1)  $\alpha = 0.1$ . 2)  $\langle r \rangle = 2.85$  nm,  $s(r) = 0.05$  nm. 3)  $\langle r \rangle = 3.74$  nm,  $s(r) = 0.08$  nm. One can argue about the correct interval for the second peak and about suppression of the small peaks. The small peaks are not related to noise, but probably to errors in background correction and to orientation selection artefacts.

## 4 Worked example: Gaussian fitting

### 4.1 Background correction as pre-processing

Whenever a reasonable correction of the intermolecular background can be made, it should be made before fitting the dipolar evolution rather than fitting the intramolecular and intermolecular distance distributions simultaneously. This is demonstrated in the following.

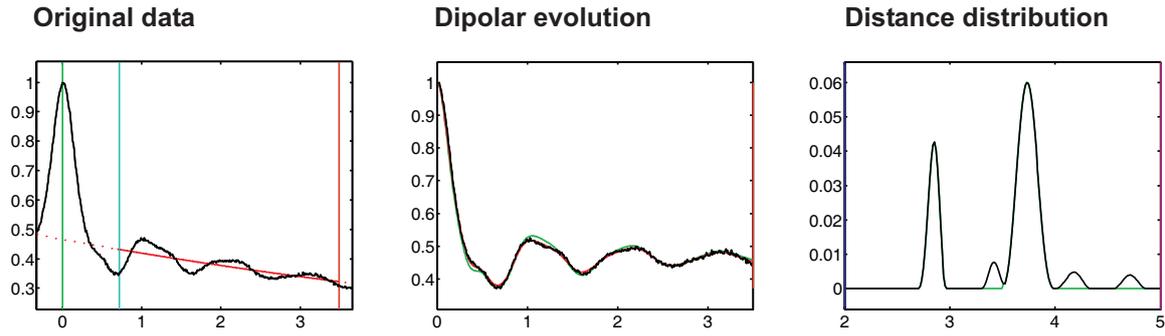


Figure 6: Result of Tikhonov regularization ( $\alpha = 0.1$ ) for the data set DEER\_mix\_50K.

- Load (or reload) data set DEER\_mix\_50K. This resets all controls of DeerAnalysis2006 to their default values.
- In panel **Distance analysis** select the radiobutton **Model fit** and from the popup menu to the right the model **Two Gaussians**. The **Distance distribution** plot displays the result of an APT (black dashed line) and the modelled distribution (red dashed line), using the starting values for the mean distance of the first peak  $\langle r1 \rangle$ , the width (standard deviation) of the first peak  $s(r1)$ , the relative contribution (integral)  $p1$  of the first peak, the mean distance of the second peak  $\langle r2 \rangle$ , and the width (standard deviation) of the second peak  $s(r2)$ . The fit of the **Dipolar evolution** data is shown as a red dotted curve.
- Improve the starting values by direct input into the edit fields. Good starting values decrease the probability to end up in a local minimum of the error hypersurface. Use the black APT curve in the **Distance distribution** plot as a guide for changing the parameters. Good values are  $\langle r1 \rangle = 2.8$  nm,  $s(r1) = 0.2$  nm,  $p1 = 0.3$ ,  $\langle r2 \rangle = 3.75$  nm,  $s(r2) = 0.2$  nm. When finished, click on the **Fit** button in the **Model fit** subpanel. The status line shows how the r.m.s. value changes during fitting.
- When the fit is completed, the **Distance distribution** plot updates to a black solid line, corresponding to the model of two Gaussian peak. The **Dipolar evolution** data corresponding to this distribution are now shown as a red solid line (Figure 7).

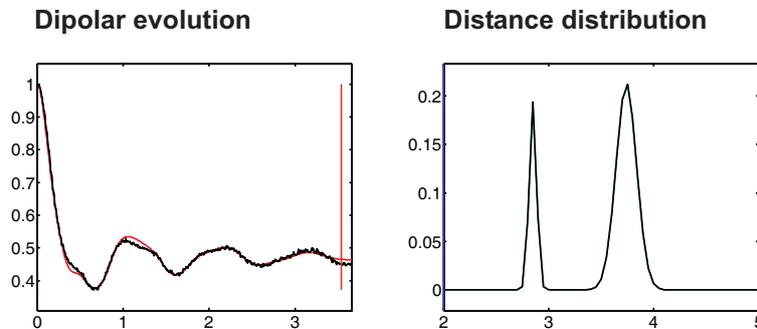


Figure 7: Result of model fitting with two Gaussian peaks for the problem data set DEER\_mix\_50K.

You should find the values  $\langle r1 \rangle = 2.85$  nm,  $s(r1) = 0.05$  nm,  $p1 = 0.2446$ ,  $\langle r2 \rangle = 3.74$  nm,  $s(r2) = 0.14$  nm.

## 4.2 Simultaneous fit of background and distance distribution

- Change the **Background model** to **No correction** and from the popup menu in the **Model fit** subpanel select **Two\_Gaussians\_hom**. Adjust the starting parameters as above. Finally adjust the concentration parameter  $c$ , so that the dotted red fit in the **Dipolar evolution** panel is reasonable. You should find a value of  $c \approx 0.4$ .
- Click on the **Fit** button. With this good set of starting parameters the program finds a reasonable solution (Figure 8:  $\langle r1 \rangle = 2.85$  nm,  $s(r1) = 0.06$  nm,  $p1 = 0.2542$ ,  $\langle r2 \rangle = 3.735$  nm,  $s(r2) = 0.17$  nm,  $c = 3636$ ).

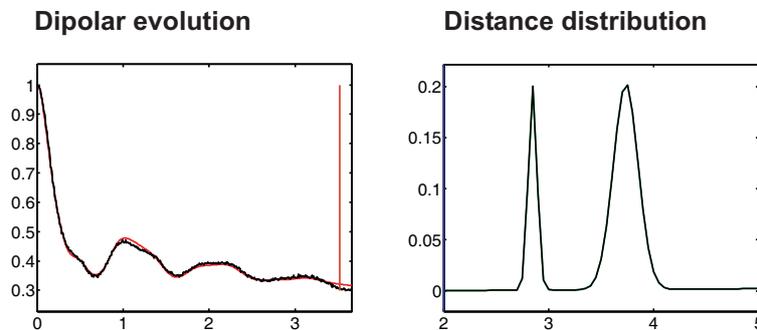


Figure 8: Result of model fitting with two Gaussian peaks plus homogeneous background for the problem data set `DEER_mix_50K`.

Simultaneous background fitting is, however, less stable.

- Again fit the model **Two\_Gaussians** using the **Homogeneous Background model** with three dimensions, but now with the slightly worse starting parameters  $\langle r1 \rangle = 2.7$  nm,  $s(r1) = 0.3$  nm,  $p1 = 0.5$ ,  $\langle r2 \rangle = 4.0$  nm,  $s(r2) = 0.3$  nm. The program finds virtually the same solution as above.
- Now fit the model **Two\_Gaussians\_hom** using the **No correction Background model**, with the same set of starting parameters and  $c = 0.4$ . The program finds a local minimum with a very broad first peak.
- Fit the same model again, with the same set of starting parameters and  $c = 0.3$ . The program now finds the global minimum.

The message is that model fitting becomes the more delicate the more fit parameters you have. Whenever you have enough information to separate a large fit problem (distance distribution and background) into two smaller ones (first background, then distance distribution), do so. For DEER data, if you do not feel confident with fitting the background separately, it is rather unlikely that simultaneous fitting of background and distance distribution will give a stable and reliable result. You then need to improve experimental data or, if this is impossible, interpret data very cautiously.

### 4.3 Problem: Fitting a single Gaussian peak

#### Problem

Load the spectrum `deer_long_50K`. Adjust zero time to 96 ns and optimize the fit range for background correction. Set the magenta **End** cursor to an appropriate distance  $r_{\text{end}}$  and fit the distance distribution after background correction by a single Gaussian peak (model `Gaussian`). Give  $r_{\text{end}}$  and the fit parameters.

#### Answer

$r_{\text{end}} = 10$  nm,  $\langle r \rangle = 7.5$  nm,  $s(r) = 0.32$  nm.

## 5 Worked example: Very broad distance distribution

### 5.1 Assuming ideal pulses

Data set `deer_broad_50K_4us` was measured on a sample of surface-labelled gold nanoparticles. Assuming spherical nanoparticles with a well-defined diameter  $d$  and a constant distance of the nitroxide labels from the surface, the distribution of distances  $P(r)$  between labels on the surface of the same particle would be shaped as a sawtooth (triangular), increasing linearly as  $P(r) = kr$  for  $0 \leq r \leq d$ , and being identical zero for  $r > d$  (see Figure 9d).

- Load data set `deer_broad_50K_4us`. Set the magenta **Start** cursor in the **Distance distribution** to 1.0 nm. Optimize the fit range for the background (Figure 9a).
- Compute the L curve for Tikhonov regularization. The program may suggest a regularization parameter of 10, but the true corner of the L curve is at  $\alpha = 10000$ . Select this regularization parameter. (Figure 9b).
- Switch to the distance distribution (Figure 9c). It is approximately zero at  $r = 1.0$  nm. Why?

### 5.2 Correcting for limited excitation bandwidth

The microwave pulses in the DEER experiment have limited excitation bandwidth. If the dipole-dipole coupling between the two spins is of the same order as or even larger than the excitation bandwidth, the dipolar modulation is suppressed. This can be accounted for by assuming an excitation profile with Gaussian shape and with a width that depends on the width of the pulses in the DEER sequence. For observer pulses of 32 ns length and a pump pulse of 12 ns length, the excitation bandwidth is 16 MHz, the default value in `DeerAnalysis2006`.

- Select the `Excit. bandwidth` checkbox in the **Dipolar evolution** panel and repeat computation of the L curve for Tikhonov regularization. *Warning:* This may take *very* long. The shape of the L curve improves (compare Figure 9b and e) and the program automatically detects the optimum regularization parameter  $\alpha = 10000$ .
- Switch to the distance distribution (Figure 9f). The remaining deviations from a triangular shape can be traced back to the size dispersion of the nanoparticles and to the conformational distribution of the spin labels on the surface.

Even with excitation bandwidth correction, DEER has a limitation towards short distances. Therefore, `DeerAnalysis2006` does not allow for extending the distance range for analysis below 1.0 nm. Contributions in the range between 1.0 and 1.5 nm should be interpreted with utmost caution.

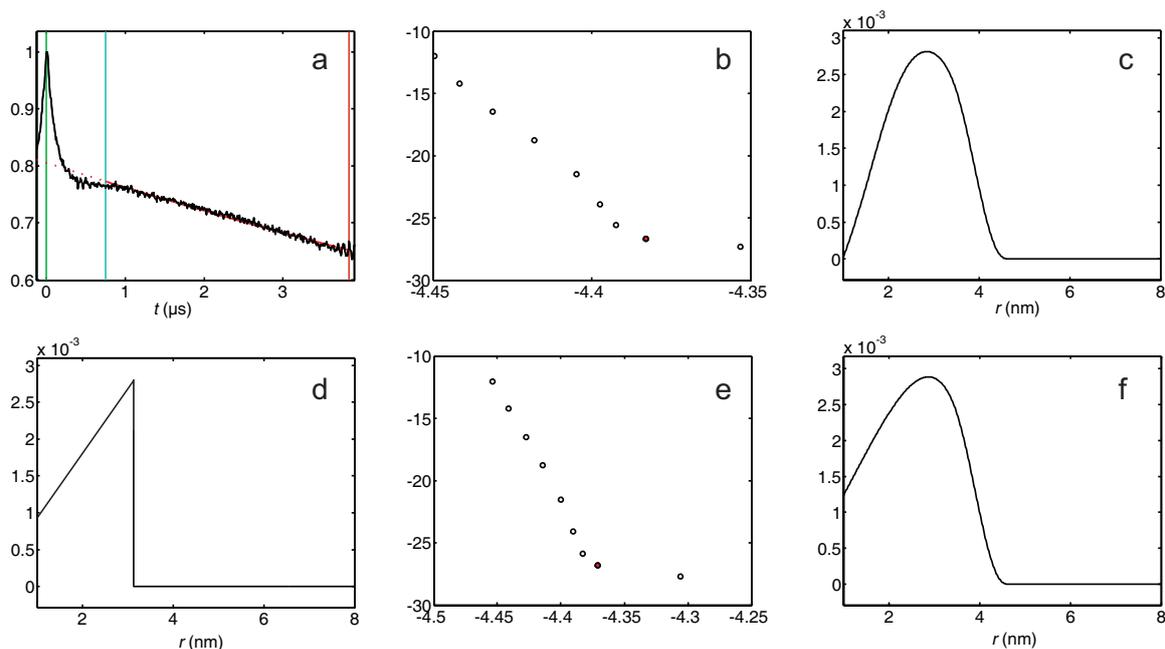


Figure 9: Analysis of data set `deer_broad_50K_4us` corresponding to a distribution of spin labels on the surface of a gold nanoparticle. (a) Original data with background fit. (b) L curve of Tikhonov regularization assuming ideal pulses. The red dot marks the optimum regularization parameter  $\alpha = 10000$ . (c) Distance distribution at  $\alpha = 10000$  assuming ideal pulses. (d) Schematic distance distribution, assuming monodisperse, ideally spherical gold nanoparticles and a uniform distance of the labels from the surface. (e) L curve of Tikhonov regularization with correction for limited excitation bandwidth of the pulses. The red dot marks the optimum regularization parameter  $\alpha = 10000$ . (f) Distance distribution at  $\alpha = 10000$  with excitation bandwidth correction.

## 6 Worked example: Dual display

In many applications it is of interest whether or not the distance distribution between two spin labels changes. Since conversion of dipolar evolution functions to distance distributions is an ill-posed problem, such comparisons should always be made on the original time-domain data. Changes in the distance distribution should only be discussed after it is established that original data are significantly different. Otherwise, they might be noise-related. For comparing original data sets, data sets after background correction, and distance distributions of two samples, `DeerAnalysis2006` features a dual display.

### 6.1 A case with no significant difference

- Load data set `LHCIIb_1`, which is a measurement on mutant `S106R1/S160R1` of light harvesting complex IIb with a particular pigment mixture. Set zero time to 128 ns, optimize the range for background fitting, and perform L curve computation for Tikhonov regularization. The corner of the L curve is difficult to determine. By looking at the fits of the **Dipolar evolution** data for different values of the regularization parameter  $\alpha$ , one finds that  $\alpha = 1000$  is a reasonable choice (Figure 10a).
- Load data set `LHCIIb_2`, which is a measurement on the same mutant but with a different

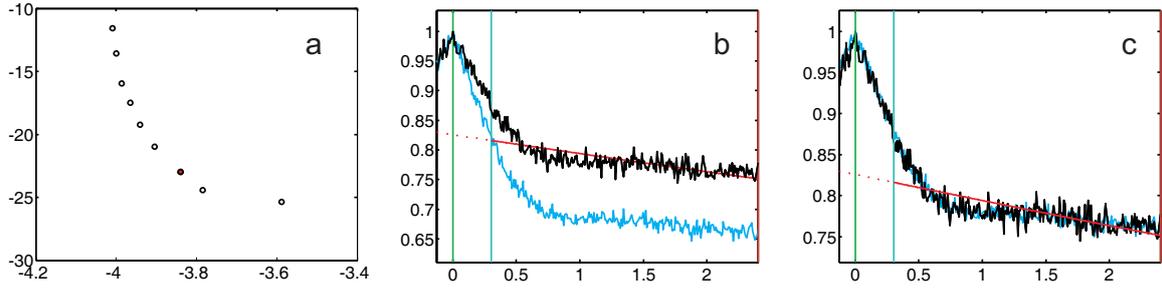


Figure 10: Dual display processing of data sets LHCIIb.1 and LHCIIb.2. (a) L curve for data set LHCIIb.1. The red dot shows the estimated optimum regularization parameter. (b) Original data sets LHCIIb.2 (black) and LHCIIb.1 (blue). The background fit corresponds to the active data set (black). (c) Original data sets after modulation depth scaling.

pigment mixture. In the **Data sets** panel, the newly loaded data is now set **A**: (black) and the previously loaded data is set **B**: (blue). Again set zero time to 128 ns, optimize the range for background fitting, and perform L curve computation for Tikhonov regularization. Select  $\alpha = 1000$  (this is a case where the L curve is misleading), and uncheck the L curve checkbox in the **Distance distribution** panel.

- Now select the **Dual display** checkbox in the **Original data** panel. Data set **B**: is displayed as a blue trace in all three plots (Figure 10b). The black and blue traces do not agree. This could be due to either of the following reasons: (a) different degree of spin labeling and thus different modulation depth or (b) a change in the distance distribution.
- To suppress effects of different modulation depth, select the **mod. depth scaling** checkbox in the **Original data** panel. There is no longer a significant difference between the original data (Figure 10c). Any difference that can be seen between the distance distributions is due to noise.

## 6.2 A case with significant difference

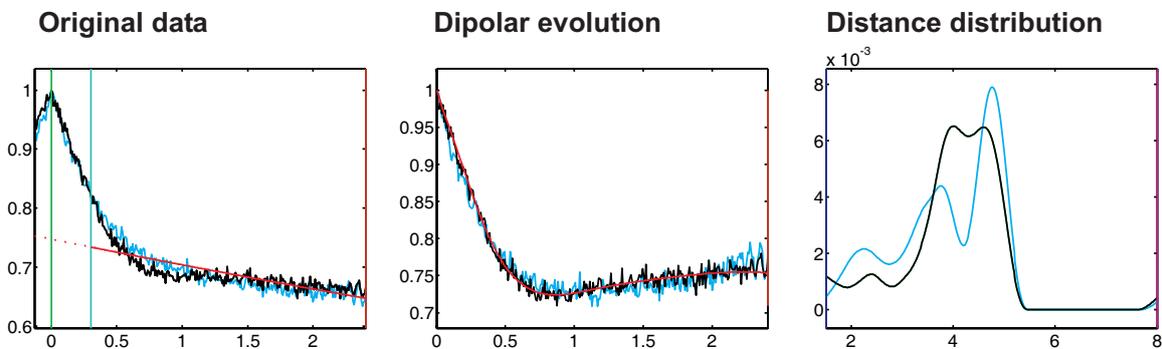


Figure 11: Dual display processing of data sets LHCIIb.1 and LHCIIb.3 with modulation depth scaling.

- Now load data set LHCIIb\_3, corresponding to yet another pigment mixture. DeerAnalysis2006 reverts to the normal display mode. In the **Data sets** panel, the data set A: (black) is now LHCIIb\_3 and data set B: (blue) becomes LHCIIb\_2. Data set LHCIIb\_1 is discarded.
- Set zero time to 128 ns, optimize the range for background fitting, and perform L curve computation for Tikhonov regularization. Here, a regularization parameter of  $\alpha = 100$  or even  $\alpha = 10$  might be appropriate, but for the sake of comparison also select  $\alpha = 1000$ . Uncheck the L curve checkbox in the **Distance distribution** panel.
- Select the **Dual display** and **mod. depth scaling** checkboxes in the **Original data** panel. The difference is subtle, but it is there.
- Load data set LHCIIb\_1 and repeat the procedure. The difference can be seen much better with this data set, as it has better signal-to-noise ratio than data set LHCIIb\_2 (Figure 11).

## 7 Spin counting via the modulation depth

The pump pulse in a DEER experiment excites only a fraction of the spins B coupled to the observer spin A. Therefore, only a fraction of the echo is modulated, the modulation depth  $\lambda$  is smaller than unity. If several B spins are coupled to the same A spin there is a higher probability that the pump pulse excites at least one of these spins and the modulation depth is thus larger. If all the other parameters that influence modulation depth (pulse lengths, flip angle of the pump pulse, spectral line shape of the nitroxide) are kept constant, the number of coupled spins in a nanoobject can thus be counted [10]. Unless the nanoobjects are highly diluted, a proper background correction is needed before the modulation depth is determined [7]. DeerAnalysis2006 incorporates the mathematics that are needed to convert background-corrected modulation depths to the number of coupled spins.

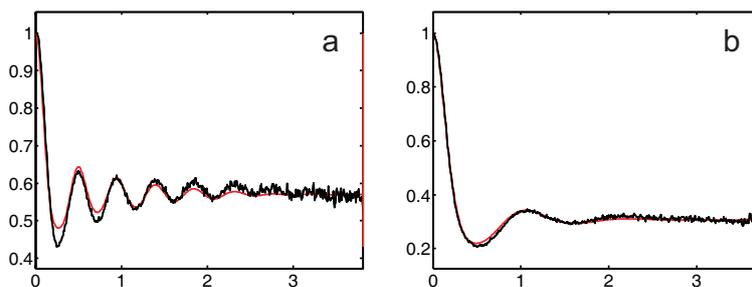


Figure 12: Spin counting via modulation depths. The modulation depth for data set **biradical** (a) is smaller than for data set **triradical** (b).

### 7.1 Worked example: Calibration of spin counting

- Load data set **biradical** and optimize the range for background fitting. Optimize cutoff time (repeated clicking of the brown ! button in the **Original data** panel) and perform a Tikhonov regularization with L curve computation. Select a regularization parameter of 0.1. The normalized background-corrected signal decays to approximately 0.57 (Figure 12a), corresponding to a modulation depth  $\lambda \approx 0.43$ .

- The `Number of spins` edit field now displays the number 2.04 in red color. This is very close to the expected number of 2 for a biradical, as this measurement was done under the standard conditions of the lab in which the program DeerAnalysis was developed. For a measurement at a different spectrometer, the number would deviate more strongly.
- Write 2.0 into the `Number of spins` edit field and press return. The updated number is displayed in green. Spin counting is now calibrated.
- Load data set `triradical`. Adjust zero time to 128 ns and optimize the background range. Optimize cutoff. With the modulation depth determined by APT, the spin count is 3.10.
- Perform a Tikhonov regularization with L curve computation. Select a regularization parameter of 100. Here, the normalized background-corrected signal decays to approximately 0.3 (Figure 12b), corresponding to a modulation depth  $\lambda \approx 0.7$ . The spin count changes only slightly to 3.11. Generally, the error in spin counting due to variations in the experimental conditions, e.g., different width of the resonator mode due to different overcoupling is larger than the difference between the APT result and optimized Tikhonov regularization.

## 7.2 Problem: Determining the number of spins in unknown samples

### Problem

Using the calibration made above and APT, determine the number of spins for data sets (a) `sampleX` and (b) `deer_broad_50K_4us`. Interpret the result for the latter data set, which stems from a sample of spin-labelled gold nanoparticles.

### Answer

(a) 2.04. The sample is a biradical. (b) 1.39. There are two aspects with respect to interpretation. First, if correct, the number would correspond to the *average* number of spin labels per nanoparticle. Second, part of the contribution at short distances will be lost. The computed number is thus expected to be lower than the actual average number of spin labels per nanoparticle. Indeed, when using Tikhonov regularization with excitation bandwidth correction and  $\alpha = 10000$ , one finds a number of 1.45. Generally, different values for the number of spins with and without excitation bandwidth correction point to the fact that contributions from short distances are underestimated. This correction cannot, however, recover the value of the modulation depth corresponding to ideal pulses.

## 8 Using single label data

There are two reasons why one may want to perform a DEER experiment on a singly labelled sample, which has only a background decay but no defined dipolar evolution. First, the background decay contains information on the concentration of spin labels. Second, for a protein the actual background function may differ from the function expected for a homogeneous distribution of point-like particles in space. By measuring data for the single mutants, one can obtain an experimental background function that can later be used for correcting the data for double mutants.

### 8.1 Worked example: Determining concentrations

Similar to modulation depths, the rate of background decay depends on pulse lengths, correct flip angle of the pump pulse, and shape of the nitroxide spectrum. Concentration measurements thus need to be calibrated.

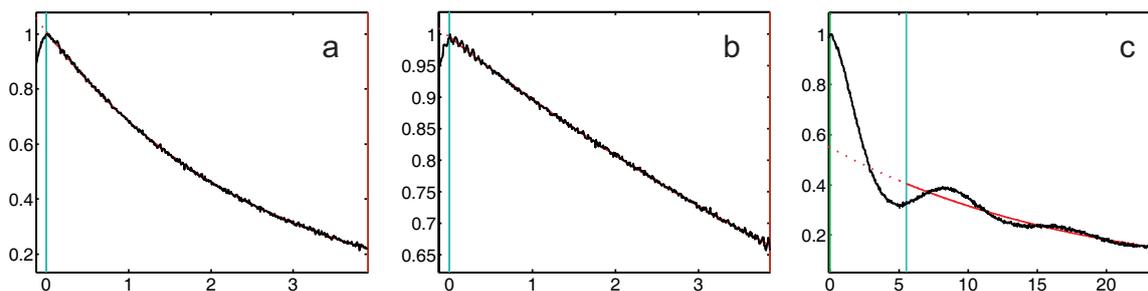


Figure 13: Determination of concentrations by exponential background fitting. (a) Calibration data set `tempol_toluene_2500uM` (b) Data set `monoradical` with a concentration of about  $670 \mu\text{mol L}^{-1}$ . (c) Data set `deer_long_50K` with a concentration of about  $350 \mu\text{mol L}^{-1}$

- Load data set `tempol_toluene_2500uM`. This sample was prepared by dissolving  $2 \text{ mmol L}^{-1}$  TEMPOL in toluene and shock-freezing the solution. As the toluene glass at  $80 \text{ K}$  has only about 80% of the volume of the solution at room temperature, the actual concentration is  $2.5 \text{ mmol L}^{-1}$ .
- Correct the zero time to 128 ns and then set the start for background fitting to 0 ns by direct input into the blue edit field in the **Original data** panel (Figure 13a). The **Density** edit field in the **Background model** panel now shows a value of 2.515. Correct this value to 2.5 by direct input into the edit field and press Return. The updated number is now shown in green colour. Concentration measurements are now calibrated for the experimental parameters used in the measurement of sample `tempol_toluene_2500uM`. Correct the zero time to 128 ns and then set the start for background fitting to 0 ns by direct input into the blue edit field in the **Original data** panel (Figure 13b). The **Density** edit field in the **Background model** panel now shows a value of 0.671. The concentration in this sample is thus about  $670 \mu\text{mol L}^{-1}$ .
- Load data set `deer_long_50K`. Correct the zero time to 96 ns and then optimize the range for background fitting (brown ! button in the **Original data** panel, Figure 13c). For this sample, the concentration is about  $350 \mu\text{mol L}^{-1}$ .

## 8.2 Worked example: Deriving and using an experimental background function

- Load data set `monolabel_41`. This data set was obtained on a membrane protein singly labelled at residue 41. Set zero time to 128 ns and start of the background fit to 0 ns. The decay cannot be fit by a homogeneous distribution in three dimensions (Figure 14a).
- Change the value in the edit field **Homogeneous dimensions** of the **Background model** panel to 2. The fit improves significantly (Figure 14a). This is because the distribution of a membrane protein in a liposome on the length scale of DEER is much closer to a homogeneous distribution in two dimensions than to a homogeneous distribution in three dimensions [11]. Yet, a close look reveals that even the two-dimensional homogeneous distribution does not perfectly fit the decay (Figure 14b).
- Select the **Polynomial** radiobutton in the **Background model** panel. The logarithm of the original data is now fitted by a 5<sup>th</sup> order polynomial. This provides an almost perfect

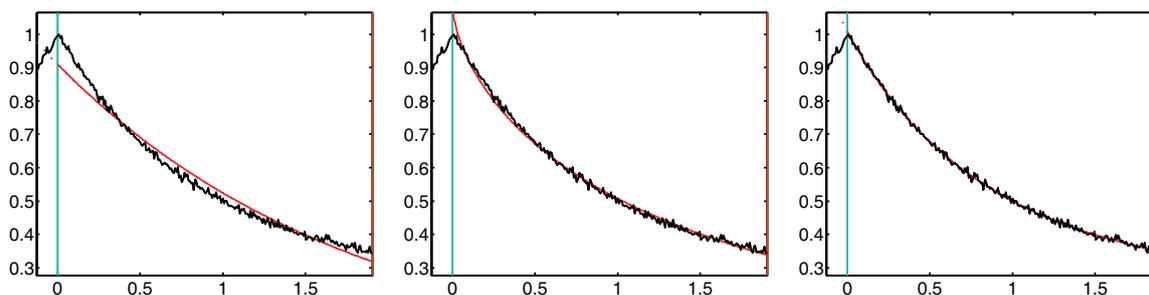


Figure 14: Different background models fitted to data set `monolabel_41` obtained on a membrane protein singly-labelled at position 41. (a) Fit by an exponential decay corresponding to a homogeneous distribution of point-like labels in space. (b) Fit by a stretched exponential decay corresponding to a homogeneous distribution of point-like labels in a plane. (c) Fit of the logarithm of the data by a 5<sup>th</sup> order polynomial.

fit (Figure 14c). To see the fit residual in the **Dipolar evolution** plot better, select **Tikhonov reg.** in the **Distance analysis** panel. Except for noise, the residual is flat.

- Decrease **Polynomial order**, watching the **r.m.s.** value in the **Background model** panel. A polynomial order as low as 3 provides a fit with the same quality, for a polynomial order of 2 the r.m.s. value increases from 0.000479 to 0.000490 and the initial part of the fit residual (first 0.4  $\mu$ s) is no longer flat. Hence, set polynomial order to 3.
- Click on the **Save** button in to the **Polynomial** subpanel and save the background polynomial as `bckg_41`.
- Load data set `monolabel_423` and repeat the procedure. In this case, the optimum order of the polynomial is 5. Save this background polynomial as `bckg_423`.
- Load data set `mix_monolabel_41_423`. This data set was obtained on a mixture of the two singly labelled mutants. Set start of the background range to 0 ns and select the **Experimental** radiobutton in the **Background model** panel. Using the **Load** button in the **Experimental** subpanel, load the background polynomial `bckg_41`. The r.m.s. value for the fit with this background polynomial is 0.000410. At the beginning and at the end of the data, there are some deviations from a flat residual.
- Load the background polynomial `bckg_423`. The r.m.s. value for the fit with this background polynomial is 0.000432. The residual exhibits a slight oscillation.
- Using the **Add** button in the **Experimental** subpanel, add the background polynomial `bckg_41`. In the dialog window that pops up, input a weighting of 1.0 for this second background polynomial. The r.m.s. value for the fit with the sum of the two single-label background polynomials improves to 0.000402. The sum of the two single-label polynomials would be almost perfect for fitting the background in a sample of the protein doubly labelled at positions 41 and 423.

Here the tutorial ends. Further features of DeerAnalysis2006 and some of the background theory are explained in the user manual.

## Acknowledgments

All shape-persistent biradicals for the example data sets were prepared by A. Godt, University of Bielefeld [9]. The surface-labelled gold nanoparticles were prepared by P. Ionita and V. Chechik, University of York. The samples of plant light harvesting complex IIb were provided by A. Bender and H. Paulsen, University of Mainz. The samples and data sets for the singly-labelled membrane proteins were provided by D. Hilger and H. Jung, LMU Munich.

## References

- [1] A.D. Milov, K.M. Salikhov, M.D. Shirov, *Fiz. Tverd. Tela (Leningrad)* **23** (1981) 957–982.
- [2] A. D. Milov, A. G. Maryasov, Yu. D. Tsvetkov, *Appl. Magn. Reson.* **15** (1998) 107–143.
- [3] M. Pannier, S. Veit, A. Godt, G. Jeschke, H. W. Spiess, *J. Magn. Reson.* **142** (2000) 331–340.
- [4] G. Jeschke, *ChemPhysChem* **3** (2002) 927–932.
- [5] G. Jeschke, *Macromol. Rapid Commun.* **23** (2002) 227–246.
- [6] G. Jeschke, A. Koch, U. Jonas, A. Godt, *J. Magn. Reson.* **155** (2001) 72–82.
- [7] G. Jeschke, G. Panek, A. Godt, A. Bender, H. Paulsen, *Appl. Magn. Reson.* **26** (2004) 223–244.
- [8] Y. W. Chiang, P. P. Borbat, J. H. Freed, *J. Magn. Reson.* **172** (2005) 279–295.
- [9] A. Godt, C. Franzen, S. Veit, V. Enkelmann, M. Pannier, G. Jeschke, *J. Org. Chem.* **65** (2000) 7575–7582.
- [10] A. D. Milov, A. B. Ponomarev, Yu. D. Tsvetkov, *Chem. Phys. Lett.* **110** (1984) 67–72.
- [11] D. Hilger, H. Jung, E. Padan, C. Wegener, K.-P. Vogel, H.-J. Steinhoff, G. Jeschke, *Biophys. J.* **89** (2005) 1328–1338.