

DipFit

User Manual

Dinar Abdullin
Olav Schiemann

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	11
4.6	Validation settings	13
4.7	Calculation settings.....	15
4.8	Output settings	16
5	Output data.....	16
5.1	Simulation output	16
5.2	Fitting output	17
6	Examples.....	19
6.1	Simulation of the dipolar spectra of a low-spin Fe^{3+} -organic radical spin system....	19
6.2	Fitting of the PDS time trace of a low-spin Fe^{3+} -trityl spin system	22
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.....	35

1 Introduction

1.1 What is it for?

Pulsed electron paramagnetic resonance dipolar spectroscopy (PDS)^[1] offers several techniques for measuring the dipolar coupling between electron spin centers. The PDS signal is 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 Fourier transform. For disordered samples, 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 isotropic g -factors of 2.0023, the PDS data analysis can be done by means of the program DeerAnalysis.^[2] This program is also applicable to $S = 1/2$ centers, which have a small g -anisotropy (difference in g -values 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 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. ^[3–5]. 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.^[4]

The 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 -tensor, 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 ζ , 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 , ζ , 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 , ζ , and φ is usually unknown and, therefore, is omitted in DipFit.

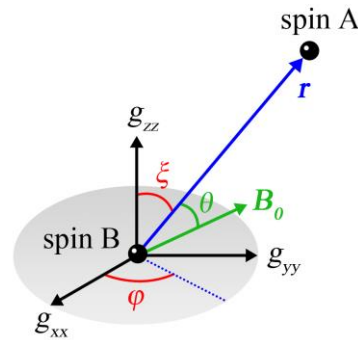


Figure 1.1. The DipFit model of a spin system.

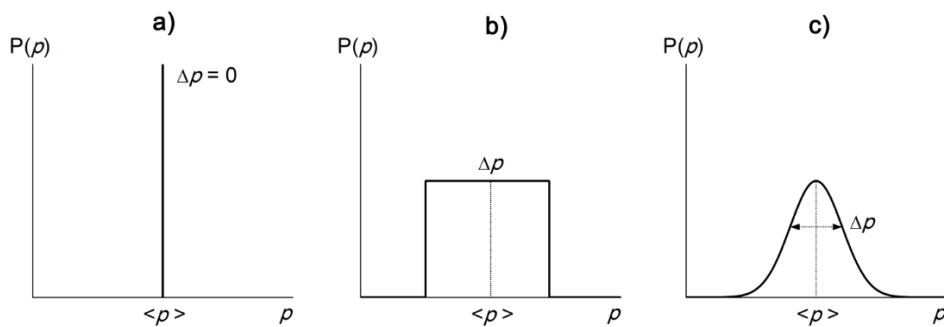


Figure 1.2. The distribution types used for the DipFit parameter $p \in (r, \zeta, \varphi)$. **a)** A single value $\langle p \rangle$. The width Δp equals to 0. **b)** Uniform distribution with the mean value $\langle p \rangle$ and the width Δp . **c)** Normal distribution with the mean value $\langle p \rangle$ 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) error analysis 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 precision of an optimized geometric model and its parameters is estimated through the error analysis. In addition, DipFit has an operation mode called “error analysis” 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 error analysis (Chapter 4.6).

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

Source: DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
mode = 1;
```

where

- | | |
|-----------------------|------------------------------------|
| <code>mode = 0</code> | Activates the simulation mode. |
| <code>mode = 1</code> | Activates the fitting mode. |
| <code>mode = 2</code> | Activates the error analysis mode. |

4.2 Experimental data

Experimental data, which can be analyzed by DipFit, can be either the background-free PDS time trace or the PDS spectrum. The file paths of both data types can be provided in the following way:

Source: *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 error analysis 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.

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:

Source: *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 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 error analysis modes.

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

Source: 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 mod_depth . |
| spc_vs_theta | If set to 1, 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 (Figure 1.1). |
| 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 RIDME data acquired at liquid helium temperatures. ^[4] |

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_parameters](#):

Source: DipFit/examples/example01_ls_iron(III)_nitroxide/config_ex01_1.cfg

```
simulation_parameters:
{
  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 shape of $P(r)$ is specified in the calculation_settings list (Chapter 7).
xi_mean	The mean value $\langle \xi \rangle$ of the angular distribution $P(\xi)$. Given in [degree].
xi_width	The width $\Delta \xi$ of the angular distribution $P(\xi)$ (Figure 1.2). Given in [degree]. The shape of $P(\xi)$ 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 shape 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](#):

Source: 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];
  mod_depth = 0.5;
  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 = [ζ_{min} , ζ_{max} , N_ζ], 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 [ζ_{min} , ζ_{max}].
phi_ranges	Defines the φ axis for the spc_vs_phi simulation. This axis is defined by three numbers: phi_ranges = [φ_{min} , φ_{max} , N_φ], 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 [φ_{min} , φ_{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}].
mod_depth	The modulation depth of the simulated time trace. It needs to be given when the simulation of the time trace is activated (timetrace = 1) and no experimental time trace was defined (path_timetrace = "").
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_parameters 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 error analysis modes but can be omitted for the simulation mode.

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

Source: DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
fitting_parameters:
{
  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 };
};
```

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.^[6–9] 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 extensive tests of the genetic algorithm on numerous PDS data sets.

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	nm
r_width	The width Δr of the inter-spin distance distribution $P(r)$	Any reasonable ranges	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]	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]	degree
temp ^[b]	The temperature of the PDS experiment	Any reasonable ranges	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.^[4]

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

Source: *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).^[9]

<code>size_generation</code>	The number of parameter sets (chromosomes) that are evaluated simultaneously by the genetic algorithm. ^[9] A good practice is to set this parameter to 10 x (the number of fitting parameters).
<code>prob_crossover</code>	The crossover rate. ^[9]
<code>prob_mutation</code>	The mutation rate. ^[9]

4.6 Error analysis settings

These settings are relevant for the fitting and error analysis 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 to perform an error analysis for all fitting parameters. Since DipFit deals with up to seven fitting parameters simultaneously, the parameter space is very large and cannot be fully covered in reasonable time. Therefore, the error analysis 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 the fitting parameters are then used to determine the parameter ranges, within which the RMSD values do not exceed the user-defined threshold (see [threshold](#)). Below this threshold, the fitting parameters are assumed to be undefined and, therefore, the obtained ranges correspond to the uncertainty ranges of the fitting parameters. In order to generalize the selection of the threshold value for different PDS data sets, the threshold is set in percent of the minimal RMSD obtained in each fitting. For example, the threshold of 120 % is used in all examples from Chapter 6. Finally, the uncertainty ranges of the fitting parameters are converted into the errors of those parameters by calculating the largest deviation of each parameters from its optimized value within the corresponding uncertainty ranges.

The settings of the error analysis have to be specified in the list called [error_analysis](#):

Source: DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_1.cfg

```
error_analysis:
{
  variables = (("r_mean", "r_width"), ("xi_mean", "xi_width"), ("phi_mean", "phi_width"));
  Ns = 10000;
  threshold = 120;
  path_optimized_parameters = "";
};
```

where

variables	The list of fitting parameters used for the error analysis. If no error analysis is wanted, this list should be left empty: <code>variables = ()</code> . Otherwise, the variables of each separate error calculation should be enclosed in round brackets, e.g., <code>variables = (("r_mean", "r_width"), ("xi_mean", "xi_width"), ("temp"))</code> . 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 $\langle \xi \rangle$ and $\Delta \xi$, 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) While recording the dependence of the RMSD on the selected fitting parameter(s), the latter one(s) will be randomly set to Ns different values within the ranges specified in the <code>fitting_parameters</code> list.
Ns	The number of points in each validation data set.
threshold	The RMSD threshold used to determine the uncertainty ranges of the fitting parameters. Given in percent of the minimal RMSD obtained in each fitting. Thus, it should be larger than 100.

As mentioned in Chapter 4.1, DipFit has an additional operation mode called “error analysis” mode. The idea behind this mode is the same as for the error analysis described above. The only difference is that the error analysis is now done post factum, i.e., using the results of the fitting, which was completed in the past. Thus, the error analysis mode can be useful in cases, when the user has done the fitting of some PDS data in the past and later wants to acquire a few additional dependencies of the RMSD on fitting parameters for the same PDS data. 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 `error_analysis` list:

Source: DipFit/examples/example02_ls_iron(III)_trityl/config_ex02_2.cfg

```
error_analysis:
{
  variables = (("xi_mean", "phi_mean"), ("r_width", "xi_mean"), ("r_width", "xi_width"),
              ("r_width", "phi_mean"), ("r_width", "phi_width"));
  Ns = 10000;
  threshold = 120;
  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`:

Source: *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;
  g_selectivity = 0;
  magnetic_field = 0;
};
```

where

<code>Ns</code>	The number of Monte-Carlo samples used in numerical integration. It determines how accurately the PDS time trace or the PDS spectra will be simulated. ^[5] Do not set this value below 10^5 !
<code>r_distr</code>	The shape of the inter-spin distance distribution $P(r)$: “uniform” or “normal” (Figure 1.2).
<code>xi_distr</code>	The shape of the angular distribution $P(\xi)$: “uniform” or “normal” (Figure 1.2).
<code>phi_distr</code>	The shape of the angular distribution $P(\varphi)$: “uniform” or “normal” (Figure 1.2).
<code>fmin</code>	The minimal frequency of the simulated spectrum. Given in [MHz].
<code>fmax</code>	The maximal frequency of the simulated spectrum. Given in [MHz].
<code>tmin</code>	The minimal time value of the simulated time trace. Given in [μ s].
<code>tmax</code>	The maximal time value of the simulated time trace. Given in [μ s].
<code>g_selectivity</code>	If set to 1, the difference in spin polarization for the different g -values of spin B will be taken into account. ^[4] This parameter is relevant only for the RIDME data acquired at liquid helium temperatures.
<code>magnetic_field</code>	The value of the applied magnetic field used in the PDS experiment. Given in [T]. This parameters has to be provided only if <code>g_selectivity = 1</code> .

4.8 Output settings

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

Source: 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/error analysis 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/error analysis will be saved into DAT files (Chapter 5).
save_figures	The graphical data obtained from simulation/fitting/error analysis 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 values in [MHz];
next columns – a spectrum which is simulated with the **X** value given in the first 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 PDS time trace or an experimental PDS spectrum.

Content: If [fitted_data](#) = “[spectrum](#)”:
column 1 – frequency values in [MHz];
column 2 – an experimental PDS spectrum;
column 3 – a fit.

If [fitted_data](#) = “[timetrace](#)”:
column 1 – time points in [μ s];
column 2 – an experimental PDS time trace;
column 3 – a fit.

- [parameters.dat](#)

Data: the optimized values of fitting parameters.

Content: column 1 – the names and the 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_parameters](#) list);
 column 4 – the errors 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 [error_analysis](#) 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 a spin system consisting of a low-spin Fe^{3+} (1s Fe^{3+}) and an organic radical. The principal g -values of the 1s 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.

The workflow of this simulation includes the following steps:

- 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 the simulation...

Running the pre-calculations... [DONE]

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

The simulation is finished

Saving the simulation results into the directory:

.../DipFit/examples/example01_ls_iron(III)_nitroxide/2020-01-01_00-00/ [DONE]

- 4) *Explore the simulation results.* During the simulation, a 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, 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 $(g_{zz}/g_e)v_0 \approx 1.45 v_0$.

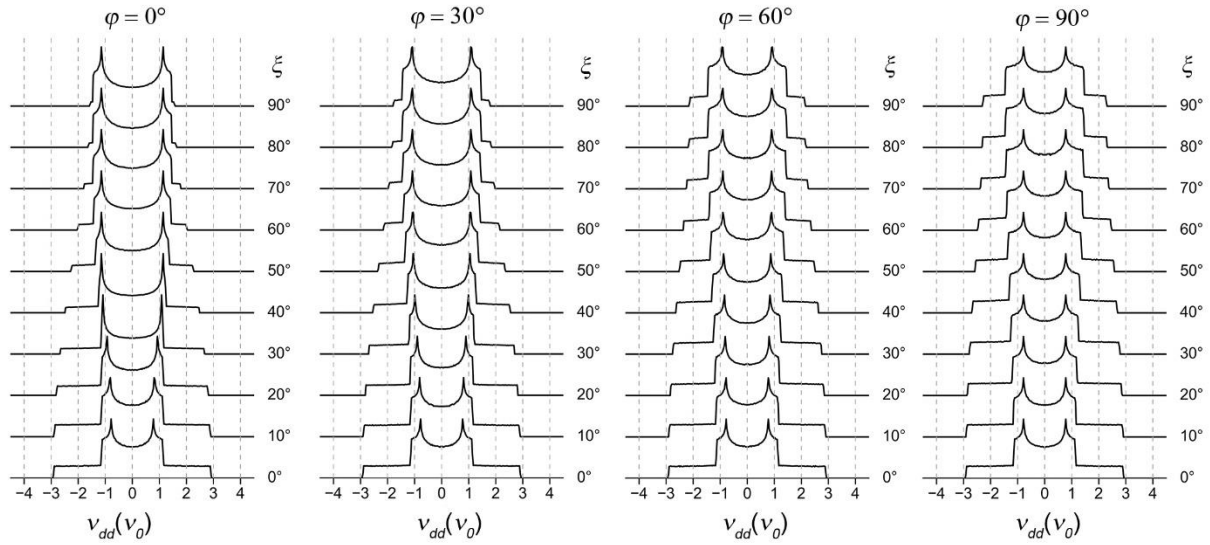


Figure 6.1. The angular dependence of the dipolar spectrum, which corresponds to the spin system consisting of an isotropic organic radical with $g = g_e$ and an anisotropic $1s \text{ Fe}^{3+}$ with $g = [1.56, 2.28, 2.91]$. The angles ζ and φ are defined in Figure 1.1.

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

In the second example, the fitting and error analysis modes of DipFit will be explored. The PDS data, which is used for this example, was acquired on the model compound **1**^[5] (Figure 6.2). This compound contains two spatially separated electron spin centers, the trityl radical and the low-spin Fe³⁺ (ls 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 pulse sequence relaxation induced dipolar modulation enhancement (RIDME)^[11,12] and yielded the background-free time trace shown in Figure 6.3b. Here, this time trace will be fitted using the parameters of $P(r)$, $P(\xi)$, and $P(\varphi)$ as fitting parameters.

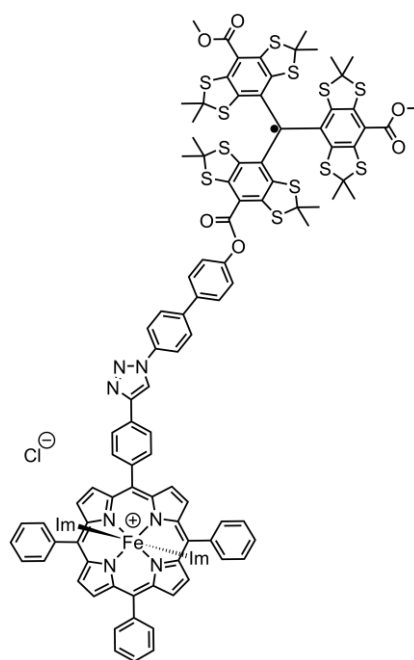


Figure 6.2. The chemical structure of the model compound **1**.

The workflow of this fitting includes the following steps:

- 1) *Prepare the configuration file.* The detailed information on how to create 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_ls_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 $P(r)$, $P(\xi)$, and $P(\varphi)$ are set to be the fitting parameters (see [fitting_parameters](#)). 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 [error analysis](#), 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 $\langle \xi \rangle$ and $\Delta \xi$, 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 RMSD level, which corresponds to 120 % of the lowest RMSD (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/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 the fitting...
```

```
Optimization step 500 / 500: RMSD = 0.003050
```

```
The fitting is finished. Total duration: 15:56:24.790875
```

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

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

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

```
Starting the error analysis...
```

```
Numerical error = 0.000283
```

```
Minimal RMSD = 0.002971
```

```
RMSD threshold = 0.003566
```

```
Calculating the RMSD in dependence of fitting parameters ...
```

```
Calculation step 3 / 3
```

```
Calculating the errors of fitting parameters ...
```

```
The error analysis is finished. Total duration: 8:57:43.418870
```

```
Saving the results of the error analysis into the directory:
```

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

```
Plotting the results of the error analysis... [DONE]
```

```
Optimized fitting parameters:
```

Parameter	Value	Optimized	Precision (+/-)
r mean (nm)	2.636	Y	0.016
r width (nm)	0.052	Y	0.027
xi mean (deg)	89.667	Y	26.570
xi width (deg)	21.766	Y	21.701
phi mean (deg)	89.128	Y	41.618
phi mean (deg)	49.394	Y	33.700
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 values of the fitting parameters might slightly differ between single fitting trials. Importantly, this deviation is below the error of the fitting parameters (see below).

- 4) *Explore the fitting results.* During the fitting, a new folder will be created inside the directory `/DipFit/examples/example02_ls_iron(III)_trityl` and the results of the 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).

Next, the results of the fitting will be briefly discussed. First, it is important to make sure that the genetic algorithm has converged to the global minimum. This information can be obtained from the output files `score.dat` and `score.png`, which contain the dependence of the RMSD on the optimization step. In the present fitting, the RMSD between the RIDME time trace and its fit fell gradually down during the first 140 optimization steps and, after this, did not change significantly during the last 360 optimization steps (Figure 6.3a). This fact shows that the global minimum was reached and, consequently, the best fit to the experimental RIDME time trace was 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 output 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, a good fit to the experimental RIDME time trace was indeed obtained.

The optimized values of the fitting parameters are stored in the output file `parameters.dat`. The content of this file is listed in Figure 6.3c. In order to estimate the errors of these parameters, the following data was 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 data sets is stored in the output file `parameter_errors.png` and shown here in Figure 6.3d. In this figure, the uncertainty ranges of the fitting parameters are depicted in dark red. Recall that the uncertainty ranges were determined at the RMSD level, which corresponds to 120 % of the minimal RMSD. In addition, the optimized values of the fitting parameters are depicted by white circles. As expected, all circles appear within the uncertainty ranges mentioned above,

meaning that the optimized values of the fitting parameters lay within the corresponding uncertainty ranges. The errors of individual fitting parameters, which were derived from the uncertainty intervals, are listed Figure 6.3c. The errors of the mean inter-spin distance $\langle r \rangle$ and its standard deviation Δr are ± 0.02 nm and ± 0.03 nm, respectively. The errors of the mean angles $\langle \zeta \rangle$ and $\langle \varphi \rangle$ equal to $\pm 27^\circ$ and $\pm 42^\circ$, respectively. The corresponding standard deviations $\Delta \zeta$ and $\Delta \varphi$ have slightly smaller errors, which equal to $\pm 22^\circ$ and $\pm 34^\circ$, respectively.

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 sub-angstrom precision. In addition, the angular distributions $P(\zeta)$ and $P(\varphi)$ could be determined with a moderate precision. Note that the precision of the angular parameters strongly depends on how anisotropic the g -factor of [spin B](#) is (in the given example, spin B is the $1s \text{ Fe}^{3+}$). It will be shown in Section 6.4 that the precision of the angular parameters can be significantly increased if spin B has a significantly larger g -anisotropy than the $1s \text{ Fe}^{3+}$.

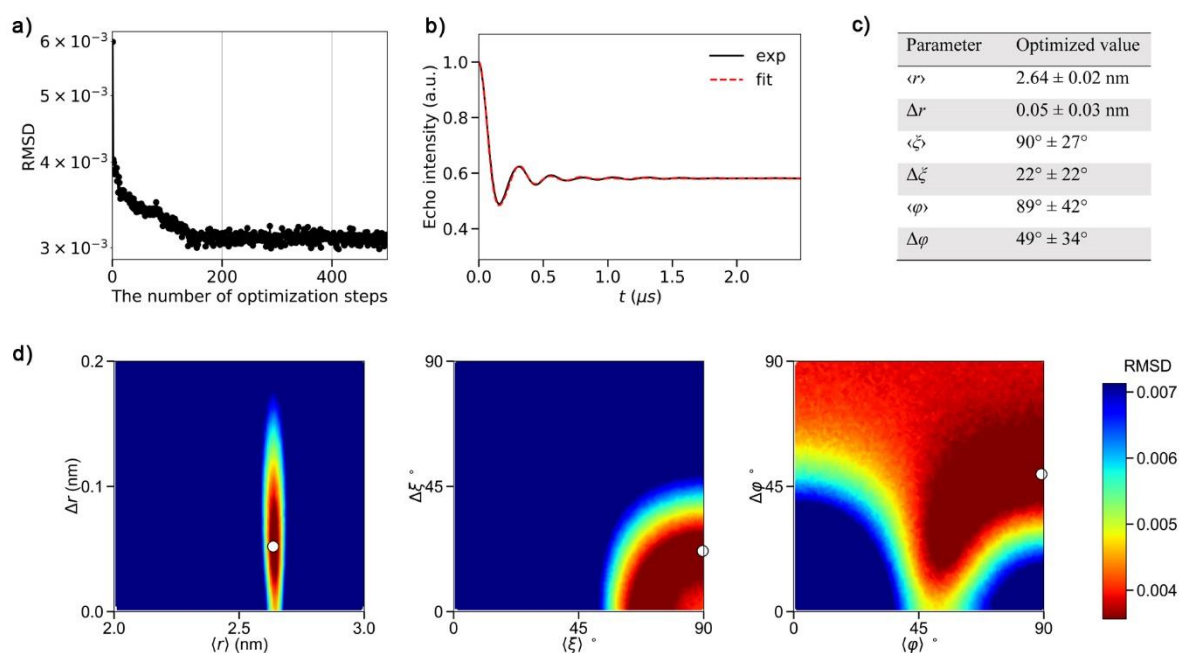


Figure 6.3. Fitting the RIDME time trace of **1** by DipFit. **a)** The RMSD in dependence of optimization step. **b)** The RIDME time trace (black line) is overlaid with the corresponding fit (red line). **c)** The optimized values of the fitting parameters. **d)** The RMSD in dependence of different pairs of fitting parameters. The dark red regions correspond to the RMSD values, which are below 120 % of the minimal RMSD. The optimized values of the fitting parameters are depicted by white dots.

In the second part of this example, the results of the fitting described above will be used to test the error analysis mode of DipFit. As was mentioned in Chapter 4.6, the idea behind this mode is identical to the already mentioned error analysis. The only difference is that this analysis will be done post factum, i.e., using the results of the fitting, which was completed in the past. In order to run DipFit in the error analysis mode, one has to do 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 “error analysis”. Second, the path to the existing [parameter.dat](#) file was specified in [path_optimized_parameters](#). Third, the recording of five new data sets is activated: 1) the 2-dimensional dependence of the RMSD on $\langle \zeta \rangle$ and $\langle \varphi \rangle$, 2) the 2-dimensional dependence of the RMSD on Δr and $\langle \zeta \rangle$, 3) the 2-dimensional dependence of the RMSD on Δr and $\Delta \zeta$, 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... \[DONE\]](#)

[Starting the error analysis...](#)

[Numerical error = 0.000241](#)

[Minimal RMSD = 0.002986](#)

[RMSD threshold = 0.003583](#)

[Calculating the RMSD in dependence of fitting parameters ...](#)

[Calculation step 5 / 5](#)

[Calculating the errors of fitting parameters ...](#)

[The error analysis is finished. Total duration: 13:03:03.284632](#)

Saving the results of the error analysis into the directory:

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

Plotting the results of the error analysis... [DONE]

Optimized fitting parameters:

Parameter	Value	Optimized	Precision (+/-)
r mean (nm)	2.636	Y	nan
r width (nm)	0.052	Y	0.041
xi mean (deg)	89.667	Y	20.029
xi width (deg)	21.766	Y	15.717
phi mean (deg)	89.128	Y	44.811
phi mean (deg)	49.394	Y	40.586
temperature (K)	10.000	N	nan

4') *Explore the validation results.* During the validation, a 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 6.4. 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.

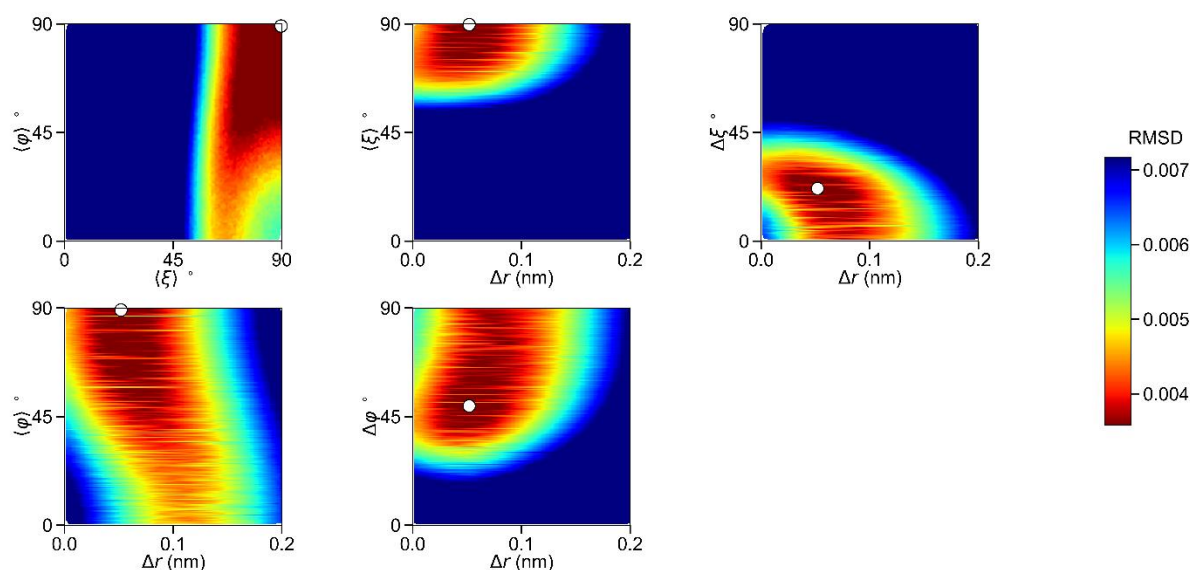


Figure 6.4. The RMSD in dependence of the different pairs of fitting parameters. The dark red regions correspond to the RMSD values, which are below 120 % of the minimal RMSD. The optimized values of the fitting parameters are depicted by white dots.

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 a spin system consisting of a high-spin Fe^{3+} (hs Fe^{3+}) and an 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,^[4] which makes DipFit still applicable. Thus, only this particular case of a large zero-field splitting is considered here. 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 of Chapter 6.1, the inter-spin distance is fixed to 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 set to 0° . In contrast, the angle ζ is expected to have an effect on the dipolar spectrum and 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 . The simulation will be done for two particular cases, $\zeta = 0^\circ$ and $\zeta = 90^\circ$. The idea behind these simulations will become clear later.

The workflow of this fitting includes the following steps:

- 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 the simulation...
```

```
Running the pre-calculations... [DONE]
```

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

```
The simulation is finished
```

```
Saving the simulation results into the directory:
```

```
.../DipFit/examples/example03_hs_iron(III)_nitroxide/2020-01-01_00-00/ [DONE]
```

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

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

```
Starting the simulation...
```

```
Running the pre-calculations... [DONE]
```

```
Calculating the dipolar spectrum vs theta... 100% [DONE]
```

```
The simulation is finished
```

```
Saving the simulation results into the directory:
```

```
.../DipFit/examples/example03_hs_iron(III)_nitroxide/2020-01-01_00-00/ [DONE]
```

- 4) *Explore the simulation results.* During the simulation, a new folder will be created within the directory `/DipFit/examples/example03_hs_iron(III)_nitroxide` and the results of the 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 6.5a). 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 appear on the screen (Figure 6.5b-c).

As can be seen in Figure 6.5a, 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 ζ . 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, $\zeta = 0^\circ$ and $\zeta = 90^\circ$. In the case of $\zeta = 90^\circ$ (Figure 6.5b), 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 6.5c), 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.

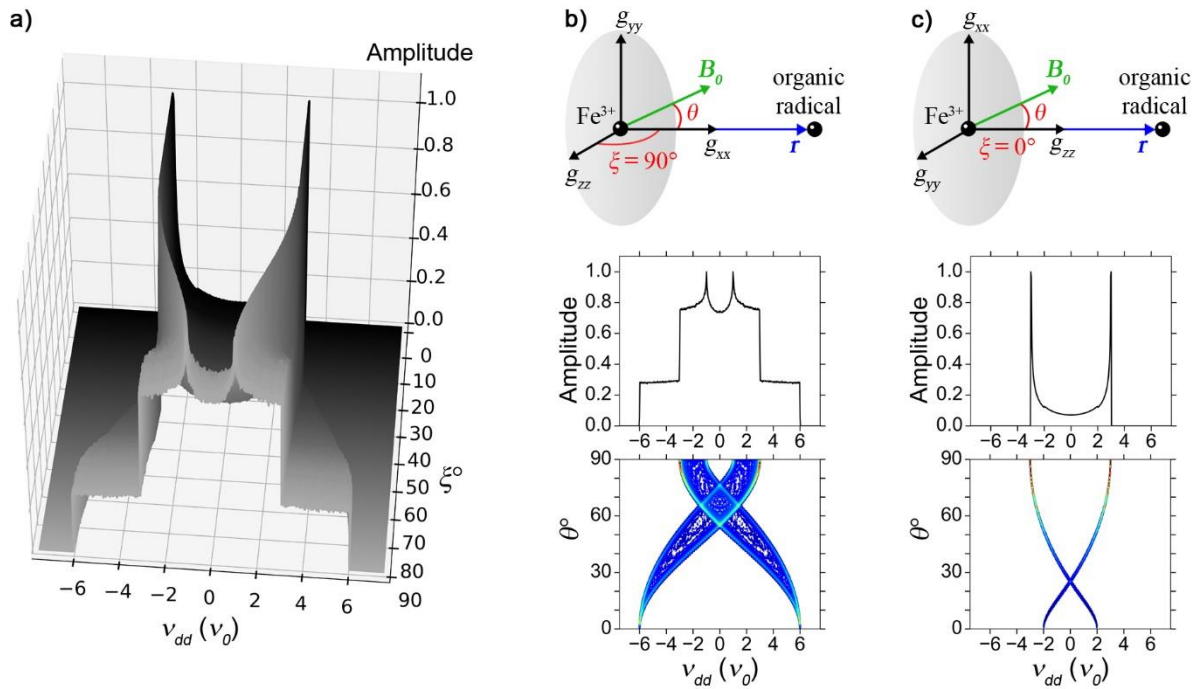


Figure 6.5. a) The angular dependence of the dipolar spectrum, which corresponds to the spin system consisting of an isotropic organic radical with $g_{iso} = g_e$ and an anisotropic hs Fe^{3+} with $g_{aniso} = [6.0, 6.0, 2.0]$. The angle ζ is defined in Figure 1.1. For the particular cases of the b) $\zeta = 0^\circ$ 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 to the fitting of a RIDME spectrum acquired on the MTSL-labeled mutant Q8R1 of the hemeprotein met-myoglobin (Figure 6.6).^[4] 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.^[4] 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 does not depend on 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. ^[4]). The RIDME measurements on Q8R1 apply to this case. 2) The error of the experimentally measured temperature was above the precision required for accurate spectral simulations.

The workflow of this fitting includes the following steps:

- 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 the widths of the distributions $P(r)$ and $P(\zeta)$, as well as the temperature, are selected to be

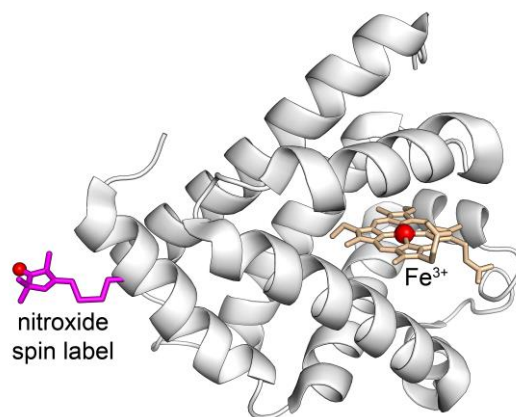


Figure 6.6. The structural model of the met-myoglobin mutant Q8R1. This model is based on the crystal structure of met-myoglobin (PDB-ID 1wla)^[14], which is shown in white. The heme group and the Fe^{3+} ion are depicted as light orange sticks and a red sphere, respectively. A single conformer of the nitroxide spin label, attached at the position Q8, is depicted by magenta sticks. The corresponding oxygen atom is shown as a red sphere.

the fitting parameters (see [fitting_parameters](#)). The distributions $P(r)$ and $P(\zeta)$ are approximated by normal distributions (see [calculation_settings](#)) and, therefore, their widths are given as 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 [error analysis](#), 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 RMSD level, which corresponds to 120 % of the lowest RMSD (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 the fitting...
```

```
Optimization step 500 / 500: RMSD = 0.027134
```

```
The fitting is finished. Total duration: 17:17:51.424179
```

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

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

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

```
Starting the error analysis...
```

```
Numerical error = 0.003214
```

```
Minimal RMSD = 0.026037
```

```
RMSD threshold = 0.031244
```

```
Calculating the RMSD in dependence of fitting parameters ...
```

```
Calculation step 3 / 3
```

```
Calculating the errors of fitting parameters ...
```

```
The error analysis is finished. Total duration: 9:29:20.916722
```

```
Saving the results of the error analysis into the directory:
```


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

Plotting the results of the error analysis... [DONE]

Optimized fitting parameters:

Parameter	Value	Optimized	Precision (+/-)
r mean (nm)	2.753	Y	0.022
r width (nm)	0.070	Y	0.036
xi mean (deg)	80.122	Y	2.162
xi width (deg)	3.129	Y	1.805
phi mean (deg)	0.000	N	nan
phi mean (deg)	0.000	N	nan
temperature (K)	1.969	Y	0.257

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 below the error of the optimized parameters (see below).

- 4) *Explore the fitting results.* During the fitting, a 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).

Next, the results of the fitting will be briefly discussed. The discussion will begin by exploring the content of the files [score.dat](#) and [score.png](#). These files contain the dependence of the RMSD on the optimization step, which is depicted in Figure 6.7a. This figure reveals that the RMSD decreased during first 150 optimization steps and, then, reached a plateau during the next 350 optimization steps. Based on this, one can conclude that the genetic algorithm has converged to the global minimum and, consequently, the best fit to the experimental RIDME spectrum was found. The latter fit is stored in the output files [fit.dat](#) and [fit.png](#) and depicted in Figure 6.7b. As can be seen, the fit reproduces quite accurately the shape of the experimental RIDME spectrum and, therefore, the fitting can be considered as successful. The corresponding optimized values of the fitting parameters are stored in the output file [parameters.dat](#), whose content is listed in Figure 6.7c. In order to estimate the errors of these parameters, the following data was 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 data sets is stored

in the output file [parameter_errors.png](#) and shown here in Figure 6.7d. In this figure, the uncertainty ranges of the fitting parameters are depicted in dark red. Recall that the uncertainty ranges were determined at the RMSD level, which corresponds to 120 % of the minimal RMSD. In addition, the optimized values of the fitting parameters are depicted by circles. As expected, all circles appear within the uncertainty ranges mentioned above, meaning that the optimized values of the fitting parameters lay within the corresponding uncertainty ranges. The errors of individual fitting parameters, which were derived from the uncertainty intervals, are listed in Figure 6.7c. The errors of the mean inter-spin distance $\langle r \rangle$ and its standard deviation Δr are ± 0.02 nm and ± 0.04 nm, respectively. The error of the angular parameters $\langle \zeta \rangle$ and $\Delta \zeta$ is $\pm 2^\circ$. Thus, both distributions, $P(r)$ and $P(\zeta)$, could be extracted from the RIDME spectrum of Q8R1 with a good precision. The optimized value of the temperature equals to 2.0 ± 0.3 K. This value deviates from the experimental temperature (3 K)^[4] by 1.0 ± 0.3 K. The obtained deviation is in agreement with the reported experimental error of the temperature measurement (~ 1 K).^[4]

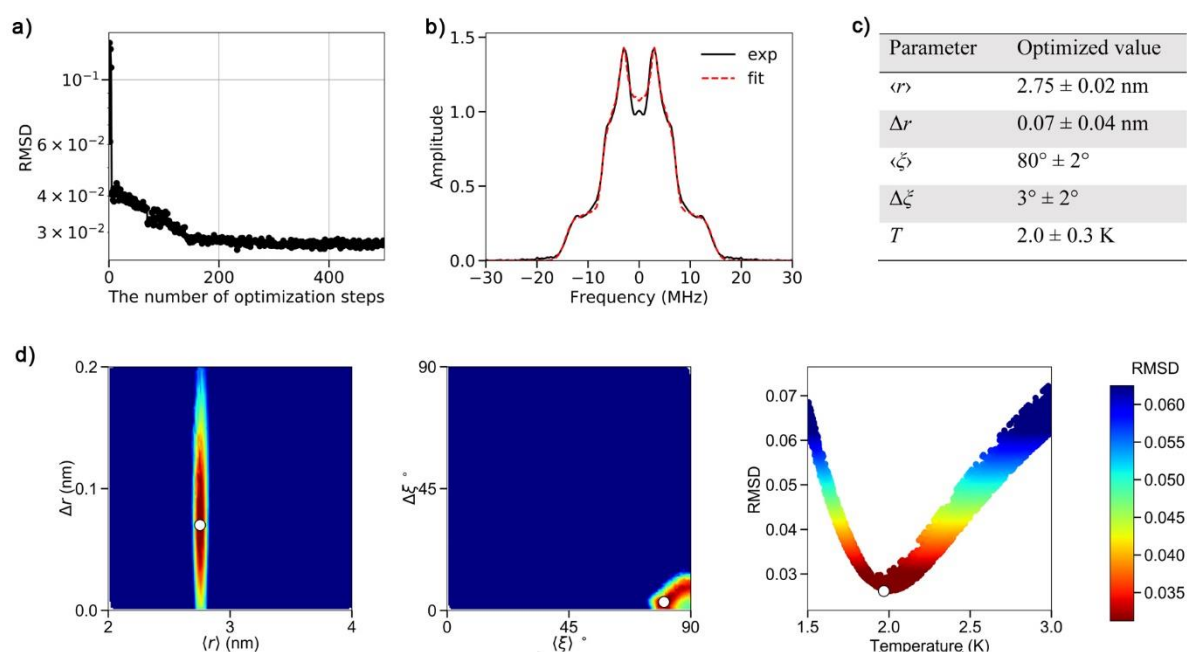


Figure 6.7. Fitting the RIDME spectrum of Q8R1 by DipFit. **a)** The RMSD in dependence of the optimization step. **b)** The RIDME spectrum (black line) is overlaid with the corresponding fit (red line). **c)** The optimized values of the fitting parameters. **d)** The RMSD in dependence of the different fitting parameters. The dark red regions correspond to the RMSD values, which are below 120 % of the minimal RMSD. The optimized values of the fitting parameters are depicted by white dots.

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] A. F. Bedilo, A. G. Maryasov, *J. Magn. Reson.* **1995**, *116*, 87–96.
- [4] 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.
- [5] D. Abdullin, P. Brehm, N. Fleck, S. Spicher, S. Grimme, O. Schiemann, *Chem. – A Eur. J.* **2019**, *25*, 14388–14398.
- [6] B. Filipič, J. Štrancar, *Appl. Soft Comput.* **2001**, *1*, 83–90.
- [7] T. Spalek, P. Pietrzyk, Z. Sojka, *J. Chem. Inf. Model.* **2005**, *45*, 18–29.
- [8] S. Stoll, A. Schweiger, *J. Magn. Reson.* **2006**, *178*, 42–55.
- [9] D. Abdullin, G. Hagelueken, R. I. Hunter, G. M. Smith, O. Schiemann, *Mol. Phys.* **2015**, *113*, 544–560.
- [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.