

DipFit

User Manual

Dinar Abdullin

Bonn 2020

Table of contents

| | | |
|-----|---|----|
| 1 | Introduction..... | 3 |
| 1.1 | What is it for? | 3 |
| 1.2 | Geometric model of a spin system | 4 |
| 1.3 | Technical information..... | 5 |
| 2 | Installation..... | 5 |
| 3 | Running the program | 5 |
| 4 | Configuration file..... | 5 |
| 4.1 | Operation mode | 6 |
| 4.2 | Experimental data | 7 |
| 4.3 | Parameters of the spin system | 7 |
| 4.4 | Simulation settings | 8 |
| 4.5 | Fitting settings | 10 |
| 4.6 | Validation settings | 12 |
| 4.7 | Calculation settings..... | 14 |
| 4.8 | Output settings | 15 |
| 5 | Output data..... | 15 |
| 5.1 | Simulation output | 15 |
| 5.2 | Fitting output | 16 |
| 6 | Examples..... | 18 |
| 6.1 | Simulation of the dipolar spectra of a low-spin Fe^{3+} -organic radical spin system.... | 18 |
| 6.2 | Fitting of the PDS time trace of a low-spin Fe^{3+} -trityl spin system | 21 |
| 6.3 | Simulation of the dipolar spectra of a high-spin Fe^{3+} -organic radical spin system... | 28 |
| 6.4 | Fitting of the dipolar spectrum of a high-spin Fe^{3+} -nitroxide spin system..... | 31 |
| 7 | References..... | 36 |

1 Introduction

1.1 What is it for?

Pulsed electron paramagnetic resonance dipolar spectroscopy (PDS)^[1] offers several techniques for measuring dipolar coupling between electron spin centers. The PDS signal is usually acquired in the form of a time trace, which is modulated by dipolar frequencies. These frequencies can be obtained from the time trace by applying the Fourier transform. For disordered samples, the Fourier transform yields a distribution of dipolar frequencies, known also as dipolar spectrum. Importantly, dipolar spectra encode information about the distance between electron spin centers and, in some cases, also their relative orientation. In order to extract this spatial information, specialized algorithms of data analysis are required. In the case of the $S = 1/2$ centers with the isotropic g -factor of 2.0023, the PDS data analysis can be done by means of the program DeerAnalysis.^[2] This program is also applicable to the $S = 1/2$ centers, which have a small g -anisotropy (the difference in g -values is smaller than 0.1), such as nitroxide radicals. However, if the g -anisotropy of electron spin centers is significant, DeerAnalysis cannot be applied and alternative algorithms of PDS data analysis are required.

The program DipFit was developed for the analysis of PDS signals, which correspond to spin systems consisting of one isotropic and one anisotropic $S = 1/2$ centers. The mathematical background of this analysis is described in Refs.^[4–6] DipFit supports two operation modes: the simulation mode and the fitting mode. In the simulation mode, the PDS signal is calculated using the pre-defined geometric model of a spin system (see Chapter 1.2) and the spectroscopic parameters of spin centers. In the fitting mode, the geometry of a spin system (see Chapter 1.2) is optimized until the simulated PDS signal provides the best fit to the experimental PDS signal. This optimization procedure is done by means of a genetic algorithm.^[5]

This manual is organized as follows. In the rest of Chapter 1, the geometric model of DipFit and the technical information about the program are provided. Chapters 2 and 3 describe how one can download and run DipFit with user data, respectively. Chapter 4 provides a comprehensive description of the DipFit configuration file. The output of the program is described in Chapter 5. Finally, Chapter 6 provides several examples of using DipFit for the simulation and fitting of PDS signals.

1.2 Geometric model of a spin system

Since there is an infinite number of possible spin system geometries, it is impossible to describe all spin systems with a single geometric model. Therefore, the DipFit model of a spin system is limited by the following assumptions:

- 1) The spin system consists of two well-localized electron spin centers denoted as spin A and spin B.
- 2) Both electron spin centers have an effective spin $S = 1/2$. Spin A has an isotropic or almost isotropic g -factor, whereas spin B has an anisotropic g -factor.
- 3) The reference coordinate system of the model is set to be coincident with the g -tensor axes of spin B. A vector connecting spin A with spin B is described by three spherical coordinates: a length r , a polar angle ξ , and an azimuthal angle φ (Figure 1.1).
- 4) In order to account for the conformational flexibility of the spin system, all three geometric parameters, namely r , ξ , and φ , are allowed to have either a uniform distribution or a normal distribution. In both cases, the distributions are described by two parameters, a mean value and a width (Figure 1.2). In the case of the normal distribution, the standard deviation is used as the width parameter.
- 5) The correlation between geometric parameters r , ξ , and φ is usually unknown and, therefore, is omitted in DipFit.

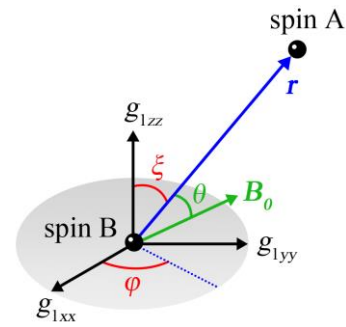


Figure 1.1. The DipFit model of a spin system.

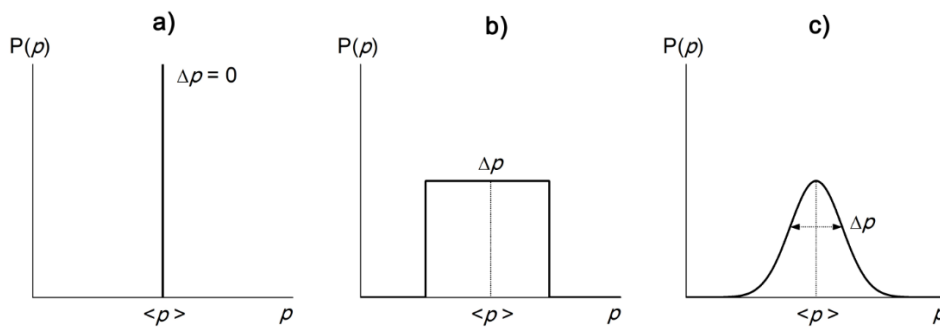


Figure 1.2. The distribution types used for the DipFit parameter $p \in (r, \xi, \varphi)$. **a)** A single value $\langle p \rangle$. The width Δp equals to 0. **b)** Uniform distribution with the mean value $\langle p \rangle$ and the width Δp . **c)** Normal distribution with the mean value $\langle p \rangle$ and the standard deviation Δp .

1.3 Technical information

DipFit is a console application written in Python3. The source code of the program can be found at <https://github.com/dinarabdullin/DipFit>. The Linux and Windows executables of the program can be found at <https://github.com/dinarabdullin/DipFit/releases>. The program is free of charge and can be distributed under GNU General Public License.

All examples in this manual were tested using a 64-core workstation from sys-Gen GmbH with 2.3 GHz processor frequency and 132 GB RAM.

2 Installation

Download the zip archive with the DipFit executables at

<https://github.com/dinarabdullin/DipFit/releases>

and unzip it into the directory where the program will be stored. That's all!

3 Running the program

- 1) Open Terminal (Linux) or Command Prompt (Windows).
- 2) Navigate into the directory with the DipFit executable:
`cd ../DipFit`
- 3) Set the permission properties (only for Linux):
`chmod 755 DipFit`
- 4) Prepare the configuration file (see Chapter 4) and run the program by the following command:
`sh DipFit.sh [path to the configuration file]` (for Linux)
`DipFit.exe [path to the configuration file]` (for Windows)

4 Configuration file

The DipFit configuration file contains all input parameters of the program. It has the usual ASCII data format and the *.cfg* extension. The configuration file can be created and edited with common text editors, such as Notepad++. Since the number of input parameters in a single configuration file is typically quite large, it is highly recommended to use the configuration files from the “examples” folder as a template to build your own configuration file. This will save time and reduce the probability of errors.

In general, the DipFit configuration file can contain up to eight information blocks:

- 1) operation mode,
- 2) experimental data and settings,
- 3) parameters of the spin system,
- 4) simulation settings,
- 5) fitting settings,
- 6) validation settings,
- 7) calculation settings,
- 8) output settings.

The information blocks 1-3 and 7-8 are mandatory. The information blocks 4-6 are optional and need to be specified only if the corresponding mode was selected. Each of these information blocks is described in detail below.

Note that comment lines in the configuration file are indicated by symbol # in the beginning of the line. Comment lines, as well as empty lines, are ignored by DipFit at execution.

4.1 Operation mode

As mentioned in Chapter 1, DipFit has two main operation modes: simulation and fitting. In the simulation mode, the PDS signal is calculated using the pre-defined geometric model of a spin system (Chapter 4.4). In the fitting mode, the geometric model of a spin system is optimized until the simulated PDS signal provides the best fit to the experimental one (Chapter 4.5). After the fitting, the parameters of the optimized model are validated via the calculation of confidence intervals. Furthermore, DipFit has an additional operation mode called validation mode. In this mode, the results of the fitting, which was done in the past, can be used to re-run or to extend the initial validation (Chapter 4.6).

The operation mode of DipFit is specified in the following line:

Taken from DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
mode = 1;
```

where

- `mode = 0` Activates the simulation mode.
- `mode = 1` Activates the fitting mode.
- `mode = 2` Activates the validation mode.

4.2 Experimental data

A data, which can be analyzed by DipFit, includes the background-free PDS time traces and the PDS spectra. The file paths of both data types can be provided in the following way:

Taken from DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
path_spectrum = "";  
path_timetrace = "examples/example02_ls_iron(III)_trityl/timetrace.dat";
```

Note that all files, which contain the experimental data, need to have the ASCII data format. If one wants to simulate/fit only the PDS time trace, `path_spectrum` has to be initialized by an empty string: `path_spectrum = ""`. If one wants to simulate/fit only the PDS spectrum, `path_timetrace` has to be initialized by an empty string: `path_timetrace = ""`. For the fitting and validation modes, at least one of the data types and the corresponding file path has to be provided. For the simulation mode, there is no need to specify any experimental data and, therefore, both file paths can be left empty. In cases, when the experimental data is still provided, the results of the simulation will be displayed together with the experimental data (Chapter 4.4).

4.3 Parameters of the spin system

A spin system, which is considered in DipFit, consists of two $S = 1/2$ centers denoted as `spinA` and `spinB` (Chapter 1.2). Currently, `spinA` can have only an isotropic or almost isotropic g -tensor, whereas the g -tensor of `spinB` can also be anisotropic. The principal g -factors of both spins can be given as follows:

Taken from DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
# trityl  
spinA:  
{  
    type = "isotropic";  
    g = [2.0032, 2.0032, 2.0032];  
};  
  
# low-spin iron(III)  
spinB:  
{  
    type = "anisotropic";  
    g = [1.56, 2.28, 2.91];  
};
```

where

- `type` The type of an electron spin center: “isotropic” or “anisotropic”.
- `g` g -factor. It should always to contain 3 components: `g = [gxx, gyy, gzz]` in the case of an orthorhombic g -factor, `g = [g⊥, g⊥, g∥]` in the case of an axial g -factor, `g = [giso, giso, giso]` in the case of an isotropic g -factor.

4.4 Simulation settings

These settings are relevant only for the simulation mode but can be omitted for the fitting and validation modes.

First, the type of simulated data is provided in the list called [simulation_modes](#):

Taken from DipFit/examples/example01_ls_iron(III)_nitroxide/config_ex01_1.cfg

```
simulation_modes:
{
    spc = 0;
    timetrace = 0;
    spc_vs_theta = 0;
    spc_vs_xi = 1;
    spc_vs_phi = 0;
    spc_vs_temp = 0;
};
```

where

- | | |
|------------------------------|--|
| spc | If set to 1, the dipolar spectrum will be simulated. If no experimental spectrum was defined (see path_spectrum), the frequency ranges will be set automatically. |
| timetrace | If set to 1, the dipolar time trace will be simulated. If no experimental time trace was defined (see path_timetrace), the time ranges will be set automatically and the modulation depth will be set to 0.5. |
| spc_vs_theta | If set to 1, the dipolar spectrum will be simulated in dependence on the angle θ between the inter-spin vector \vec{r} and the direction of the applied magnetic field \vec{B}_0 . |
| spc_vs_xi | If set to 1, the dipolar spectrum will be simulated in dependence of the angle ξ between the inter-spin vector \vec{r} and the g_{zz} -axis of spin B (Figure 1.1). |
| spc_vs_phi | If set to 1, the dipolar spectrum will be simulated in dependence of the angle φ between the projection of the inter-spin vector \vec{r} on the xy -plane of spin B and the g_{xx} -axis of spin B (Figure 1.1). |
| spc_vs_temp | If set to 1, the dipolar spectrum will be simulated in dependence of temperature. This dependence is relevant only for the RIDME data acquired at liquid helium temperatures. ^[5] |

In order to switch off unwanted simulations, the corresponding setting should be set to 0.

Next, the parameters of the simulation(s) need to be specified in the list [simulation_variables](#):

Taken from DipFit/examples/example01_ls_iron(III)_nitroxide/config_ex01_1.cfg

```
simulation_variables:
{
  r_mean = 2.50;
  r_width = 0.00;
  xi_mean = 0.0;
  xi_width = 0.0;
  phi_mean = 0.0;
  phi_width = 0.0;
  temp = 300.0;
}
```

where

| | |
|---------------------------|--|
| r_mean | The mean value $\langle r \rangle$ of the inter-spin distance distribution $P(r)$. Given in [nm]. |
| r_width | The width Δr of the inter-spin distance distribution $P(r)$ (Figure 1.2). Given in [nm]. The type of $P(r)$ is specified in the calculation_settings list (Chapter 7). |
| xi_mean | The mean value $\langle \zeta \rangle$ of the angular distribution $P(\zeta)$. Given in [degree]. |
| xi_width | The width $\Delta \zeta$ of the angular distribution $P(\zeta)$ (Figure 1.2). Given in [degree]. The type of $P(\zeta)$ is specified in the calculation_settings list (Chapter 7). |
| phi_mean | The mean value $\langle \varphi \rangle$ of the angular distribution $P(\varphi)$. Given in [degree]. |
| phi_width | The width $\Delta \varphi$ of the angular distribution $P(\varphi)$ (Figure 1.2). Given in [degree]. The type of $P(\varphi)$ is specified in the calculation_settings list (Chapter 7). |
| temp | The temperature of the PDS experiment in [K]. Do not set this parameter to 0! Default value is 300 K. |

Finally, several simulation settings have to be provided in the list [simulation_settings](#):

Taken from DipFit/examples/example01_ls_iron(III)_nitroxide/config_ex01_1.cfg

```
simulation_settings:
{
  theta_ranges = [0.0, 90.0, 91];
  xi_ranges = [0.0, 90.0, 10];
  phi_ranges = [0.0, 90.0, 91];
  temp_ranges = [1.5, 300.0, 91];
  faxis_normalized = 1;
  plot_3d = 0;
}
```

where

| | |
|------------------------------|---|
| theta_ranges | Defines the θ axis for the spc_vs_theta simulation. This axis is defined by three numbers: theta_ranges = $[\theta_{min}, \theta_{max}, N_{\theta}]$, where θ_{min} and θ_{max} are the minimal and maximal values of θ , respectively, and N_{θ} is the number of evenly spaced samples in the interval $[\theta_{min}, \theta_{max}]$. |
| xi_ranges | Defines the ζ axis for the spc_vs_xi simulation. This axis is defined by three numbers: xi_ranges = $[\zeta_{min}, \zeta_{max}, N_{\zeta}]$, where ζ_{min} and ζ_{max} are the |

minimal and maximal values of ζ , respectively, and N_ζ is the number of evenly spaced samples in the interval $[\zeta_{min}, \zeta_{max}]$.

phi_ranges Defines the φ axis for the **spc_vs_phi** simulation. This axis is defined by three numbers: **phi_ranges** = $[\varphi_{min}, \varphi_{max}, N_\varphi]$, where φ_{min} and φ_{max} are the minimal and maximal values of φ , respectively, and N_φ is the number of evenly spaced samples in the interval $[\varphi_{min}, \varphi_{max}]$.

temp_ranges Defines the temperature axis (T axis) for the **spc_vs_temp** simulation. This axis is defined by three numbers: **temp_ranges** = $[T_{min}, T_{max}, N_T]$, where T_{min} and T_{max} are the minimal and maximal values of T , respectively, and N_T is the number of evenly spaced samples in the interval $[T_{min}, T_{max}]$.

faxis_normalized If set to 1, the frequency axis of the simulated dipolar spectra will be normalized by the dipolar coupling constant of an isotropic spin pair:

$$\nu_0 = \frac{\mu_0}{4\pi} \frac{\beta_e^2 g_e^2}{r^3},$$

where μ_0 is the vacuum permeability, β_e is the Bohr magneton, g_e is the g -factor of free electron, and r is set to the value of **r_mean** in the **simulation_variables** list.

plot_3d If set to 1, the results of the **spc_vs_xi**, **spc_vs_phi**, and **spc_vs_temp** simulations will be depicted as 3-dimensional plots.

4.5 Fitting settings

These settings are relevant for the fitting and validation modes but can be omitted for the simulation mode.

A list with all possible fitting parameters of DipFit is called **fitting_variables** and has the following structure:

Taken from DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
fitting_variables:
{
  r_mean :    opt = 1;   range = [2.00, 3.00]; value = 0.0  };
  r_width :   opt = 1;   range = [0.00, 0.10]; value = 0.0  };
  xi_mean :   opt = 1;   range = [ 0.0, 90.0]; value = 0.0  };
  xi_width :  opt = 1;   range = [ 0.0, 30.0]; value = 0.0  };
  phi_mean :  opt = 1;   range = [ 0.0, 90.0]; value = 0.0  };
  phi_width : opt = 1;   range = [ 0.0, 30.0]; value = 0.0  };
  temp :      opt = 0;   range = [ 9.0, 11.0]; value = 10.0 };
};
```

Table 4.1. The fitting parameters of DipFit.

| Parameter keyword | Parameter description ^[a] | Ranges | Units |
|-------------------------------------|---|--|--------|
| r_mean | The mean value $\langle r \rangle$ of the inter-spin distance distribution $P(r)$ | Any reasonable ranges inside (0.0, 15.0] | nm |
| r_width | The width Δr of the inter-spin distance distribution $P(r)$ | Any reasonable ranges inside (0.0, 15.0] | nm |
| xi_mean | The mean value $\langle \xi \rangle$ of the angular distribution $P(\xi)$ | [0, 90] | degree |
| xi_width | The width $\Delta \xi$ of the angular distribution $P(\xi)$ | [0, 90] for unimodal distribution [0, 30] for normal distribution | degree |
| phi_mean | The mean value $\langle \varphi \rangle$ of the angular distribution $P(\varphi)$ | [0, 90] | degree |
| phi_width | The width $\Delta \varphi$ of the angular distribution $P(\varphi)$ | [0, 90] for unimodal distribution [0, 30] for normal distribution | degree |
| temp ^[b] | The temperature of the PDS experiment | Any reasonable ranges inside (0, 300] | K |

^[a] The type of the distributions $P(r)$, $P(\xi)$, and $P(\varphi)$ is specified in the [calculation_settings](#) list (Chapter 7).

^[b] This parameter is relevant only for the RIDME data acquired at liquid helium temperatures.^[5]

In this list, each text line corresponds to one of the fitting parameters (Table 4.1) and contains three settings called [opt](#), [range](#), and [value](#). The most important setting is [opt](#), since it determines whether the corresponding parameter will be included into the fitting or not:

[opt](#) If [opt](#) = 0, the parameter is excluded from the fitting.

 If [opt](#) = 1, the parameter is included into the fitting.

In the case of [opt](#) = 1, the ranges for the corresponding parameter have to be specified in the [range](#) field. The appropriate ranges of all fitting parameters are given Table 4.1. In the case of [opt](#) = 0, the corresponding parameter will be set to a constant value, which has to be provided in the [value](#) field. The units of the [range](#) and [value](#) settings for each of the fitting parameters are given in Table 4.1.

Another important aspect of the fitting is an optimization algorithm. Currently, DipFit allows using only one optimization algorithm, a genetic algorithm. This algorithm has been shown to be very efficient when one deals with a large number of optimization parameters and needs to find a global minimum.^[7–9] Importantly, the genetic algorithm has its own internal parameters, which determine its ability to find a global minimum. Optimal values of these parameters may vary depending on a particular PDS data set. Therefore, the configuration file allows users to set these parameters manually. However, it is highly recommended to use the genetic algorithm’s parameters from the configuration files given in the “examples” folder, because these parameters were obtained after the extensive tests of the genetic algorithm on numerous PDS data sets.

The parameters of the optimization algorithm are provided in the list [fitting_settings](#):

Taken from DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
fitting_settings:
{
  fitted_data = "timetrace";
  display_graphics = 0;
  method = "genetic";
  num_generations = 500;
  size_generation = 128;
  prob_crossover = 0.5;
  prob_mutation = 0.01;
};
```

where

| | |
|----------------------------------|---|
| fitted_data | The type of the fitted data: “timetrace” or “spectrum”. If fitted_data = “timetrace”, the file path of the experimental PDS time trace has to be specified in path_timetrace . If fitted_data = “spectrum”, the file path of the experimental PDS spectrum has to be specified in path_spectrum . |
| display_graphics | If set to 1, the results of fitting will be displayed after each optimization step. In order to fasten the fitting procedure, display_graphics is recommended to be set to 0. |
| method | The type of the optimization algorithm. Currently, only the genetic algorithm can be chosen: method = “genetic”. |
| num_generations | The total number of optimization steps (generations). |
| size_generation | The number of parameter sets (chromosomes) that are evaluated simultaneously by the genetic algorithm. A good practice is to set this parameter to 10 x (the number of fitting parameters). |
| prob_crossover | The crossover rate. |
| prob_mutation | The mutation rate. |

4.6 Validation settings

These settings are relevant for the fitting and validation modes but can be omitted for the simulation mode.

The fitting of PDS data by DipFit yields the optimized values of fitting parameters, whose precision is unknown. To fill this gap, DipFit allows the validation of fitting parameters. Since DipFit deals with up to seven fitting parameters simultaneously, the parameter space is very large and cannot be fully covered during the validation. Therefore, the validation is done in a less general way by recording the goodness of fit in dependence of single fitting parameters or different pairs of fitting parameters, while setting all other fitting parameters to their optimized values. As a measure for the goodness of fit, the RMSD between the experimental and simulated

PDS time traces or the experimental and simulated PDS spectra is used. The calculated dependencies of the RMSD on fitting parameters are then used to determine the 1σ confidence intervals of fitting parameters. Note that the 1σ confidence interval, as well as its multiples, 2σ and 3σ confidence intervals, are commonly used as error estimates of fitting parameters. According to statistics, the 1σ , 2σ , and 3σ confidence intervals contain the true value of a fitting parameter with the probabilities of 68 %, 95 %, and 99.7 %, respectively. In DipFit, user can choose which of those intervals will be used to determine the parameters' errors (see [threshold](#)). The validation settings are given in the list called [validation](#):

Taken from DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
validation:
{
  variables = (("r_mean", "r_width"), ("xi_mean", "xi_width"), ("phi_mean", "phi_width"));
  Ns = 10000;
  threshold = 1;
  path_optimized_parameters = "";
};
```

where

[variables](#) The list of fitting parameters used for the validation. If no validation is wanted, this list should be left empty: [variables](#) = (). Otherwise, the variables of each separate validation run should be enclosed in round brackets, e.g., [variables](#) = (("r_mean", "r_width"), ("xi_mean", "xi_width"), ("temp")). In the given example, the following three dependences will be recorded: 1) the 2-dimensional dependence of the RMSD on $\langle r \rangle$ and Δr , 2) the 2-dimensional dependence of the RMSD on $\langle \zeta \rangle$ and $\Delta \zeta$, and 3) the 1-dimensional dependence of the RMSD on temperature. Note that the names of fitting parameters are given in accordance to their keywords (column 1 in Table 4.1) During the validation, the selected fitting parameter(s) will be randomly set to [Ns](#) different values within the ranges specified in the [fitting_variables](#) list.

[Ns](#) The number of points in each validation data set.

[threshold](#) The multiple of 1σ , which sets the error level for fitting parameters.

As mentioned in Chapter 4.1, DipFit has an additional operation mode called validation mode. This mode can be useful in cases, when the user has already done the fitting of the PDS data earlier and now just wants to use the previous results for validation. Thus, the validation mode allows avoiding the unnecessary repetition of the fitting procedure, which is usually time expensive. In order to activate this mode, one has to set [mode](#) = 2 (Chapter 4.1) and provide the path to the file [parameters.dat](#), which stores the optimized values of the fitting parameters (Chapter 5.2). The latter file path has to be given in the [path_optimized_parameters](#) field of the [validation](#) list:

```
validation:
{
  variables = (("xi_mean", "phi_mean"), ("r_width", "xi_mean"), ("r_width", "xi_width"),
              ("r_width", "phi_mean"), ("r_width", "phi_width"));
  Ns = 10000;
  threshold = ;
  path_optimized_parameters = "examples/example02_ls_iron(III)_trityl/for_validation/parameters.dat"
};
```

4.7 Calculation settings

Calculation settings, which are relevant for all three operation modes, have to be provided in the list [calculation_settings](#):

Taken from DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
calculation_settings:
{
  Ns = 1000000;
  r_distr = "normal";
  xi_distr = "normal";
  phi_distr = "normal";
  fmin = 0;
  fmax = 0;
  tmin = 0;
  tmax = 0;
  magnetic_field = 1.20197;
  g_selectivity = 0;
};
```

where

| | |
|--------------------------------|--|
| Ns | The number of Monte-Carlo averages used in numerical integration. It determines how accurately the powder averaging will be done. |
| r_distr | The type of the inter-spin distance distribution $P(r)$: “uniform” or “normal”. |
| xi_distr | The type of the angular distribution $P(\xi)$: “uniform” or “normal”. |
| phi_distr | The type of the angular distribution $P(\varphi)$: “uniform” or “normal”. |
| fmin | The minimal frequency of the simulated spectrum. Given in [MHz]. |
| fmax | The maximal frequency of the simulated spectrum. Given in [MHz]. |
| tmin | The minimal time value of the simulated time trace. Given in [μ s]. |
| tmax | The maximal time value of the simulated time trace. Given in [μ s]. |
| g_selectivity | If set to 1, the difference in spin polarization for the different g -values of spin B will be taken into account. ^[5] This parameter is relevant only for the RIDME data acquired at liquid helium temperatures. |
| magnetic_field | The value of the applied magnetic field used in the PDS experiment. Given in [T]. This parameters has to be provided only if g_selectivity = 1. |

4.8 Output settings

In the last part of the configuration file, output settings are provided in the list called [output](#):

Taken from DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
output:
{
  directory = "";
  save_data = 1;
  save_figures = 1;
};
```

where

- | | |
|------------------------------|---|
| directory | The output directory of DipFit. DipFit will create a new folder inside this directory and name it in accordance to the current date and time. All simulation/fitting/validation results will then be saved into this folder. When left empty, i.e. directory = "" , the directory of the configuration file will be used as the output directory. |
| save_data | The data obtained from simulation/fitting/validation will be saved into DAT files (Chapter 5). |
| save_figures | The graphical data obtained from simulation/fitting/validation will be saved into PNG files (Chapter 5). |

5 Output data

The output data of DipFit is stored in two forms: numerical and graphical. The numerical data is saved as DAT files, whereas the graphical data as PNG files (see [output](#)). The detailed description of all output files is given below.

5.1 Simulation output

Depending on the simulation (Chapter 4.4) and output (Chapter 4.8) settings of the configuration file, the following output files will be created:

- [spc.dat](#)
Data: a simulated dipolar spectrum.
Content: column 1 – frequency values in [MHz];
column 2 – a simulated spectrum;
column 3 – an experimental spectrum (optional).
- [timetrace.dat](#)
Data: a simulated dipolar time trace.
Content: column 1 – time points in [μ s];
column 2 – a simulated time trace;

column 3 – an experimental time trace (optional).

- [spc_vs_X.dat](#), where **X** = [theta](#), [xi](#), [phi](#), or [T](#)

Data: dipolar spectra simulated for the different values of parameter **X**.

Content: column 1 – frequency in [MHz];
next columns – a spectrum which is simulated with the **X** value given in the 1st row of the column.

- [spc.png](#)

Content: the picture of a simulated dipolar spectrum.

- [timetrace.png](#)

Content: the picture of a simulated dipolar time trace.

- [spc_vs_X.png](#), where **X** = [theta](#), [xi](#), [phi](#), or [T](#)

Content: the picture of dipolar spectra simulated for the different values of parameter **X**.

5.2 Fitting output

Depending on the fitting (Chapter 4.5), validation (Chapter 4.6), and output (Chapter 4.8) settings of the configuration file, the following output files will be created:

- [score.dat](#)

Data: the goodness of fit in dependence of optimization step.

Content: column 1 – optimization steps;
column 2 – RMSD values.

- [fit.dat](#)

Data: a fit to an experimental data.

Content: If [fitted_data](#) = “[spectrum](#)”:
column 1 – frequency values in [MHz];
column 2 – an experimental spectrum;
column 3 – a simulated spectrum.
If [fitted_data](#) = “[timetrace](#)”:
column 1 – time points in [μ s];
column 2 – an experimental time trace;
column 3 – a simulated time trace.

- [parameters.dat](#)

Data: optimized values of fitting parameters.

Content: column 1 – the names and units of fitting parameters;

column 2 – the values of fitting parameters;
 column 3 – information about activated (“Y”) and deactivated (“N”) fitting parameters (identical to the [opt](#) setting of the [fitting_variables](#) list);
 column 4 – the estimated precision of fitting parameters;

- [parameter_errors-X.dat](#) or [parameter_errors-X-Y.dat](#), where **X** and **Y** are the keywords used for the fitting parameters (see column 1 in Table 4.1)

Data: the goodness of fit in dependence of fitting parameter(s) **X** (and **Y**)

Content: For one fitting parameter **X**:

column 1 – **X** values;

column 2 – RMSD values;

For two fitting parameter **X** and **Y**:

column 1 – **X** values;

column 2 – **Y** values;

column 3 – RMSD values.

- [score.png](#)

Content: the picture of the goodness of fit in dependence of optimization step.

- [fit.png](#)

Content: the picture of a fit to an experimental data.

- [parameter_errors.png](#)

Content: the picture of the goodness of fit in dependence of the fitting parameters specified in the [validation](#) list.

6 Examples

This chapter provides several examples of how one can use DipFit to simulate and to fit the PDS data.

6.1 Simulation of the dipolar spectra of a low-spin Fe^{3+} -organic radical spin system

The first example deals with the spectral simulations for the spin system consisting of the low-spin Fe^{3+} (ls Fe^{3+}) and the organic radical. The principal g -values of the ls Fe^{3+} are set to $g_{1xx} = 1.56$, $g_{1yy} = 2.28$, and $g_{1zz} = 2.91$. The organic radical is assumed to have an isotropic g -factor which is equal to $g_e = 2.0023$. The distance between both spin centers is fixed at 2.50 nm, whereas the values of the ζ and φ angles are varied in the range $[0^\circ, 90^\circ]$ with steps of 10° and 30° , respectively. Thus, the angular dependence of the dipolar spectrum will be explored.

Now, let's go step-by-step through this simulation:

- 1) *Prepare the configuration file.* The detailed information on how one creates such a configuration file is given in Chapter 4. The ready-to-use configuration files can be found in the following directory:

[/DipFit/examples/example01_ls_iron\(III\)_nitroxide](#)

This directory contains four configuration files [config_ex01_i.cfg](#), where $i = 1, 2, 3$, and 4. Each of these configuration files sets up the spectral simulation in dependence of ζ (the [spc_vs_xi](#) simulation mode), while setting φ to one of four different values: $\varphi = 0^\circ$ ($i = 1$), $\varphi = 30^\circ$ ($i = 2$), $\varphi = 60^\circ$ ($i = 3$), and $\varphi = 90^\circ$ ($i = 4$). Note that the [faxis_normalized](#) parameter is set to 1 in all four configuration files, meaning that the frequency axis of the simulated spectra will be normalized by the dipolar coupling constant ν_0 of an isotropic spin pair with the same inter-spin distance as mentioned above.

- 2) *Run the program.* First, open Terminal (Linux) or Command Prompt (Windows) and navigate to the directory:

[cd ../DipFit](#)

If needed, set the permission properties (only for Linux):

[chmod 755 DipFit](#)

Finally, run the program by the following command:

[sh DipFit.sh examples/example01_ls_iron\(III\)_nitroxide/config_ex01_i.cfg](#) (for Linux)

[DipFit.exe examples/example01_ls_iron\(III\)_nitroxide/config_ex01_i.cfg](#) (for Windows)

- 3) *Monitor the progress status.* During the program operation, the following messages should appear in Terminal or Command Prompt:

Reading out the configuration file... [DONE]

Starting simulation...

Running pre-calculations... [DONE]

Calculating dipolar spectrum vs xi... 100% [DONE]

Simulation is finished

- 4) *Explore the simulation results.* During the simulation, the new folder will be created inside the directory `/DipFit/examples/example01_ls_iron(III)_nitroxide` and the results of simulations will be saved in that folder. In the given example, the program will create two output files, `spc_vs_xi.dat` and `spc_vs_xi.png`. The content of `spc_vs_xi.png` will pop up on the screen after the simulation.

Dipolar spectra, obtained from all four simulations, are shown in Figure 6.1. This figure reveals a prominent deviation of the simulated spectra from the Pake doublet.^[10] Whereas the Pake doublet has two characteristic singularities, often referred to as perpendicular ($\theta = 90^\circ$) and parallel ($\theta = 0^\circ$) components, all calculated spectra here display three singularities instead. In analogy to the Pake doublet, these singularities can be subdivided into two perpendicular components, which correspond to $\theta = 90^\circ$, and one parallel component, which corresponds to $\theta = 0^\circ$. Moreover, the frequencies, at which the singularities appear in Figure 6.1, do not have a fixed ratio, as in the case of the Pake doublet, but depend on the principal g -values of the ls Fe^{3+} center and the angles ζ and φ . This dependence can be readily explained on example of four spectra corresponding to the angular combinations $(\zeta, \varphi) = (0^\circ, 0^\circ), (0^\circ, 90^\circ), (90^\circ, 0^\circ)$ and $(90^\circ, 90^\circ)$. For the angular combination $(0^\circ, 0^\circ)$, the inter-spin vector \vec{r} is collinear to the g_{zz} -axis of the ls Fe^{3+} g -tensor. Consequently, the parallel component of the spectrum is scaled by g_{zz} , yielding a singularity at $2(g_{zz}/g_e)v_0 \approx 2.91 v_0$. Then, the other two components of the ls Fe^{3+} g -tensor give rise to two perpendicular components, which appear at $(g_{xx}/g_e)v_0 \approx 0.78 v_0$ and $(g_{yy}/g_e)v_0 \approx 1.14 v_0$. The same assignment of singularities also holds for the spin pair geometry with $(\zeta, \varphi) = (0^\circ, 90^\circ)$, because the shape of the dipolar spectrum does not depend on the φ angle as soon as the ζ angle equals 0° . For the angular combination $(90^\circ, 0^\circ)$, \vec{r} is aligned along the g_{xx} -axis of the ls Fe^{3+} g -tensor. Thus, the parallel component of the spectrum is scaled by g_{xx} and appears at components $2(g_{xx}/g_e)v_0 \approx 1.56 v_0$, whereas two perpendicular components are scaled by g_{yy} and g_{zz} and appear at $(g_{yy}/g_e)v_0 \approx 1.14 v_0$ and $(g_{zz}/g_e)v_0 \approx 1.45 v_0$, respectively. Finally, the angular combination $(90^\circ, 90^\circ)$ corresponds to the case where \vec{r} is collinear to the g_{yy} -axis of the ls Fe^{3+} g -tensor. In this case, the parallel component of the spectrum is determined by the value of g_{yy} , which yields the singularity at $2(g_{yy}/g_e)v_0 \approx 2.28 v_0$, whereas the

perpendicular components of the spectrum are scaled by g_{xx} and g_{zz} and appear at $(g_{xx}/g_e)v_0 \approx 0.78 v_0$ and $(g_{zz}/g_e)v_0 \approx 1.45 v_0$.

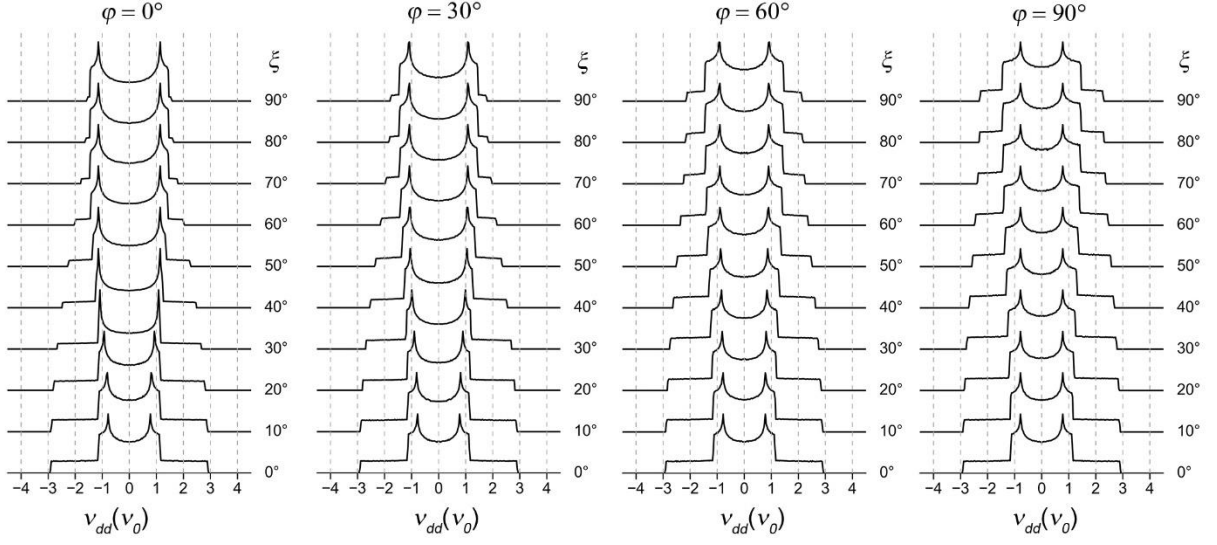


Figure 6.1. The angular dependence of the dipolar spectrum, which corresponds to the spin system consisting of an isotropic organic radical with $g = g_e$ and an anisotropic ls Fe^{3+} with $g = [1.56, 2.28, 2.91]$. The angles ζ and φ are defined in Figure 1.1.

6.2 Fitting of the PDS time trace of a low-spin Fe³⁺-trityl spin system

In the second example, the fitting and validation modes of DipFit will be explored. The PDS data, which is used for this example, was acquired on the model compound **1**^[6] (Figure 6.2). This compound contains two spatially separated electron spin centers, the trityl radical and the low-spin Fe³⁺ (1s Fe³⁺). The former spin center has an almost isotropic *g*-factor of 2.0032, whereas the latter one shows a pronounced *g*-anisotropy with three principal components $g_{xx} = 1.56$, $g_{yy} = 2.28$, and $g_{zz} = 2.91$. The PDS measurements on **1** were done using the relaxation induced dipolar modulation enhancement (RIDME)^[11,12] technique and yielded the background-free time trace shown in Figure 6.3b. Here, this time trace will be fitted using the parameters of $P(r)$, $P(\zeta)$, and $P(\varphi)$ as fitting parameters. After the fitting, the errors of the optimized parameters will be determined using the validation procedure.

Now, let's go step-by-step through the fitting:

- 1) *Prepare the configuration file.* The detailed information on how one creates such a configuration file is given in Chapter 4. The ready-to-use configuration file [config_ex02_1.cfg](#) can be found in the following directory:

[/DipFit/examples/example02_1s_iron\(III\)_trityl](#)

As follows from this configuration file, the RIDME time trace of **1** is going to be fitted (see [fitted_data](#)). The mean values and the widths of the distributions $P(r)$, $P(\zeta)$, and $P(\varphi)$ are set to be the fitting parameters (see [fitting_variables](#)). All three distributions are approximated by normal distributions (see [calculation_settings](#)) and, therefore, their widths are given by standard deviations. The genetic algorithm is set to be the fitting method. Its internal parameters are provided in [fitting_settings](#). Among them, the total number of optimization steps is set to 500. For [validation](#), the recording of the following data sets is activated: 1) the 2-dimensional dependence of the RMSD on $\langle r \rangle$ and Δr , 2) the 2-dimensional dependence of the RMSD on $\langle \zeta \rangle$ and $\Delta \zeta$, and 3) the 2-dimensional dependence of the RMSD on $\langle \varphi \rangle$ and $\Delta \varphi$. Each of these data sets will contain 10^4 data points (see [Ns](#)). The errors of the fitting parameters will be determined at the 1σ level (see [threshold](#)).

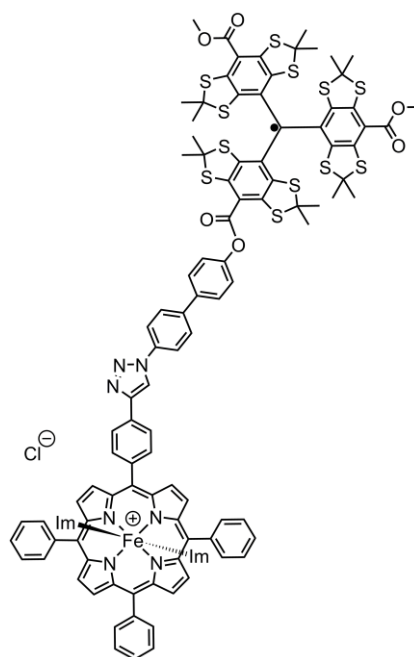


Figure 6.2. The chemical structure of the model compound **1**.

- 2) *Run the program.* First, open Terminal (Linux) or Command Prompt (Windows) and navigate to the directory:

```
cd ../DipFit
```

If needed, set the permission properties (only for Linux):

```
chmod 755 DipFit
```

Finally, run the program by the following command:

```
sh DipFit.sh examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg (for Linux)
```

```
DipFit.exe examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg (for Windows)
```

- 3) *Monitor the progress status.* During the program operation, the following messages should appear in Terminal or Command Prompt:

```
Reading out the configuration file... [DONE]
```

```
Starting fitting...
```

```
Optimization step 500 / 500: RMSD = 0.003090
```

```
Fitting is finished. Total duration: 16:09:17.278718
```

```
Saving fitting results into the directory:
```

```
../DipFit/examples/example02_ls_iron(III)_trityl/2020-01-01_00-00/ [DONE]
```

```
Plotting fitting results... [DONE]
```

```
Validating the optimized fitting parameters...
```

```
Numerical error = 0.000265
```

```
Validation is finished. Total duration: 8:39:15.294717
```

```
Saving validation results into the directory:
```

```
../DipFit/examples/example02_ls_iron(III)_trityl/2020-01-01_00-00/ [DONE]
```

```
Plotting validation results... [DONE]
```

```
Optimized fitting parameters:
```

| Parameter | Value | Optimized | Precision (+/-) |
|-----------------|--------|-----------|-----------------|
| r mean (nm) | 2.639 | Y | 0.195 |
| r width (nm) | 0.056 | Y | 0.065 |
| xi mean (deg) | 67.837 | Y | 24.898 |
| xi width (deg) | 0.020 | Y | nan |
| phi mean (deg) | 54.094 | Y | 27.080 |
| phi mean (deg) | 29.483 | Y | 21.426 |
| temperature (K) | 10.000 | N | nan |

Here, a few comments have to be made. First, the calculation time strongly depends on the used hardware (Chapter 1.3). Second, the optimized fitting parameters might slightly differ

between single fitting trials. Importantly, this deviation is well below the error of the optimized parameters (see below).

- 4) *Explore the fitting results.* During the fitting, the new folder will be created inside the directory `/DipFit/examples/example02_ls_iron(III)_trityl` and the results of fitting will be saved in that folder. The output files include `score.dat`, `fit.dat`, `parameters.dat`, `parameter_errors-r_mean-r_width.dat`, `parameter_errors-xi_mean-xi_width.dat`, `parameter_errors-phi_mean-phi_width.dat`, and the corresponding graphical files (see Chapter 5.2).

Let's briefly discuss the content of the output files. First, it is important to make sure that the genetic algorithm has converged to the global minimum. This information is contained in the files `score.dat` and `score.png`. In the present fitting, the RMSD between the RIDME time trace and its fit fell gradually down during the first 300 optimization steps and, after this, did not change anyhow significantly during the last 200 optimization steps (Figure 6.3a). This shows that the global minimum was reached and that the optimal values of the fitting parameters were found.

Although DipFit keeps track of the goodness of fit by calculating the RMSD value, an additional check of the obtained fit by eye is advisable. The fit to the RIDME time trace is stored in the files `fit.dat` and `fit.png`. As can be seen from Figure 6.3b, the shape of the RIDME time trace of **1** is well reproduced by the obtained fit. Thus, one can conclude that the optimized parameters of $P(r)$, $P(\xi)$, and $P(\varphi)$ provide a good agreement with the experimental RIDME data.

The values of optimized fitting parameters are stored in the file `parameters.dat`. The content of this file is listed in Figure 6.3c. In order to estimate how defined the optimized parameters are, three validation data sets were recorded: 1) the 2-dimensional dependence of the RMSD on $\langle r \rangle$ and Δr (`parameter_errors-r_mean-r_width.dat`), 2) the 2-dimensional dependence of the RMSD on $\langle \xi \rangle$ and $\Delta \xi$ (`parameter_errors-xi_mean-xi_width.dat`), and 3) the 2-dimensional dependence of the RMSD on $\langle \varphi \rangle$ and $\Delta \varphi$ (`parameter_errors-phi_mean-phi_width.dat`). The picture of these validation data sets is stored in the file `parameter_errors.png` and shown here as Figure 6.3d. As can be seen, the RMSD plot for the parameters of $P(r)$ has a single well-defined minimum, whose coordinates are in excellent agreement the optimized values of $\langle r \rangle$ and Δr . In contrast, the RMSD plots for the parameters of $P(\xi)$ and $P(\varphi)$ have less defined minima. For $P(\xi)$, the RMSD minimum appears on the $\langle \xi \rangle$ interval (70° , 90°) and the $\Delta \xi$ interval (0° , 30°). For $P(\varphi)$, the RMSD plot shows the minimum on the $\langle \varphi \rangle$ interval (45° , 65°) and the $\Delta \varphi$ interval (15° , 30°). Note that the optimized values of all four angular parameters are within the intervals given above.

In order to estimate the errors of the obtained parameters, the 1σ confidence intervals were calculated by DipFit. In these calculations, the dependence of the RMSD on each single fitting parameter was extracted from Figure 6.3d and, then, fitted by a Gaussian (Figure 6.4). This yielded the parameters' errors given in Figure 6.3c. The errors of the mean inter-spin distance $\langle r \rangle$ and its standard deviation Δr are ± 0.19 nm and ± 0.06 nm, respectively. The errors of both mean angles, $\langle \zeta \rangle$ and $\langle \varphi \rangle$, are similar and equal to $\pm 25^\circ$ and $\pm 27^\circ$, respectively. The errors of the standard deviations $\Delta \zeta$ and $\Delta \varphi$ differ from the errors of the mean angles and from each other. In case of $\Delta \zeta$, the error is as large as the variation range of this parameter, meaning that this parameter is undefined. In case of $\Delta \varphi$, the error is smaller than the corresponding variation range and equals to $\pm 21^\circ$. Taken all that, one can conclude that the DipFit analysis of the RIDME time trace of **1** allowed determining the distance distribution $P(r)$ with a good precision. In addition, the mean values of the angles ζ and φ could be determined with a moderate precision. However, the distribution widths of both angles were either undefined or poorly defined.

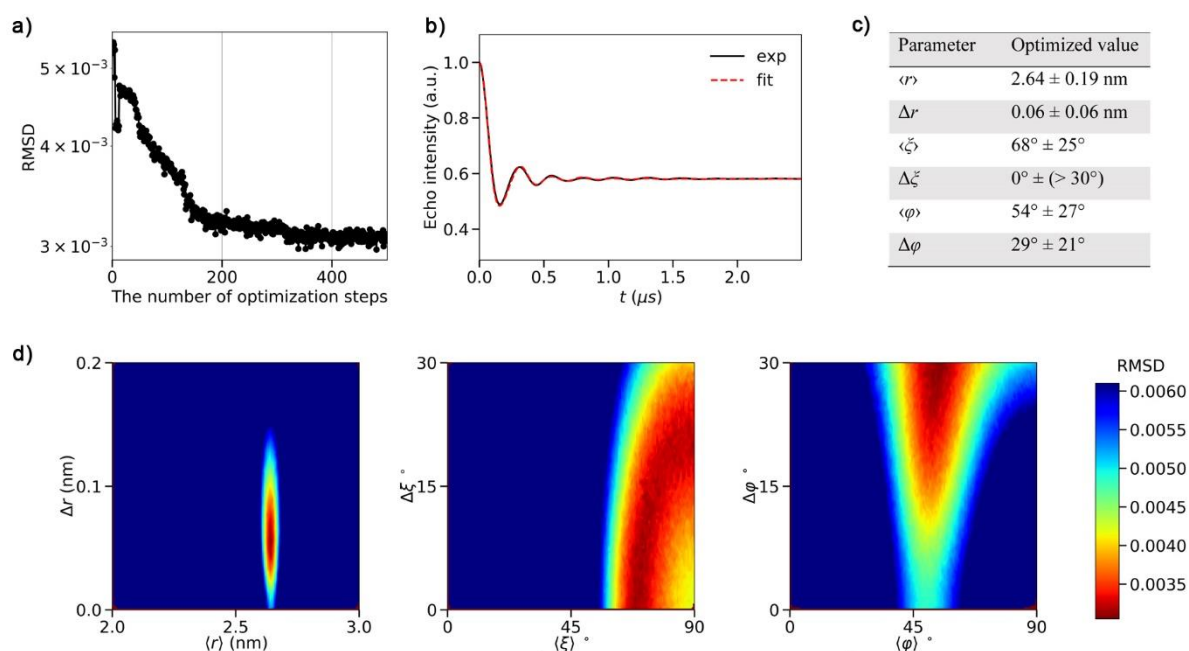


Figure 6.3. Fitting the RIDME time trace of **1** by DipFit. **a)** The RMSD in dependence of optimization step. **b)** The RIDME time trace (black line) is overlaid with the corresponding fit (red line). **c)** The optimized values of the fitting parameters. **d)** The RMSD in dependence of different pairs of fitting parameters.

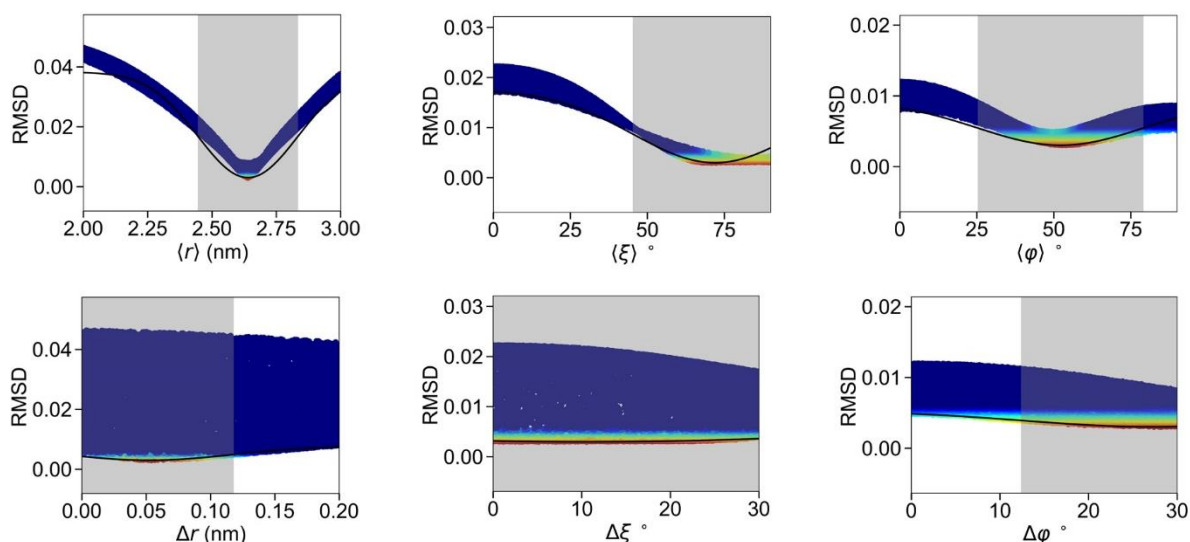


Figure 6.4. The RMSD in dependence of single fitting parameters (colored dots). The color code is identical to Figure 6.3. The Gaussian fits to this data is depicted by black lines. The parameters' ranges, which correspond to the 1σ confidence intervals, are depicted by gray bars.

Now, let's use the results of the fitting described above and run DipFit in the validation mode. To do this, one has to go through the following steps:

- 1') *Prepare the new configuration file.* The ready-to-use configuration file [config_ex02_2.cfg](#) can be found in the following directory:

[/DipFit/examples/example02_ls_iron\(III\)_trityl](#)

The configuration file [config_ex02_2.cfg](#) is based on the configuration file [config_ex02_1.cfg](#) used for the fitting. However, several changes were made to the latter file. First, the [operation mode](#) was changed from “fitting” to “validation”. Second, the path to the [parameter.dat](#) file was specified in the [path_optimized_parameter](#) field of the [validation](#) settings. Third, five new validation data sets were included: 1) the 2-dimensional dependence of the RMSD on $\langle \xi \rangle$ and $\langle \varphi \rangle$, 2) the 2-dimensional dependence of the RMSD on Δr and $\langle \xi \rangle$, 3) the 2-dimensional dependence of the RMSD on Δr and $\Delta \xi$, 4) the 2-dimensional dependence of the RMSD on Δr and $\langle \varphi \rangle$, 5) the 2-dimensional dependence of the RMSD on Δr and $\Delta \varphi$. Again, each of these data sets will contain 10^4 data points (see [Ns](#)).

- 2') *Run the program.* First, open Terminal (Linux) or Command Prompt (Windows) and navigate to the directory:

[cd ../DipFit](#)

If needed, set the permission properties (only for Linux):

[chmod 755 DipFit](#)

Finally, run the program by the following command:

sh DipFit.sh examples/example02_ls_iron(III)_trityl/config_ex02_2.cfg (for Linux)

DipFit.exe examples/example02_ls_iron(III)_trityl/config_ex02_2.cfg (for Windows)

- 3') During the program operation, the following messages should appear in Terminal or Command Prompt:

Reading out the configuration file... [DONE]

Validating the optimized fitting parameters...

Numerical error = 0.000294

Validation is finished. Total duration: 13:00:51.385063

Saving validation results into the directory:

.../DipFit/examples/example02_ls_iron(III)_trityl/2020-01-01_00-00/ [DONE]

Plotting validation results... [DONE]

Optimized fitting parameters:

| Parameter | Value | Optimized | Precision (+/-) |
|-----------------|--------|-----------|-----------------|
| r mean (nm) | 2.639 | Y | nan |
| r width (nm) | 0.056 | Y | nan |
| xi mean (deg) | 67.837 | Y | 28.503 |
| xi width (deg) | 0.020 | Y | nan |
| phi mean (deg) | 54.094 | Y | 34.365 |
| phi mean (deg) | 29.483 | Y | 17.836 |
| temperature (K) | 10.000 | N | nan |

- 4') *Explore the validation results.* During the validation, the new folder will be created in the directory /DipFit/examples/example02_ls_iron(III)_trityl and the results of validation will be saved in that folder. The output files include [parameter_errors-xi_mean-phi_mean.dat](#), [parameter_errors-r_width-xi_mean.dat](#), [parameter_errors-r_width-xi_width.dat](#), [parameter_errors-r_width-phi_mean.dat](#), [parameter_errors-r_width-phi_width.dat](#), and the corresponding graphical file [parameter_errors.png](#).

The content of [parameter_errors.png](#) is depicted in Figure 6.5. As can be seen, the minima of all RMSD plots appear at the same values of the fitting parameters as the minima in Figure 6.3c. Thus, the results of two validations are consistent with each other. In addition, Figure 6.5 reveals one interesting fact. The RMSD plot for the parameters Δr and $\langle\varphi\rangle$ shows not only one global minimum, which appears around the optimized values of both parameters ($\Delta r = 0.06$ nm, $\langle\varphi\rangle = 54^\circ$), but also one local minimum, which appears in the region with $\Delta r > 0.1$ nm and $\langle\varphi\rangle < 45^\circ$. The origin of the local minimum is likely to be explained by the fact that Δr and $\langle\varphi\rangle$ have similar effect of the simulated RIDME time trace. Thus, these two parameters might be

interdependent. In the given example, the interdependence between Δr and $\langle\varphi\rangle$ is small enough to allow finding the global minimum for both parameters.

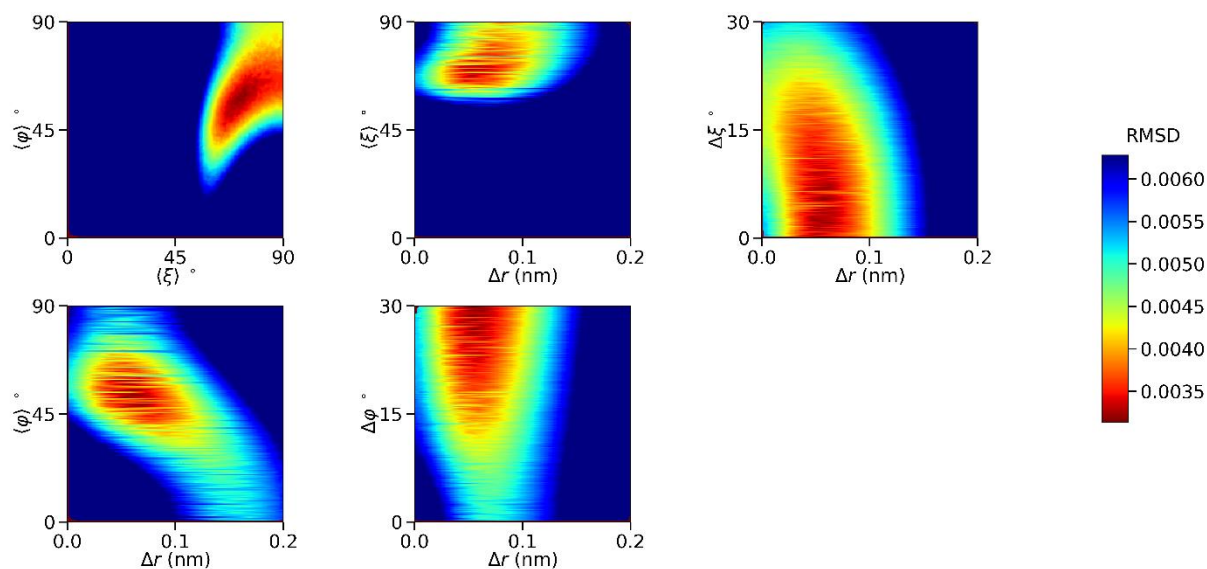


Figure 6.5. The RMSD in dependence of the different pairs of fitting parameters.

6.3 Simulation of the dipolar spectra of a high-spin Fe^{3+} -organic radical spin system

In the next example, the simulation mode of DipFit is used once again. The spectral simulations are done for the spin system consisting of the high-spin Fe^{3+} (hs Fe^{3+}) and the organic radical. In contrast to the ls Fe^{3+} considered in Chapter 6.1, the hs Fe^{3+} has $S = 5/2$ and, therefore, does not satisfy the requirements of DipFit (Chapter 1.2). However, if the zero-field splitting (ZFS) of the hs Fe^{3+} significantly exceeds its Zeeman energy and the thermal energy, the hs Fe^{3+} can be considered an effective $S = 1/2$ center,^[5] which makes DipFit still applicable. Thus, only this particular case is considered below. The g -factor of the hs Fe^{3+} is set to be axial with the principal components $g_{xx} = g_{yy} = 6.0$ (also denoted as g_{\perp}) and $g_{zz} = 2.0$ (also denoted as g_{\parallel}). The g -factor of the organic radical is assumed to be isotropic and equal to $g_e = 2.0023$. In analogy to the example from Chapter 6.1, the inter-spin-distance distance is fixed at 2.50 nm. Due to the axial symmetry of the hs Fe^{3+} center, the angle φ has no effect on the dipolar spectrum and, therefore, is fixed at 0° . In contrast, the angle ζ is expected to have an effect on the dipolar spectrum. Therefore, the spectral simulations will be done in dependence on the angle ζ , which will be varied in the range $[0^\circ, 90^\circ]$ with a constant step of 0.25° .

Additionally, the dipolar spectrum will be simulated in dependence of the angle θ between the inter-spin vector \vec{r} and the direction of the applied magnetic field \vec{B}_0 . This dependence will be acquired for two particular cases, $\zeta = 0^\circ$ and $\zeta = 90^\circ$. The idea behind of these simulations will become clear later.

Now, let's go step-by-step through the outlined simulations:

- 1) *Prepare the configuration file.* The detailed information on how one creates such a configuration file is given in Chapter 4. The ready-to-use configuration files can be found in the following directory:

[/DipFit/examples/example03_hs_iron\(III\)_nitroxide](#)

This directory contains three configuration files [config_ex03_i.cfg](#), where $i = 1, 2$, and 3 . The configuration file [config_ex03_1.cfg](#) initializes the spectral simulations in dependence of ζ , whereas the configuration files [config_ex03_2.cfg](#) and [config_ex03_3.cfg](#) initialize the spectral simulations in dependence of θ and set ζ to 0° and 90° , respectively. Note that the [faxis_normalized](#) parameter is set to 1 in all three configuration files, meaning that the frequency axis of the simulated spectra will be normalized by the dipolar coupling constant ν_0 of an isotropic spin pair with the same inter-spin distance as mentioned above.

- 2) *Run the program.* First, open Terminal (Linux) or Command Prompt (Windows) and navigate to the directory:

[cd .../DipFit](#)

If needed, set the permission properties (only for Linux):

`chmod 755 DipFit`

Finally, run the program by the following command:

`sh DipFit.sh examples/example03_hs_iron(III)_nitroxide/config_ex03_i.cfg` (for Linux)

`DipFit.exe examples/example03_hs_iron(III)_nitroxide/config_ex03_i.cfg` (for Windows)

- 3) *Monitor the progress status.* During the program operation, the following messages should appear in Terminal or Command Prompt:

For `config_ex03_1.cfg`:

Reading out the configuration file... [DONE]

Starting simulation...

Running pre-calculations... [DONE]

Calculating dipolar spectrum vs ξ ... 100% [DONE]

Simulation is finished

For `config_ex03_2.cfg` and `config_ex03_3.cfg`:

Reading out the configuration file... [DONE]

Starting simulation...

Running pre-calculations... [DONE]

Calculating dipolar spectrum vs θ ... 100% [DONE]

Simulation is finished

- 4) *Explore the simulation results.* During the simulation, the new folder will be created within the directory `/DipFit/examples/example03_hs_iron(III)_nitroxide` and the results of simulations will be saved in that folder. For the simulation based on `config_ex03_1.cfg`, the folder will contain the files `spc_vs_xi.dat` and `spc_vs_xi.png`. The content of `spc_vs_xi.png` will pop up on the screen after the simulation (Figure 6.6a). For the simulations based on `config_ex03_2.cfg` and `config_ex03_3.cfg`, the folder will contain the files `spc_vs_theta.dat` and `spc_vs_theta.png`. Again, the content of the graphical file will be appear on the screen (Figure 6.6b-c).

As can be seen in Figure 6.6a, the simulated spectra have prominent differences to the Pake doublet.^[10] Firstly, they are significantly broader than the Pake doublet. This stems from the fact that the dipolar coupling frequencies are proportional to the effective g -factors of the hs Fe^{3+} , which are 3 times larger than g_e for two out of three canonical orientations. Secondly, the shapes of the simulated spectra differ from the shape of the usual Pake doublet and, as expected, depend strongly on the angle ξ . In order to provide a deeper insight into the obtained shapes, the dipolar spectrum is plotted as a function of θ for two particular cases, $\xi = 0^\circ$ and $\xi = 90^\circ$. In

the case of $\zeta = 90^\circ$ (Figure 6.6b), the parallel component ($\theta = 90^\circ$) of the spectrum scales with g_{yy} of the hs Fe^{3+} , which is about 3 times larger than g_e . Consequently, the parallel component appears in the spectrum at $\pm 6\nu_0$ and not at $\pm 2\nu_0$ as for the Pake doublet. The perpendicular component ($\theta = 0^\circ$) of the spectrum scales with the g -values of the hs Fe^{3+} ranging from $g_{zz} \approx g_e$ to $g_{xx} \approx 3g_e$. This gives rise to two features in the spectrum at $\pm\nu_0$ and $\pm 3\nu_0$, respectively. In the case of $\zeta = 0^\circ$ (Figure 6.6c), the perpendicular component ($\theta = 90^\circ$) of the spectrum scales with g_{xx} and g_{yy} of the hs Fe^{3+} . Since $g_{xx} = g_{yy} \approx 3g_e$, the perpendicular component of the spectrum appears at $\pm 3\nu_0$. The parallel component ($\theta = 0^\circ$) of this spectrum scales with g_{zz} of the hs Fe^{3+} and, thus, should appear at $\pm 2\nu_0$. Since the parallel component has a smaller probability and appears at lower frequencies than the perpendicular component, it does not yield any prominent features in the spectrum.

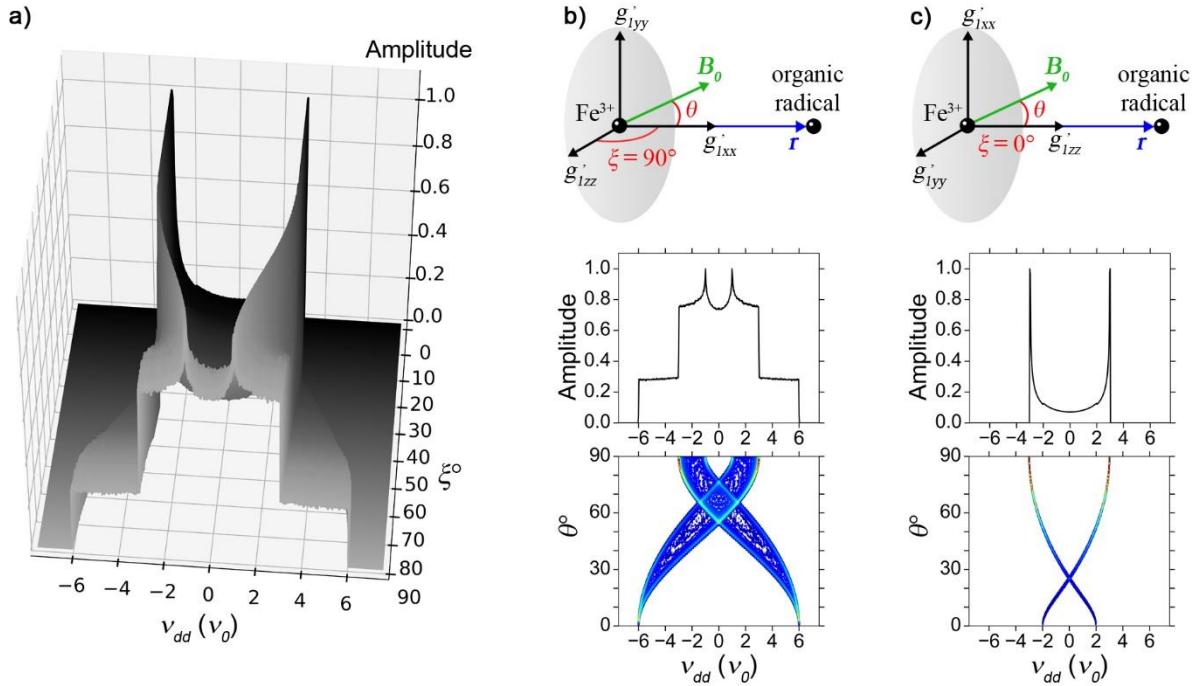


Figure 6.6. a) The angular dependence of the dipolar spectrum, which corresponds to the spin system consisting of an isotropic organic radical with $g_{iso} = g_e$ and an anisotropic hs Fe^{3+} with $g_{aniso} = [6.0, 6.0, 2.0]$. The angle ζ is defined in Figure 1.1. For the particular cases of the b) $\zeta = 0^\circ$ and c) $\zeta = 90^\circ$, the dipolar spectrum is also shown in dependence of the angle θ . The geometric model of the spin system with the fixed ζ angle. Middle: The corresponding dipolar spectrum. Bottom: The corresponding dipolar spectrum plotted against θ .

6.4 Fitting of the dipolar spectrum of a high-spin Fe^{3+} -nitroxide spin system

In the last example, DipFit is applied for the fitting of the RIDME spectrum, which was acquired on the MTSL-labeled mutant Q8R1 of the hemeprotein met-myoglobin (Figure 6.7).^[5] This met-myoglobin mutant contains two spatially separated electron spin centers, the nitroxide and the high-spin Fe^{3+} (hs Fe^{3+}). The hs Fe^{3+} has the large axial ZFS ($D \sim 9.26 \text{ cm}^{-1}$, $E = 0.0023 \text{ cm}^{-1}$),^[13] which allows considering this ion as an effective $S = 1/2$ center at Q-band and at the temperatures below 3 K.^[5] The g -factor of this center is very anisotropic and has the principal components $g_{xx} = 5.93$, $g_{yy} = 5.94$, and $g_{zz} = 2.00$. In contrast, the g -anisotropy of the nitroxide center is so small that it can be

neglected in the data analysis by setting all g -values to g_e . Since the g -factor of the hs Fe^{3+} is almost axial, the RIDME spectrum of Q8R1 is independent from the angle φ . Therefore, this angle will be excluded from the fitting and fixed at a constant value of 0° . Thus, the fitting of the RIDME spectrum will be done using the parameters of only two distributions, $P(r)$ and $P(\zeta)$, as fitting parameters. In addition, the temperature of the RIDME experiment will be used as a fitting parameter. This has the following reasons: 1) if the Zeeman energy of the hs Fe^{3+} becomes comparable to the thermal energy, the corresponding RIDME spectrum becomes temperature dependent (for explanation see Ref. ^[5]), 2) the conditions, used for the RIDME measurements on Q8R1, apply to the case described in 1), and 3) the precision of the experimentally measured temperature was above 1 K, which is not sufficient for accurate spectral simulations.

Now, let's go step-by-step through the fitting:

- 1) *Prepare the configuration file.* The detailed information on how one creates such a configuration file is given in Chapter 4. The ready-to-use configuration file [config_ex04.cfg](#) can be found in the following directory:

[/DipFit/examples/example04_hs_iron\(III\)_nitroxide](#)

As follows from this configuration file, the data, which is going to be fitted (see [fitted_data](#)), is the RIDME spectrum of Q8R1 (see [path_spectrum](#)). The mean values and

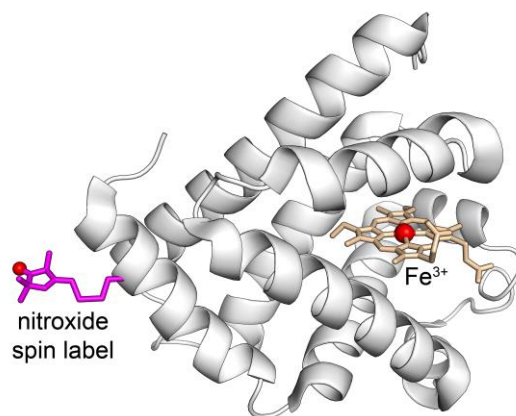


Figure 6.7. The structural model of the met-myoglobin mutant Q8R1. This model is based on the crystal structure of met-myoglobin (PDB-ID 1wla)^[14], which is shown in white. The heme group and the Fe^{3+} ion are depicted as light orange sticks and a red sphere, respectively. A single conformer of the nitroxide spin label, attached at the position Q8, is depicted by magenta sticks. The corresponding oxygen atom is shown as a red sphere.

the widths of the distributions $P(r)$ and $P(\zeta)$, as well as the temperature, are selected to be the fitting parameters (see [fitting_variables](#)). The distributions $P(r)$ and $P(\zeta)$ are approximated by normal distributions (see [calculation_settings](#)) and, therefore, their widths are given by standard deviations. The genetic algorithm is set to be the fitting method. Its internal parameters are provided in [fitting_settings](#). Among them, the total number of optimization steps is set to 300. For [validation](#), the recording of the following data sets is activated: 1) the 2-dimensional dependence of the RMSD on $\langle r \rangle$ and Δr , 2) the 2-dimensional dependence of the RMSD on $\langle \zeta \rangle$ and $\Delta \zeta$, and 3) the 2-dimensional dependence of the RMSD on the temperature. Each of these data sets will contain 10^4 data points (see [Ns](#)). The errors of the fitting parameters will be determined at the 1σ level (see [threshold](#)).

- 2) *Run the program.* First, open Terminal (Linux) or Command Prompt (Windows) and navigate to the directory:

```
cd ../DipFit
```

If needed, set the permission properties (only for Linux):

```
chmod 755 DipFit
```

Finally, run the program by the following command:

```
sh DipFit.sh examples/example04_hs_iron(III)_nitroxide/config_ex04.cfg (for Linux)
```

```
DipFit.exe examples/example04_hs_iron(III)_nitroxide/config_ex04.cfg (for Windows)
```

- 3) *Monitor the progress status.* During the program operation, the following messages should appear in Terminal or Command Prompt:

```
Reading out the configuration file... [DONE]
```

```
Starting fitting...
```

```
Optimization step 300 / 300: RMSD = 0.027326
```

```
Fitting is finished. Total duration: 10:00:49.371723
```

```
Saving fitting results into the directory:
```

```
../DipFit/examples/example04_hs_iron(III)_nitroxide/2019-01-01_00-00/ [DONE]
```

```
Plotting fitting results... [DONE]
```

```
Validating the optimized fitting parameters...
```

```
Numerical error = 0.002344
```

```
Validation is finished. Total duration: 9:06:06.032284
```

```
Saving validation results into the directory:
```

```
../DipFit/examples/example04_hs_iron(III)_nitroxide/2020-01-01_00-00/ [DONE]
```


Plotting validation results... [DONE]

Optimized fitting parameters:

| Parameter | Value | Optimized | Precision (+/-) |
|-----------------|--------|-----------|-----------------|
| r mean (nm) | 2.745 | Y | 0.297 |
| r width (nm) | 0.086 | Y | 0.052 |
| xi mean (deg) | 80.236 | Y | 16.118 |
| xi width (deg) | 1.630 | Y | 19.781 |
| phi mean (deg) | 0.000 | N | nan |
| phi mean (deg) | 0.000 | N | nan |
| temperature (K) | 2.018 | Y | 0.323 |

Here, a few comments have to be made. First, the calculation time strongly depends on the PC used (Chapter 1.3). Second, the optimized fitting parameters might slightly differ between single fitting trials. Importantly, this deviation is well below the error of the optimized parameters (see below).

- 4) *Explore the fitting results.* During the fitting, the new folder will be created within the directory [/DipFit/examples/example04_hs_iron\(III\)_nitroxide](#) and the results of fitting will be saved in that folder. The output files include [score.dat](#), [fit.dat](#), [parameters.dat](#), [parameter_errors-r_mean-r_width.dat](#), [parameter_errors-xi_mean-xi_width.dat](#), [parameter_errors-temp.dat](#), and the corresponding graphical files (see Chapter 5.2).

Let's briefly discuss the content of the output files. The content of the files [score.dat](#) and [score.png](#) is depicted in Figure 6.8a. This figure shows that the RMDS decreased during first 50 optimization steps and, then, reached a plateau during next 250 optimization steps. Based on this, one can conclude that genetic algorithm has converged to the global minimum. The content of the files [fit.dat](#) and [fit.png](#) reveals that the good fit to the RIDME spectrum was obtained (Figure 6.8b) and, consequently, the optimized distributions $P(r)$ and $P(\xi)$ have a good agreement with the experimental RIDME data. The parameters of these distributions, as well as the optimized value of the temperature, are stored in the file [parameters.dat](#) and listed here in Figure 6.8c. In order to estimate how defined these parameters are, three validation data sets were recorded: 1) the 2-dimensional dependence of the RMSD on $\langle r \rangle$ and Δr ([parameter_errors-r_mean-r_width.dat](#)), 2) the 2-dimensional dependence of the RMSD on $\langle \xi \rangle$ and $\Delta \xi$ ([parameter_errors-xi_mean-xi_width.dat](#)), and 3) the 1-dimensional dependence of the RMSD on the temperature ([parameter_errors-temp.dat](#)). The picture of these validation data sets is stored in the file [parameter_errors.png](#) and shown here as Figure 6.8d. As can be seen, all three RMSD plots have a single well-defined minimum, whose coordinates are in excellent agreement the optimized fitting parameters.

In order to estimate the errors of the optimized fitting parameters, the 1σ confidence intervals of these parameters were calculated. In these calculations, the dependence of the RMSD on each single fitting parameter was extracted from Figure 6.8d and, then, fitted by a Gaussian (Figure 6.9). This yielded the parameters' errors given in Figure 6.8c. The errors of the mean inter-spin distance $\langle r \rangle$ and its standard deviation Δr are ± 0.3 nm and ± 0.05 nm, respectively. The error of the angular parameters $\langle \zeta \rangle$ and $\Delta \zeta$ is about $\pm 20^\circ$. Thus, both distributions, $P(r)$ and $P(\zeta)$, could be extracted from the RIDME spectrum of Q8R1 with a reasonable precision. The error obtained for the temperature is 0.3 K. Consequently, the optimized value of the temperature deviates from the experimentally measured value by 1.0 ± 0.3 K. This deviation can be assigned to the error of the experimentally measured temperature, which was expected to be above 1 K.

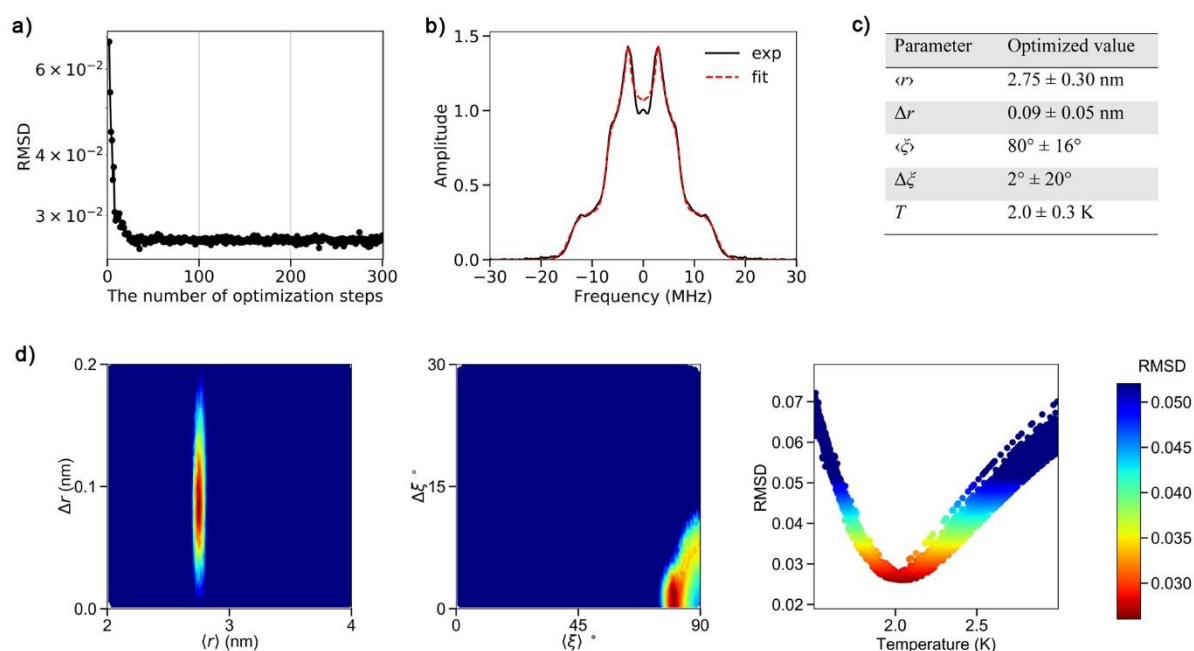


Figure 6.8. Fitting the RIDME spectrum of Q8R1 by DipFit. **a)** The RMSD in dependence of the optimization step. **b)** The RIDME spectrum (black line) is overlaid with the corresponding fit (red line). **c)** The optimized values of the fitting parameters. **d)** The RMSD in dependence of the different fitting parameters.

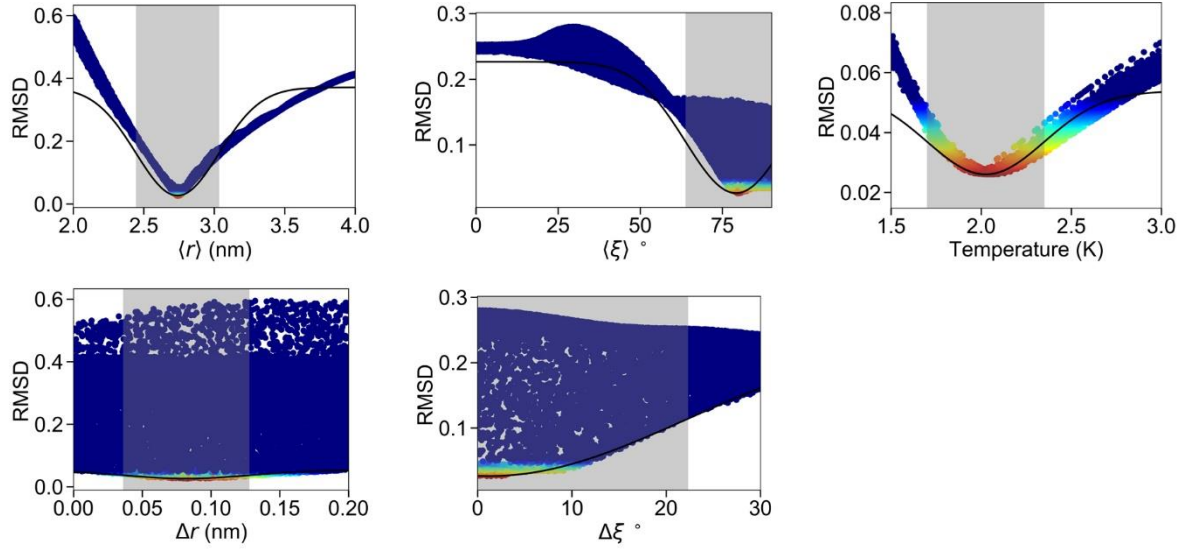


Figure 6.9. The RMSD in dependence of single fitting parameters (colored dots). The color code is identical to Figure 6.8. The Gaussian fits to this data is depicted by black lines. The parameters' ranges, which correspond to the 1σ confidence intervals, are depicted by gray bars.

7 References

- [1] P. P. Borbat, J. H. Freed, *Struct. Bond.* **2013**, *152*, 1–82.
- [2] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, *Appl. Magn. Reson.* **2006**, *30*, 473–498.
- [3] D. Abdullin, O. Schiemann, *ChemPlusChem* **2020**, 353–372.
- [4] A. F. Bedilo, A. G. Maryasov, *J. Magn. Reson.* **1995**, *116*, 87–96.
- [5] D. Abdullin, H. Matsuoka, M. Yulikov, N. Fleck, C. Klein, S. Spicher, G. Hagelueken, S. Grimme, A. Luetzen, O. Schiemann, *Chem. Eur. J.* **2019**, *25*, 8820–8828.
- [6] D. Abdullin, P. Brehm, N. Fleck, S. Spicher, S. Grimme, O. Schiemann, *Chem. – A Eur. J.* **2019**, *25*, 14388–14398.
- [7] B. Filipič, J. Štrancar, *Appl. Soft Comput.* **2001**, *1*, 83–90.
- [8] T. Spalek, P. Pietrzyk, Z. Sojka, *J. Chem. Inf. Model.* **2005**, *45*, 18–29.
- [9] S. Stoll, A. Schweiger, *J. Magn. Reson.* **2006**, *178*, 42–55.
- [10] O. Schiemann, T. F. Prisner, *Q. Rev. Biophys.* **2007**, *40*, 1–53.
- [11] L. V. Kulik, S. A. Dzuba, I. A. Grigoryev, Y. D. Tsvetkov, *Chem. Phys. Lett.* **2001**, *343*, 315–324.
- [12] S. Milikisyants, F. Scarpelli, M. G. Finiguerra, M. Ubbink, M. Huber, *J. Magn. Reson.* **2009**, *201*, 48–56.
- [13] M. Fittipaldi, I. García-Rubio, F. Trandafir, I. Gromov, A. Schweiger, A. Bouwen, S. Van Doorslaer, *J. Phys. Chem. B* **2008**, *112*, 3859–3870.
- [14] R. Maurus, C. M. Overall, R. Bogumil, Y. Luo, A. G. Mauk, M. Smith, G. D. Brayer, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1997**, *1341*, 1–13.