

Application of spherical harmonics for DEER data analysis in systems with a conformational distribution.

*Alexey Potapov^{*1},*

*¹School of Physics and Astronomy, University of Nottingham, University Park,
Nottingham, NG7 2RD, UK,*

**corresponding author: alexey.potapov@nottingham.ac.uk, tel. +44 115 951 4739*

Keywords:

EPR theory, orientation selection, data processing, distance measurements, high-field EPR

Abstract

Double electron-electron resonance (DEER) and other pulse electron paramagnetic resonance (EPR) techniques are valuable tools for determining distances between paramagnetic centres. DEER theory is well developed for a scenario where relative orientations of paramagnetic centres do not affect the DEER data. In particular, such theory enables a number of approaches for extracting distance distributions. However, in a more general case, when orientation selection effects become substantial, the analytical theory of DEER is less well developed, therefore quite commonly researchers rely on a comparison of some model-based simulations with experimental data. This work elaborates the theory of DEER with orientation selection effects, focusing on a scenario of a moderate conformational disorder, leading to an orientation distribution in a pair of paramagnetic centres. The analytical treatment based on expansions into spherical harmonics, provides important insights into the structure of DEER data. As follows from this treatment, DEER spectra with orientation selection can be represented as a linear combination of modified Pake pattern (MPP) components. The conformational disorder has a filtering effect on the weights of MPP components, specifically by significantly suppressing MPP components of higher degrees. The developed theory provides a pathway for model-based simulations of DEER data where orientation distribution is defined by analytical functions with parameters. The theory based on spherical harmonics expansions was also applied to develop an iterative processing algorithm based on Tikhonov regularization, which disentangles the distance and orientation information in a model-free manner. As an input, this procedure takes several DEER datasets measured at various positions of an EPR line, and outputs a distance distribution and orientation distribution information encoded in a set of coefficients related to the weights of MPP components. The model-based and model-free approaches based on the developed theory were validated for a nitroxide biradical and a spin-labelled protein.

Introduction.

Double electron-electron resonance (DEER) and other pulse electron paramagnetic resonance (EPR) techniques have become widely used for measuring inter-spin distances in a typical range of 2-8 nm[1,2], which can be extended up to 10 nm in deuterated samples[3]. Such experiments are particularly instrumental for structural studies of biomolecules, such as proteins and nucleic acids[4–6]. There, the measurements are carried out to determine distances between pairs of paramagnetic species, such as intrinsic paramagnetic sites in a biomolecule or between chemically linked spin labels. In principle, the data produced by pulse DEER experiments encodes information both on the distance between paramagnetic species and on their orientations with respect to a biomolecule[7]. More generally, due to a conformational flexibility of linkers tethering the spin labels to a biomolecule, and a conformational flexibility of a biomolecule itself, the DEER data collected in such systems encodes information both on the distance and orientation distributions in pairs of spins.

The theory of DEER is rather well developed and enables calculation of DEER traces for a given inter-spin distance and orientation of a spin pair[4,8,9]. A scenario, where distance orientations do not have an effect on the DEER data, is the most commonly encountered in the EPR literature (e.g. for many nitroxide spin-labelled proteins at X-band and Q-band). There, the distance distribution can be obtained by solving an inverse problem. The distance distribution can then be found from an integral equation with a known kernel using Tikhonov regularization, fitting to a particular distance model or by processing using neural networks[10–13]. In general, however, DEER traces actually do depend on the relative orientation distributions of paramagnetic species in a molecule, because microwave (MW) pulses of a DEER sequence excite (or select) only a subset of molecular orientations contributing to an

EPR line. For brevity this phenomenon is further referred to as orientation selection effect in DEER. Extracting distance and orientation information in this case is more complicated.

For nitroxide spin-labels often used in biomolecular studies, the inhomogeneously broadened EPR spectrum has a moderate width (~ 200 MHz and ~ 300 MHz at the base in the magnetic fields of ~ 0.35 T and ~ 1.2 T respectively). Therefore at X-band pulse EPR spectrometers with wide enough resonator bandwidth, the pump pulse may be set to the maximum of the spectrum, while the observer pulses can be applied at almost any position across the EPR line. For such a scenario Marko et al.[14] have previously demonstrated a model-free approach for extracting distance and orientation information. The distance distribution can be obtained then using Tikhonov regularization from a synthetic dataset made by a summation over many DEER traces, which were collected with an observer pulse set to many possible spectral positions of the EPR line. The orientation information encoded in a so called orientation intensity function $\tilde{\lambda}(\cos\theta)$ can also be obtained using Tikhonov regularization from the corresponding integral equation. However, for systems where their EPR spectrum is too wide and the resonator bandwidth is too narrow for setting DEER pulses in this manner (e.g. Cu^{2+} centres at X-band[15] or nitroxides at some W-band spectrometers[16]), such an approach is generally not applicable.

The orientation selection effects can also be studied by direct numerical simulations of DEER traces, which are then compared with experimental results[15–24]. While numerical approaches are useful, using analytical treatment is often a preferred way for obtaining revealing insights. Previously, DEER data with orientation selection were studied theoretically to obtain an angle between rigid labels in a DNA molecule[25]. However, no further developed analytical treatment of DEER with orientation selection has been presented so far.

This work elaborates the theory of DEER experiments with orientation selection. The treatment focuses on a scenario where orientations of paramagnetic species with respect to a molecular frame have a moderate degree of disorder, such as common in spin-labelled biomolecules. The necessary degree of this disorder will later be detailed quantitatively, while at this stage it is defined as neither too large to render DEER data independent of the orientations, nor too small as in a rigid biradical. At the basis of this approach are rotational transformations of excitation probability density functions $f(\theta, \phi)$, defined on a spherical surface and represented as their spherical harmonic series. The developed theory enables both model-based and model-free analysis of DEER data with orientation selection as demonstrated using a nitroxide biradical and a spin-labelled τ_{C14} protein.

The paper is organized as follows. First, it demonstrates that representation with spherical harmonics enables *analytical* calculation of effects produced by orientation distributions, described by parametrized analytical functions. In addition, DEER spectra are shown to consist of a linear combination of components, further referred to as *modified Pake patterns* (MPPs). The weights of MPP components encode the orientation distributions and quite importantly, a moderate conformational disorder suppresses MPP components of higher degrees.

Second, model-based simulations using this theory are shown to provide a very good agreement with experimental data obtained for a model nitroxide biradical. Such simulations use orientation distribution parameters, such that a search in the parameter space is enabled. The theory then is applied to treat the same nitroxide biradical DEER in a model-free manner, using a Tikhonov regularization-based iterative processing algorithm for finding both the distance and orientation information. This approach yields orientation information, which is encoded in the form of coefficients, related to MPP component weights, and a separate distance distribution using as an input several DEER datasets collected at various positions of the EPR line.

Finally, the model-based approach is applied to simulate DEER data of a spin-labelled τ_C14 protein, the conformational space of which is obtained using the MMM software[26]. The simulations and the experiment are shown to agree fairly well with one another. The model-free approach then is applied to obtain the distance distribution and the encoded orientation information for the protein.

Theoretical background

DEER experiment

In a DEER experiment[2] the refocused echo intensity produced by the observer pulses is recorded as a function of the time delay between the pump pulse and a primary echo position, denoted as t . For an electron spin pair where one of the spins is observed, while the other one is inverted by a pump pulse, the DEER sequence produces a signal[8]:

$$s(t) = \cos[\omega_{\text{dd}}(r, \theta)t]. \quad (1)$$

In a point-dipole approximation, the frequency of the dipolar interaction $\omega_{\text{dd}}(r, \theta)$ depends on a distance r between the two electron spins, and a polar angle θ determining the orientation of the magnetic field in the dipolar frame:

$$\omega_{\text{dd}}(r, \theta) = \frac{\mu_0 g_1 g_2 \beta_e^2}{4\pi\hbar} \frac{(1 - 3 \cos^2 \theta)}{r^3} = \omega_{\text{dd},0}(1 - 3 \cos^2 \theta), \quad (2)$$

where g_1 and g_2 are g -factors of the two spins respectively. The z -axis of the dipolar frame here is aligned with a vector connecting the two spins, while the x -axis and y -axis can be chosen in an arbitrary manner.

In a typical EPR experiment, the sample usually consists of many molecules containing pairs of paramagnetic species. The DEER signal $S(t)$, also referred to as a DEER trace, arises due

to an excitation of spin pairs within one molecule (intramolecular contribution), and due to an excitation of spin pairs in different molecules (intermolecular contribution):

$$S(t) = \prod_{j, \text{ all pairs}} s_j(t) = S_{\text{inter}}(t)S_{\text{intra}}(t). \quad (3)$$

The intermolecular contribution denoted as $S_{\text{inter}}(t)$ is referred to as a background signal in EPR literature[8,27]. It often has a known shape and is usually removed during DEER data processing. The intramolecular part $S_{\text{intra}}(t)$ depends on the geometry of spin systems within a molecule and when orientation selection effects are negligible it is also referred to as a form-factor in EPR literature. However, in a more general scenario $S_{\text{intra}}(t)$ encodes both the molecular geometry and the excitation probabilities by various pulses of the DEER sequence. If a sample consists of molecules containing only two spin systems labelled A and B (e.g. a biradical or a pairwise spin-labelled protein) the intramolecular part $S_{\text{intra}}(t)$ in turn can be written as:

$$S_{\text{intra}}(t) = \sum_{i, \text{ molecules}} \left[p_A \left(1 - \lambda_{AB}^{(i)} \cos \omega_{\text{dd}}^{(i)} t \right) + p_B \left(1 - \lambda_{BA}^{(i)} \cos \omega_{\text{dd}}^{(i)} t \right) \right], \quad (4)$$

where p_A and p_B are probabilities for the observer pulses to excite spin systems A and B respectively and summation proceeds over all the molecules in a sample. Symbol $\lambda_{AB}^{(i)}$ denotes a probability that in a particular intramolecular spin pair “ i ” the pump pulse excites the spin system B , given that the spin system A is already excited by an observer pulse, whereas $\lambda_{BA}^{(i)}$ denotes the same thing with the swapped A and B . Finally, $\omega_{\text{dd}}^{(i)}$ is a dipolar interaction in a pair of A and B .

The summation over all intramolecular pairs can be replaced with an integration over all possible orientations of the external magnetic field, which direction in the dipolar frame is given by a polar angle θ and an azimuthal angle ϕ . The probabilities to excite a particular pair

$\lambda_{AB}^{(i)}$ and $\lambda_{BA}^{(i)}$ can then be replaced with continuous functions $\lambda_{AB}(\theta, \phi)$ and $\lambda_{BA}(\theta, \phi)$, thus giving an intramolecular signal:

$$\begin{aligned}
S_{\text{intra}}(t) &= \frac{S(t)}{S_{\text{inter}}(t)} = \\
&= 1 - \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \int_0^\infty \lambda(\theta, \phi) [1 - \cos \omega_{\text{dd}}(r, \theta)t] f(r) dr, \quad (5) \\
\lambda(\theta, \phi) &= p_A \lambda_{AB}(\theta, \phi) + p_B \lambda_{BA}(\theta, \phi),
\end{aligned}$$

where $f(r)$ is a distance distribution. The distance and orientation distributions here are assumed to be independent from one another, which means that a distribution $f(r)$ has no dependence on angles θ and ϕ . While strictly speaking unphysical in many cases, this assumption helps to provide useful insights as will be shown further. Functions $\lambda_{AB}(\theta, \phi)$ and $\lambda_{BA}(\theta, \phi)$ in Eq.(5) have a physical meaning of a probability density that a pair, where magnetic field is directed at (θ, ϕ) is excited by both the pump and the observer pulses. Thus for brevity $\lambda(\theta, \phi)$ is referred to as a pair excitation probability density function (PDF). When orientation selection effects are significant, this function $\lambda(\theta, \phi) \neq \text{const}$ and as follows from the definition, its shape depends on the probabilities p_A and p_B , i.e. on the positioning of the MW pulses within an EPR spectrum.

A model of a fictitious biradical consisting of two nitroxide radicals tethered via a flexible linker, shown in Fig. 1, is used to visualize and highlight the meaning of the PDF $\lambda(\theta, \phi)$. The orientations of the two nitroxide spin systems labelled *A* and *B* are somewhat disordered as schematically shown in Fig. 1a. The specific disorder model in this example will be explained further in this paragraph. Generally each conformer “*i*” in the resulting conformational ensemble can be characterized by their Euler angles $(\alpha_i^{(A)}, \beta_i^{(A)}, \gamma_i^{(A)})$ and $(\alpha_i^{(B)}, \beta_i^{(B)}, \gamma_i^{(B)})$ transforming the principal axes frame of the *g*-anisotropy tensor of *A* and *B* (called *g*-frame for

brevity) into the principal axes frame of the dipolar interaction (called dipolar frame for brevity). The EPR spectra of *A* and *B* nitroxide spin systems at W-band are dominated by the *g*-anisotropy and the nitrogen nucleus hyperfine coupling. The EPR spectrum shown in Fig. 1b is simulated with the EasySpin software[28] using the *g*-anisotropy and hyperfine couplings reported earlier by Savitsky et al.[29] The observer and pump pulses, which frequencies are set as shown by arrows in Fig. 1b, excite (or *select*) a subset of molecular orientations of spin systems *A* and *B* respectively, i.e. such system features *orientation selection* by the MW pulses. The probability of this excitation can be given by the excitation probability density functions (PDF) $f_{\text{pump}}^{(A,g)}(\theta, \phi)$ and $f_{\text{obs}}^{(B,g)}(\theta, \phi)$, which depend on the angles θ and ϕ determining the direction of the external magnetic field in the *g*-anisotropy tensor frames of the two spin systems. The values of both these PDFs are plotted as various colours in a surface plot shown in Fig. 1c. For illustration purposes it is convenient to split the coordinate transformations from *A* and *B* *g*-frames to the dipolar frame in two parts. The first part, carries out some *average* transformation by the three Euler angles $(\bar{\alpha}^{(A)}, \bar{\beta}^{(A)}, \bar{\gamma}^{(A)})$ for the spin system *A* and $(\bar{\alpha}^{(B)}, \bar{\beta}^{(B)}, \bar{\gamma}^{(B)})$ for *B* respectively. The second part contains the remaining *random* rotation with the angles $(\Delta\alpha_i^{(A)}, \Delta\beta_i^{(A)}, \Delta\gamma_i^{(A)})$ and $(\Delta\alpha_i^{(B)}, \Delta\beta_i^{(B)}, \Delta\gamma_i^{(B)})$. In this example, they are distributed according to Gaussian distributions (*vide infra*) centred at zero with the widths chosen as $(\sigma_\alpha^{(A/B)}, \sigma_\beta^{(A/B)}, \sigma_\gamma^{(A/B)}) = (30^\circ, 30^\circ, 30^\circ)$ for both *A* and *B*. PDFs $f_{\text{obs}}^{(A,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B,g)}(\theta, \phi)$ after the first average transformation are shown in Fig. 1c(ii). Then the second rotation is applied and the result is summed up over all the conformers “*i*”, giving PDFs $f_{\text{obs}}^{(A,\text{dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(B,\text{dip})}(\theta, \phi)$ in the dipolar frame, shown in Fig. 1c(iii). While the PDFs $f_{\text{obs}}^{(A,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B,g)}(\theta, \phi)$ have some noticeable sharp features, after summing up over the conformers the resulting PDFs $f_{\text{obs}}^{(A,\text{dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(B,\text{dip})}(\theta, \phi)$ are significantly smoothed. In

the following sections of this paper, it will be shown that such smoothing effect is a rather general property and its mathematical description will be presented. Finally, assuming that orientations of spin systems A and B are independent from one another, the probability to excite both spins in a pair for which the magnetic field is directed at θ and ϕ is given by:

$$\lambda_{AB}(\theta, \phi) = f_{\text{obs}}^{(A,\text{dip})}(\theta, \phi) f_{\text{pump}}^{(B,\text{dip})}(\theta, \phi). \quad (6)$$

Since the EPR spectra of the nitroxide spin systems A and B are the same, the observer and pump pulses also excite spins B and A respectively with the same probabilities and corresponding excitation PDFs $f_{\text{pump}}^{(A,g)}(\theta, \phi)$ and $f_{\text{obs}}^{(B,g)}(\theta, \phi)$. The pair excitation PDF is then given by:

$$\begin{aligned} \lambda(\theta, \phi) &= \frac{1}{2} (\lambda_{AB}(\theta, \phi) + \lambda_{BA}(\theta, \phi)) \\ &= \frac{1}{2} \left(f_{\text{obs}}^{(A,\text{dip})}(\theta, \phi) f_{\text{pump}}^{(B,\text{dip})}(\theta, \phi) + f_{\text{pump}}^{(A,\text{dip})}(\theta, \phi) f_{\text{obs}}^{(B,\text{dip})}(\theta, \phi) \right), \end{aligned} \quad (7)$$

which is plotted in Fig. 1c(iv).

As follows from Eq.(5), a DEER signal at very long times $t \rightarrow \infty$, consists of a large number of components oscillating with t , which cancel each other out upon the integration. For that reason, a DEER signal levels out for large t , i.e. $S_{\text{intra}}(\infty) = S(\infty)/S_{\text{inter}}(\infty) = (1 - \lambda)$, where λ is a total probability to excite any spin by a pump pulse, among pairs in which the first spin is already excited by observer pulses. It is related to the pair excitation PDF $\lambda(\theta, \phi)$ as:

$$\lambda = \int_0^{2\pi} d\phi \int_0^\pi \lambda(\theta, \phi) \sin \theta d\theta. \quad (8)$$

Experimentally, the value of λ can be determined as a modulation depth of a DEER trace, when $t \rightarrow \infty$.

Eqs. (5) and (6) show that calculations of DEER traces require knowledge of the pair excitation PDF $\lambda(\theta, \phi)$ which in turn depends on the dipolar frame excitation PDFs $f_{\text{pump}}^{(\text{A,dip})}(\theta, \phi)$ and $f_{\text{obs}}^{(\text{B,dip})}(\theta, \phi)$. Those PDFs emerge as a result of rotations of coordinate systems between the g -frame and the dipolar frame, transforming the PDFs $f_{\text{pump}}^{(\text{A,g})}(\theta, \phi)$ and $f_{\text{obs}}^{(\text{B,g})}(\theta, \phi)$. The next section focuses on a technique for carrying out such rotations.

Excitation of molecular orientations by a microwave pulse

The shape of an EPR spectrum in a frozen solution is mostly determined by the g - and hyperfine-tensor anisotropies. The strength of an oscillating magnetic field attainable in a typical pulse EPR spectrometer is often sufficient to excite only a fraction of the spectrum, i.e. microwave (MW) pulse there excite (or *select*) only a subset of molecular orientations.

Let's consider the effect of this orientation selection in the principal axis frame of the g -anisotropy tensor (i.e. g -frame as defined above) of one of the spin systems. For all the spin systems in an ensemble, in their g -frames the probability for an external magnetic field to point in the direction (θ, ϕ) is described by a uniform probability density function (PDF), i.e. $f(\theta, \phi) = \text{const}$. However, among the spin systems excited by a MW pulse, the probability for a certain direction of the magnetic field is no longer uniform, i.e. $f(\theta, \phi) \neq \text{const}$. In practice, this excitation PDF $f(\theta, \phi)$ can be found by solving numerically the system spin Hamiltonian, in which the spin system parameters, such as a g -tensor and hyperfine coupling (and excluding the dipolar coupling term) should be obtained elsewhere. In particular, for DEER measurements between pairs of nitroxide spin-labels, the spin Hamiltonian parameters are usually well known[30].

The same excitation PDF can be also be expressed in another frame, which arises as a result of a coordinate system rotation. Using an operator $\hat{D}(\alpha, \beta, \gamma)$ which rotates the coordinate

system by the three Euler angles (α, β, γ) , between the old a new coordinate systems (θ, ϕ) and (θ', ϕ') , in a new frame the result of rotation can be written as[31]:

$$f'(\theta, \phi) = \widehat{D}(\alpha, \beta, \gamma)f(\theta, \phi) = f(\theta', \phi') = f(\theta'(\theta, \phi, \alpha, \beta, \gamma), \phi'(\theta, \phi, \alpha, \beta, \gamma)). \quad (9)$$

Carrying out this rotation is easier by using an expansion of $f(\theta, \phi)$ into a series of complex spherical harmonics $Y_{lm}(\theta, \phi)$. Any function $f(\theta, \phi)$ can then be represented as:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_{lm}(\theta, \phi), \quad (10)$$

where f_{lm} are complex spherical harmonic coefficients (SHC). Generally, a function $f(\theta, \phi)$ can be said to be uniquely represented by a set of its SHCs f_{lm} . The Supplementary Information section ‘‘Spherical harmonics basics’’ outlines some basic properties of spherical harmonics and provides expressions for obtaining f_{lm} .

An individual spherical harmonic in the rotated frame transforms into:

$$\widehat{D}(\alpha, \beta, \gamma)Y_{lm}(\theta, \phi) = \sum_{m'=-l}^l D_{m',m}^{(l)}(\alpha, \beta, \gamma)Y_{lm'}(\theta, \phi), \quad (11)$$

where $D_{m',m}^{(l)}(\alpha, \beta, \gamma)$ is an element of the Wigner D -matrix. It is important to note, that coordinate system rotation produces a linear combination of spherical harmonics with the same value of l , i.e. spherical harmonics of degree l form a subspace upon rotations. The matrix elements for a rotation matrix in this subspace $D_{m',m}^{(l)}(\alpha, \beta, \gamma)$ are given by:

$$D_{m',m}^{(l)}(\alpha, \beta, \gamma) = e^{im'\alpha} d_{m',m}^{(l)}(\beta) e^{im\gamma}, \quad (12)$$

where $d_{m',m}^{(l)}(\beta)$ is an element of the Wigner (small) d -matrix which analytical form is known[32].

Combining Eq.(10) and (11) allows expressing a function $f(\theta, \phi)$ in a new frame:

$$\begin{aligned}
f'(\theta, \phi) &= \widehat{D}(\alpha, \beta, \gamma)f(\theta, \phi) = \\
&= \sum_{l=0}^{\infty} \sum_{m=-l}^l \sum_{m'=-l}^l f_{lm} Y_{lm'}(\theta, \phi) D_{m',m}^{(l)}(\alpha, \beta, \gamma) = \\
&= \sum_{l=0}^{\infty} \sum_{m'=-l}^l Y_{lm'}(\theta, \phi) \sum_{m=-l}^l D_{m',m}^{(l)}(\alpha, \beta, \gamma) f_{lm}.
\end{aligned} \tag{13}$$

After swapping the indices m and m' , $f'(\theta, \phi)$ can be rewritten as a linear combination of spherical harmonics:

$$f'(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f'_{lm} Y_{lm}(\theta, \phi), \tag{14}$$

which SHCs f'_{lm} can be obtained using:

$$f'_{lm} = \sum_{m'=-l}^l D_{m,m'}^{(l)}(\alpha, \beta, \gamma) f_{lm'}. \tag{15}$$

To summarize, the SHCs f'_{lm} of a function $f'(\theta, \phi)$ transformed into a rotated frame from a function $f(\theta, \phi)$, can be obtained by multiplying the Wigner D -matrix by a vector consisting of the SHCs f_{lm} . Such transformation scrambles the f_{lm} coefficients within a subspace with the same value of l to produce SHCs f'_{lm} .

Averaging due to distributed rotations

In most EPR measurements, the sample consists of many molecules having some degree of conformational disorder. It means that the relative orientation between any two frames in a molecule (for instance, between a g -frame and a dipolar frame) varies from one molecule to another due to differences of the molecular conformations. In other words, a rotation operator $\widehat{D}(\alpha_i, \beta_i, \gamma_i)$ for each molecule " i " in an ensemble is characterized by its own Euler angles $(\alpha_i, \beta_i, \gamma_i)$. The excitation PDF $f'(\theta, \phi)$ in the rotated frame can be calculated from the initial PDF $f(\theta, \phi)$ by summing over all molecules of the conformational ensemble:

$$f'(\theta, \phi) = \sum_{i, \text{ molecules}} f'_i(\theta, \phi) = \sum_{i, \text{ molecules}} \widehat{D}(\alpha_i, \beta_i, \gamma_i) f(\theta, \phi). \quad (16)$$

The SHCs $f'_{lm,i}$ for each molecule "i" can be found by applying Eq.(15):

$$f'_{lm,i} = \sum_{m'=-l}^l D_{m,m'}^{(l)}(\alpha_i, \beta_i, \gamma_i) f_{lm}. \quad (17)$$

The PDF $f'(\theta, \phi)$ can also be expanded into a series of spherical harmonics:

$$f'(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f'_{lm} Y_{lm}(\theta, \phi). \quad (18)$$

Using brackets $\langle \rangle$ as a shorthand for averaging, SHCs f'_{lm} can be compactly written as:

$$\begin{aligned} f'_{lm} &= \sum_i \sum_{m'=-l}^l D_{m,m'}^{(l)}(\alpha_i, \beta_i, \gamma_i) f_{lm'} = \sum_{m'=-l}^l \left(\sum_i D_{m,m'}^{(l)}(\alpha_i, \beta_i, \gamma_i) \right) f_{lm'} \\ &= \sum_{m'=-l}^l \langle D_{m,m'}^{(l)}(\alpha, \beta, \gamma) \rangle f_{lm'}. \end{aligned} \quad (19)$$

The effect of averaging due to random rotations can be illustrated using as an example the fictitious biradical in Fig. 1a mentioned above. The excitation PDFs for the pump and observer pulses $f_{\text{obs}}^{(A,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B,g)}(\theta, \phi)$, plotted in Fig. 1c(i), can be expanded as a series of spherical harmonics. The power carried by a set of spherical harmonics with the same degree l can be calculated from their SHCs as $P_l = \sum_{m=-l}^l |f_{lm}|^2$, as explained in the ‘‘Spherical harmonics basics’’ section of the SI. This quantity serves as a measure of l -th degree harmonics contribution to the total function. Fig. 1d shows power spectra, where the power of the l -th degree harmonics is plotted as a function of l for the PDFs $f_{\text{obs}}^{(A,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B,g)}(\theta, \phi)$ in the fictitious biradical. As seen from this graph, most of the contribution to $f_{\text{obs}}^{(A,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B,g)}(\theta, \phi)$ arises from the harmonics with degrees $l < 10$. Conformational disorder in the

biradical produces dipolar frame PDFs $f_{\text{obs}}^{(\text{A,dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(\text{B,dip})}(\theta, \phi)$, which are shown in Fig. 1c(iii). For these PDFs, their power spectra, shown in Fig. 1e, demonstrate a rather significant suppression of the harmonics with higher degrees of l , more specifically, most of the intensity arises only in the degrees up to $l \leq 4$. As will be shown later this suppression of higher degree harmonics, leading to a significant “smoothing” of PDFs, is a rather common property arising due to the conformational disorder. In addition, the PDFs $f_{\text{obs}}^{(\text{A,dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(\text{B,dip})}(\theta, \phi)$ (together with their pairs $f_{\text{obs}}^{(\text{B,dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(\text{A,dip})}(\theta, \phi)$) are used to calculate $\lambda(\theta, \phi)$, which is shown in Fig. 1c(iv). Its power spectrum shown in Fig. 1e, demonstrates that most intensity in $\lambda(\theta, \phi)$ also arises from harmonics with $l \leq 4$. Note that any excitation PDF must be even with respect to a sign change of the external magnetic field, due to a symmetry of the spin Hamiltonian. This means that their spherical harmonics expansions can only contain functions $Y_{lm}(\theta, \phi)$ with even l due to symmetry properties of the latter (shown in “Spherical harmonics basics” of the SI). Rotations retain this symmetry because D -matrix can only produce linear transformations in a subspace of spherical functions $Y_{lm}(\theta, \phi)$ with the same value of l .

As evident from Eq. (12), the Wigner D -matrix elements can be expressed in a form of analytical functions. Therefore for conformational distributions known in an analytical form, the rotated and averaged SHCs f'_{lm} can also be calculated analytically using Eq. (19). The averaging can be done by replacing the summation in Eq. (19) with an integration over all possible rotations:

$$\langle D_{m,m'}^{(l)}(\alpha, \beta, \gamma) \rangle = \int_0^{2\pi} d\gamma \int_0^\pi d\beta \int_0^{2\pi} D_{m,m'}^{(l)}(\alpha, \beta, \gamma) g(\alpha, \beta, \gamma) d\alpha \quad (20)$$

where $g(\alpha, \beta, \gamma)$ is a probability to find a molecule where the transformation between its coordinate systems (such as between a g -frame and a dipolar frame) is given by the three Euler

angles α, β, γ . In the following, three common scenarios of random rotations will be considered in more detail.

Uniformly distributed random rotation is a trivial case. In this scenario the two frames (such as a g -frame and a dipolar frame) have a random relative orientation with respect to one another. This may be the case for spin-labels tethered to a biomolecule via sufficiently long and flexible linkers. Taking into account all possible orientations is done by using the weighting function:

$$g(\alpha, \beta, \gamma) = \frac{1}{8\pi^2} \sin(\beta) \quad (21)$$

Averaging the D -matrix elements over angles α and γ produces zeros for all elements $\langle D_{m,m'}^{(l)} \rangle$ except those with $m = m' = 0$. Since the Wigner d -matrix element $d_{00}^{(l)}(\beta) \sim Y_{l0}(\beta, 0)$, the integration over β also produces zeros for all l values except $l = 0$ [32]. Therefore, as follows from Eq.(19) and (20), the only non-zero SHC remaining after the integration is $f'_{00} = f_{00}$. Physically, it means that as a result of such random rotations any PDF $f(\theta, \phi)$ in the initial frame (e.g. g -frame) is equivalent to a uniform PDF in the rotated frame (e.g. dipolar frame), i.e. $f'(\theta, \phi) = f_{00} Y_{00}(\theta, \phi) = \text{const}$. The PDF $f'(\theta, \phi)$ can then be regarded as a result of applying a *low-pass filter* to the PDF $f(\theta, \phi)$, such that only a single spherical harmonic component with $l = m = 0$ is retained.

Uniformly distributed rotation about one axis. The Wigner rotation D -matrix dependent on the three Euler angles (α, β, γ) can be represented as three consecutive rotations about axis z, y' and z'' :

$$\widehat{D}(\alpha, \beta, \gamma) = \widehat{D}(0, 0, \gamma) \widehat{D}(0, \beta, 0) \widehat{D}(\alpha, 0, 0), \quad (22)$$

thereby making it easier to consider the effect of individual random rotations about the corresponding axis.

1. First, consider only the rotation $\widehat{D}(\alpha, 0, 0)$, where an angle α is uniformly distributed in a range $0 < \alpha < 2\pi$, i.e. $g(\alpha) = 1/2\pi$. Substituting Eq. (12) into Eq. (20) produces non-zero $\langle D_{m,m'}^{(l)}(\alpha, 0, 0) \rangle$ only for $m = 0$. Furthermore, since $d_{m,m'}^{(l)}(0) = \delta_{m,m'}$, where $\delta_{m,m'}$ is a Kronecker symbol, the averaged values become $\langle D_{m,m'}^{(l)}(\alpha, 0, 0) \rangle = \delta_{0,m} \delta_{0,m'}$. As follows from Eq. (19), the only remaining non-zero SHCs in this case are $f'_{l0} = f_{l0}$. In other words, the averaged PDF $f'(\theta, \phi)$ in Eq. (18) can be regarded as a result of filtering out all spherical harmonics with $m \neq 0$ in the initial PDF $f(\theta, \phi)$.

2. Averaging the rotation $\widehat{D}(0, 0, \gamma)$ over an angle γ uniformly distributed in a range $0 < \gamma < 2\pi$ can be done in a similar manner with the same result, i.e. such averaging suppresses all the spherical harmonics with $m \neq 0$.

3. Averaging the rotation $\widehat{D}(0, \beta, 0)$ over a uniformly distributed angle β can be qualitatively evaluated by noticing that rotation about y-axis is the same as rotation about z-axis in some other frame, specifically:

$$\widehat{D}(0, \beta, 0) = \widehat{D}(-\pi/2, 0, 0) \widehat{D}(0, -\pi/2, 0) \widehat{D}(\beta, 0, 0) \widehat{D}(0, \pi/2, 0) \widehat{D}(\pi/2, 0, 0). \quad (23)$$

Averaging the above expression over angle β has already been done in item 1, giving $\langle D_{m,m'}^{(l)}(\beta, 0, 0) \rangle = \delta_{0,m} \delta_{0,m'}$. This means that $l \times l$ subspace matrix $\langle D_{m,m'}^{(l)}(\beta, 0, 0) \rangle$ has only a single non-zero element $\langle D_{0,0}^{(l)}(\beta, 0, 0) \rangle$. The transformation under $\widehat{D}(0, \pi/2, 0) \widehat{D}(\pi/2, 0, 0)$ then scrambles the elements with the same l value, meaning that the value of $\langle D_{0,0}^{(l)}(\beta, 0, 0) \rangle$ is “spread” over the entire $l \times l$ matrix. As a result, when $\langle D_{m,m'}^{(l)}(0, \beta, 0) \rangle$ is used to calculate SHCs, their values will effectively be reduced by a factor of $\sim 1/l^2$, i.e. $f'_{lm} \sim f_{lm}/l^2$. In other words, the spherical harmonics with larger degrees l are suppressed by such averaging. The exact

expressions for f'_{lm} are rather cumbersome, however Eq. (23) provides a straightforward way of implementing them in a numerical computation using a precalculated matrix $d_{m,m'}^{(l)}(\pi/2)$ [31].

Narrow distribution of Euler angles. In this scenario Euler angles α, β, γ transforming between some frames are distributed according to a zero-centred Gaussian distribution following:

$$g(\alpha, \beta, \gamma) = \frac{1}{N_g} \exp \left[- \left(\frac{\alpha^2}{\sigma_\alpha^2} + \frac{\beta^2}{\sigma_\beta^2} + \frac{\gamma^2}{\sigma_\gamma^2} \right) \right], \quad (24)$$

where N_g is a normalization factor, such that $\int g(\alpha, \beta, \gamma) d\alpha d\beta d\gamma = 1$. Parameters $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$ characterize the width of the angle distribution. For the case of a sufficiently narrow distribution with parameters $\sigma_\alpha, \sigma_\beta, \sigma_\gamma \ll 2\pi$, the average values of $\langle D_{m,m'}^{(l)}(\alpha, \beta, \gamma) \rangle$ can be obtained analytically by substituting Eq.(12) and Eq. (24) into Eq.(20). The exact analytical expressions derived in the ‘‘Notes on averaging due to random rotations’’ section of the SI, show that the averaged values of Wigner D -matrix elements $\langle D_{m',m}^{(l)}(0, \beta, 0) \rangle$ rapidly decay with an increase of m, m', l . This means that SHCs f'_{lm} rapidly decay with an increase of l and m .

As can be seen in all the described scenarios, a transformation of a PDF $f(\theta, \phi)$ into another frame via random rotations is equivalent to applying a *low-pass filter* suppressing spherical harmonics with large values of l and m . This has already been visualized for the case of a fictitious biradical Fig. 1, where the conformational disorder smooths the PDFs in Fig. 1c(i) to produce those in Fig. 1c(iii). In the language of SHCs, it means that the spherical harmonic components of $f_{\text{obs}}^{(A,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B,g)}(\theta, \phi)$ which power spectra are shown in Fig. 1d, are being suppressed for larger values of l to produce PDFs $f_{\text{obs}}^{(A,\text{dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(B,\text{dip})}(\theta, \phi)$ with power spectra shown in Fig. 1e.

Modulation depth and DEER spectra

Once the excitation PDFs $f_{\text{obs}}^{(A/B,\text{dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(B/A,\text{dip})}(\theta, \phi)$ are known, they can be used to find $\lambda(\theta, \phi)$, which expansion into spherical harmonics is given by:

$$\lambda(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \lambda_{lm} Y_{lm}(\theta, \phi). \quad (25)$$

Therefore, as follows from Eq. (8), the modulation depth is simply:

$$\lambda = \lambda_{00}. \quad (26)$$

Calculation of DEER signals as a function of time delay t can be done using Eq. (5). This expression contains terms oscillating with frequencies $\omega_{\text{dd}}(r, \theta)$ weighted by a function $\lambda(\theta, \phi)$. However, it is more convenient to carry out the calculation in the frequency domain, i.e. for DEER spectra.

For brevity of mathematical expressions, consider the case of a definite distance r between the two spins, i.e. when the distance distribution is very narrow. The spectrum of a DEER trace $S_{\text{intra}}(t)$, further referred to as a DEER spectrum, can be obtained by taking a Fourier transform of the time dependent part in Eq.(5):

$$\tilde{S}(\omega) = \int_0^{2\pi} \int_0^{\pi} \frac{1}{2} [\delta(\omega - \omega_{\text{dd}}(\theta, \phi)) + \delta(\omega + \omega_{\text{dd}}(\theta, \phi))] \lambda(\theta, \phi) \sin \theta d\theta d\phi, \quad (27)$$

where $\delta(\omega)$ is a Dirac δ -function. Using the spherical harmonic expansion of $\lambda(\theta, \phi)$, the DEER spectrum $\tilde{S}(\omega)$ can be rewritten as a linear combination of individual spectral components:

$$\begin{aligned} \tilde{S}(\omega) &= \sum_{l,m} \lambda_{lm} \int_0^{2\pi} \int_0^\pi \frac{1}{2} [\delta(\omega - \omega_{dd}(\theta, \phi)) \\ &+ \delta(\omega + \omega_{dd}(\theta, \phi))] Y_{lm}(\theta, \phi) \sin \theta d\theta d\phi = \sum_{l,m} \lambda_{lm} \tilde{S}_{lm}(\omega). \end{aligned} \quad (28)$$

Due to the symmetry of $\omega_{dd}(r, \theta)$ only some of the $\tilde{S}_{lm}(\omega)$ components have non-zero contribution. First, since $\omega_{dd}(r, \theta)$ is independent of azimuthal angle ϕ , the integration over it in Eq. (28) removes all the components with $m \neq 0$. In addition, since $\omega_{dd}(\theta) = \omega_{dd}(-\theta)$, the integration over θ retains only components with even $l = 2k$, where $k = 0, 1, 2 \dots$ etc. As shown in the ‘‘DEER spectra’’ section of the SI, the spectral components $\tilde{S}_{2k,0}(\omega)$ can be expressed as:

$$\tilde{S}_{2k,0}(\omega) = \begin{cases} \pi\sqrt{4k+1} \sqrt{\frac{3}{1-\frac{\omega}{\omega_{dd,0}}}} P_{2k}^0 \left(\sqrt{\frac{1-\frac{\omega}{\omega_{dd,0}}}{3}} \right), & \text{for } -2 \leq \frac{\omega}{\omega_{dd,0}} < -1 \\ \pi\sqrt{4k+1} \left[\sqrt{\frac{3}{1+\frac{\omega}{\omega_{dd,0}}}} P_{2k}^0 \left(\sqrt{\frac{1+\frac{\omega}{\omega_{dd,0}}}{3}} \right) + \sqrt{\frac{3}{1-\frac{\omega}{\omega_{dd,0}}}} P_{2k}^0 \left(\sqrt{\frac{1-\frac{\omega}{\omega_{dd,0}}}{3}} \right) \right], & \text{for } -1 \leq \frac{\omega}{\omega_{dd,0}} \leq 1 \\ \pi\sqrt{4k+1} \sqrt{\frac{3}{1+\frac{\omega}{\omega_{dd,0}}}} P_{2k}^0 \left(\sqrt{\frac{1+\frac{\omega}{\omega_{dd,0}}}{3}} \right), & \text{for } 1 < \frac{\omega}{\omega_{dd,0}} \leq 2. \end{cases} \quad (29)$$

where $P_l^m(z)$ denotes an associated Legendre polynomial with a degree l and an order m . Fig. 2a shows plots of $\tilde{S}_{2k,0}(\omega)$ for several values of $k = 0, 1, \dots, 5$. For $k = 0$ the component $\tilde{S}_{0,0}(\omega)$ has a shape of a regular Pake pattern. Components with $k > 0$ represent the same Pake patterns modulated by the corresponding polynomial functions, and for brevity such components are further referred to as *modified Pake patterns* (MPPs). The DEER traces corresponding to each of these DEER spectra are shown in Fig. 2b.

Any DEER spectrum can therefore be represented using a linear combination of modified Pake pattern components:

$$\tilde{S}(\omega) = \sum_{k=0}^{\infty} w_k^{\text{MPP}} S_k^{\text{MPP}}(\omega), \quad (30)$$

where

$$w_k^{\text{MPP}} = \lambda_{2k,0} \quad (31)$$

is a weight of the k -th degree MPP component $\tilde{S}_k^{\text{MPP}}(\omega) = \tilde{S}_{2k,0}(\omega)$. Obviously, since $w_k^{\text{MPP}} = \lambda_{00}$ is a total probability to excite any pair, it is always positive, while the signs of other weights may have either positive or negative signs.

To summarize, if the distance and orientations distributions are independent from one another and orientations are moderately distributed, the DEER spectra have three important properties:

1. MPP components $\tilde{S}_k^{\text{MPP}}(\omega)$ themselves depend only on the dipolar interaction, and consequently on the underlying distance distribution, whereas orientation information is encoded only in the weights w_k^{MPP} . In other words, distance and orientation information can be mathematically disentangled.
2. As shown above, a conformational disorder significantly suppresses spherical harmonics with large degrees l in the PDFs for the pump and observer pulses. Since $\lambda(\theta, \phi)$ is a result of their product, its SHCs $\lambda_{2k,0} = w_k^{\text{MPP}}$ also rapidly decay with an increase of k .
3. The presence of a distance distribution broadens the MPP spectra and smooths their discontinuities. As an example Fig. 2c shows the MPP components, where the distance is distributed according to a Gaussian distribution. The components have several common features: a) sharp horns at $\omega/\omega_{0,\text{dd}} = \pm 1$, b) a broad intense feature filling the interval $|\omega/\omega_{0,\text{dd}}| \leq 1$ and c) small intensity shoulders spanning $1 < |\omega/\omega_{0,\text{dd}}| < 2$. Individual MPP components are not orthogonal to one another. For that reason, a linear

combination of the first three components with $k = 0, 1, 2$ can be used to approximate MPP components with higher degrees $k \geq 3$.

Overall, these properties form a basis for a procedure for disentangling the distance and orientation information from the DEER data, which will be outlined further in the text.

Methods

EPR. The sample of a nitroxide biradical dissolved in o-terphenyl matrix is a courtesy of Prof. Gunnar Jeschke[33]. The DEER datasets for this biradical were collected using a W-band pulse EPR spectrometer described elsewhere[34]. There, the $t_{90}/t_{180} = 100/200$ ns observer pulses of a DEER experiment sequence were used to produce a refocused echo, while the 180° pump pulse frequency is offset by +20 MHz and its duration is set to $t_{\text{pump}} = 200$ ns. The τ_{C14} protein was spin-labelled with 1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl methanethiosulfonate (MTSL) at Cys56 and via site-directed mutagenesis at Ser31. The DEER measurements with τ_{C14} , which results were reported earlier in ref.[35], were obtained using a W-band spectrometer at the Weizmann Institute of Science (Rehovot, Israel). The $t_{90}/t_{180} = 30/60$ ns pulses were used to produce the refocused echo in the DEER experiment sequence. The pump pulse frequency is offset by +65 MHz and it has a duration of $t_{\text{pump}} = 25$ ns. These experimental parameters were used for simulations of excitation PDFs $f_{\text{obs}}^{(A/B,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B/A,g)}(\theta, \phi)$.

Simulations. All simulation programs were written in Python programming language. The simulations of EPR spectra were verified by comparing with results obtained using the EasySpin software[28]. The PDFs $f_{\text{pump}}^{(A/B,g)}(\theta, \phi)$ and $f_{\text{obs}}^{(B/A,g)}(\theta, \phi)$ at various positions within an EPR spectrum were precalculated for further use. These PDFs were calculated on a uniform

grid of 400×400 points in θ and ϕ , $0^\circ < \theta < 180^\circ$ and $-180^\circ < \phi < 180^\circ$, which is necessary for a proper functioning of the spherical harmonics expansion algorithm[36].

The spherical harmonics expansions and rotation routines were imported from SHTools package, which was originally designed to provide tools for spherical harmonic analysis in geosciences[37]. The routines of SHTools are essentially Python wrappers around Fortran-compiled libraries, which ensures their fast runtime. The spherical harmonics expansion algorithm used in SHTools employs the scheme by Driscoll and Healy[36], which is based on the fast Fourier transform. Spherical harmonic expansion of a function defined on a simple Driscoll-Healy 400×400 grid mentioned above, yields spherical harmonics coefficients up to a degree of $l_{\max} = 200$. After the effects of random averaging are calculated, the spherical harmonics expansions are only conservatively truncated to the degree $l_{\max} = 50$, which is possible due to a described low-pass filtering effect. This truncation is necessary because the size of some temporary data structures involved in the calculation of $\lambda(\theta, \phi)$ scale as $\sim l_{\max}^3$, which takes up a lot of computer RAM. Note, that theory outlined in the “Theoretical background” section uses complex spherical harmonics, thereby making mathematical expressions rather compact, whereas SHTools operates with real spherical harmonics, which relationship to the complex ones is described in the “Spherical harmonics basics” section of the SI. The effects of random rotations described in the “Theoretical Background” were rewritten and coded accordingly, as sketched in the “Notes on averaging due to random rotations” section of the SI. All the calculations were performed using a desktop computer equipped with Intel® Core™ i3-4150 CPU running a virtual machine with Ubuntu 16 operating system.

Model biradical

Model-based simulation of nitroxide biradical DEER traces. A model biradical, which structure is shown in Fig. 3a is used to demonstrate applicability of the spherical harmonics-

based theory. There, the two nitroxide radicals are tethered via a rigid linker making a distance of 3.75 ± 0.13 nm between paramagnetic centres[16]. The N-O bond in each radical is tilted by $\approx 25^\circ$ with respect to the axis of the rigid linker. In a simple geometric model suggested earlier[16], the distribution of conformations arises due to unrestrained rotation around acetylene bonds of the linker, which is schematically shown by dashed cones in Fig. 3a. The bending angle of the linker is responsible for an additional contribution to the conformational space. Earlier molecular dynamics simulations have demonstrated a slight flexibility in the rigid linker[19], producing a Gaussian distribution of bending angles with a characteristic width of about $\approx 5^\circ$, as shown schematically by dashed red lines in Fig. 3a.

A field-sweep EPR spectrum of the biradical, shown in Fig. 3b, shows a very good agreement with an EasySpin simulation, which uses spin Hamiltonian parameters reported for a similar system by Savitsky et al.[29]. For that reason, the same spin Hamiltonian parameters were used for calculations of excitation PDFs. Calculation of modulation depth values and weights of modified Pake patterns is then carried out using the following steps.

Step 1. For each value of the experimental magnetic field, orientations excited by the pump and observer pulse are described by the PDFs $f_{\text{obs}}^{(A/B,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B/A,g)}(\theta, \phi)$, which are calculated numerically.

Step 2. For these PDFs, defined on a grid of θ and ϕ angles, the sets of SHCs are calculated using routines from the SHTools package.

Step 3. Now, as the PDFs $f_{\text{obs}}^{(A/B,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B/A,g)}(\theta, \phi)$ are represented by their SHCs, the effect of random rotations transforming them from their g -frames into a dipolar frame, described by the PDFs $f_{\text{pump}}^{(A/B,dip)}(\theta, \phi)$ and $f_{\text{obs}}^{(B/A,dip)}(\theta, \phi)$ is calculated numerically. This is coded using a combination of SHC rotation routines provided by SHTools and some home-

written routines based on analytical expressions for SHC transformations as explained already in the section “Averaging due to random rotations”. In particular for the nitroxide biradical, such calculation involves averaging due to random rotations around the biradical axis and accounting for a Gaussian distribution of the linker bending angles.

Step 4. The PDFs $f_{\text{pump}}^{(\text{A/B,dip})}(\theta, \phi)$ and $f_{\text{obs}}^{(\text{B/A,dip})}(\theta, \phi)$ represented by their SHCs are used to calculate the SHCs of $\lambda(\theta, \phi)$ using a multiplication routine from SHTools. These SHCs are used to produce modulation depth values and weights of MPP components $w_k^{\text{MPP}}(B_j)$ for each value of the experimental magnetic field B_j as shown by Eqs. (26) and (31).

Step 5. DEER traces are produced based on a given distance distribution and the weights of modified Pake patterns calculated at the previous step.

Fig. 3c shows the background-corrected experimental DEER traces, normalized to unity, while their corresponding modulation depth values are shown in Fig. 3b. The background correction of DEER datasets was done manually, which gives an uncertainty of about 20% as displayed in the error bars of the experimental modulation depth values. The orientation selection effect is prominent both in the DEER trace shape and in their modulation depths. As seen in Fig. 3b the shape of the DEER traces varies as a function of the external magnetic field, i.e. it depends on the position where MW pulses are applied. In addition, as seen in Fig. 3b, the experimental modulation depth values are rather distinct from the ones expected for a pair of nitroxide radicals with entirely random relative orientations, which is another signature that orientation selection is prominent. Simulation carried out using the simple geometric model in Fig. 3a according to the procedure outlined in steps 1 to 5, produces DEER traces and modulation depth values depicted in Fig. 3c and b respectively. Both represent a rather good agreement with experimental results, thus showing the validity of model-based simulations using spherical harmonics approach.

The intermediate **step 4** of the simulation procedure calculates the weights of MPP components $w_k^{\text{MPP}}(B_j)$ contributing to the DEER spectra at each specific value of the experimental magnetic field B_j . As shown in Fig. 3d, the coefficient values rapidly decrease with an increase of k , which means that the number of MPP components contributing to a DEER trace in this case is actually rather small. This result agrees with the filtering effect described in the “Theoretical Background”, where the distribution of orientations is shown to act as a low-pass filter applied to excitation PDFs. As a result, the contributions of higher degree MPP components is also significantly reduced.

Exploring the parameter space in model-based simulations. The orientation distributions can often be easily parametrized and the corresponding filters affecting SHCs as described in **step 3**, can be calculated using analytical expressions. For that reason, calculations using SHCs provide a convenient tool for exploring many models, where distribution parameters are varied. In contrast, a more conventional approach presented in EPR literature[18–21,38], generates individual members of a conformational ensemble, and sums up their contributions to produce a DEER trace. The approach based on SHCs in turn accounts for the distributions using analytical expressions.

To demonstrate this, the simple geometric model shown in Fig. 3a is modified to include two parameters: 1) a NO bond angle Ψ with respect to the linker axis, and 2) a characteristic width $\Delta\Phi$ of a Gaussian distribution for the backbone bending angle. Fig. 4 shows a 2D map plotting a goodness of fit, expressed in a form of a reduced chi-square χ_r^2 [39], as a function of Ψ and $\Delta\Phi$, where the parameters vary in a very generous and *physically unrealistic* range $0^\circ < \Psi < 90^\circ$ and $0^\circ < \Delta\Phi < 180^\circ$. As can be seen from the graph, the angles $\Psi = 30^\circ$ and $\Delta\Phi = 15^\circ$ are close to the original model with $\Psi = 25^\circ$ and $\Delta\Phi = 5^\circ$ and provide a slightly better

agreement with the experimental data than the latter. The DEER traces simulated with $\Psi = 30^\circ$ and $\Delta\Phi = 15^\circ$ are shown in Fig. S1.

Since the smallest value of $\chi_r^2 \approx 3.5 > 1$, it means that even though the overall quality of fit is fairly good, such a model still does not perfectly fit the experimental data. This points to the fact that orientation distributions cannot be adequately described by the simple geometric model presented in Fig. 3a and the molecular conformational space most likely has a somewhat more complicated structure. Fig. 4 also demonstrates that χ_r^2 has a rather large flat region, where $\chi_r^2 < 4.5$, showing that experimental data can be fit with a similar quality by rather different sets of model parameters. This result generally agrees with findings of Marko and Prisner[18], who pointed out to a similarity of some DEER traces corresponding to rather different geometries of a spin pair. For that reason, interpretation of orientation information in general cannot be done without employing a comparison with some physically realistic model. Nevertheless, the presented results demonstrate well how the spherical harmonics-based theory can be applied to exploring the space of parameters describing orientation distributions.

Model-free disentangling of the distance and orientation information. As shown in Fig. 3d the number of MPP components contributing to the DEER spectra of the nitroxide biradical is rather small, with components $k = 0, 1, 2$ having the greatest contributions. As mentioned in the “Modulation depth and DEER spectra” section, the shape of MPP components with $k \geq 3$ can also be approximated to a certain degree with a linear combination of the first three components with $k = 0, 1, 2$. Given that experimental data always contain some finite noise, the difference between the actual component and its approximation is likely to be indistinguishable.

Further evidence for the small number of contributing MPP components can be obtained using principal component analysis (PCA). The experimental DEER traces of the biradical

measured at various magnetic fields were subjected to the PCA using singular value decomposition[40]. PCA treats each DEER trace recorded at a certain value of the magnetic field as a linear combination of so called principal components. If the number of DEER datasets in the analysis is N , the PCA produces N principal components and their contributing weights. Fig. S2a shows all 6 principal components extracted from the 6 experimental DEER traces, shown in Fig. 3c. As seen from the graph, only the first three principal components are distinguishable from the noise level. Fig. S2b demonstrates that reconstructing the DEER traces using only the first two principal components produces a rather good representation of original data. Adding a third component as shown in Fig. S2c, produces only a slight improvement. Overall the PCA shows that most of the variation in the DEER dataset arises due to the first *three* principal components, which implies that only *three* MPP components with degrees $k = 0, 1, 2$ would be necessary to describe the biradical DEER traces.

As follows from Eqs. (5) and (25), the background corrected and normalized to unity DEER trace $S_j(t)$ recorded at a magnetic field B_j can be found using a distance distribution $f(r)$ and an infinite series of kernels $K_k(r, t)$:

$$S_j(t) = \sum_{k=0}^{k_{\max}=\infty} p_k(B_j) \int_{r_{\min}}^{r_{\max}} K_k(r, t) f(r) dr, \quad (32)$$

$$K_k(r, t) = 2\pi \int_0^\pi \cos(\omega(r, \theta)t) Y_{2k,0}(\theta, \phi) \sin \theta d\theta.$$

The coefficients $p_k(B_j)$ encode orientation information and converge to the MPP component weights at the magnetic field B_j , i.e. $p_k(B_j) \rightarrow w_k^{\text{MPP}}(B_j)$ when $k_{\max} \rightarrow \infty$. As discussed above, the number of contributing MPP components is limited, therefore the largest contributing component degree can be set to some $k_{\max} \neq \infty$, specifically, for reasons discussed above, for the biradical DEER data it is most practical to truncate the sum in Eq. (32) to $k_{\max} = 2$. Both the distance distribution $f(r)$ and the orientation encoding coefficients

$p_k(B_j)$, can be found by minimizing a functional arising from Eq. (32). However, it is more convenient to split minimization with respect to $f(r)$ and $p_k(B_j)$ into two parts, and treat them in an iterative manner. This produces a Tikhonov regularization-based **iterative processing algorithm**, which consists of repeating the following two steps:

Step 1. The distance distribution satisfying Eq. (32) for all the experimental magnetic fields B_j simultaneously, can be found by finding $f(r)$ which minimizes the following expression:

$$\sum_{j \text{ over all } B_j} \left[\int_{r_{\min}}^{r_{\max}} \sum_{k=0}^{k_{\max}} p_k(B_j) K_k(r, t) f(r) dr - S^j(t) \right]^2 + \zeta \int_{r_{\min}}^{r_{\max}} |f(r)|^2 dr \rightarrow \min, \quad (33)$$

where a regularization parameter ζ is introduced to limit the norm of $f(r)$, thereby stabilizing the found solution. Essentially, this is a version of a Tikhonov regularization procedure for solving all the integral equations in Eq. (32). The coefficients $p_k(B_j)$ are taken as some random numbers between 0 and 1 at the first iteration and for further iterations they are produced at **step 2**. In practice, finding a minimum of the expression in Eq. (33) is carried out using a non-negative least squares fit, because a physically realistic distance distribution $f(r)$ cannot be negative.

Step 2. The resulting $f(r)$ is plugged into the equations Eq. (32), where $p_k(B_j)$ are now treated as parameters in a linear regression with Tikhonov regularization, minimizing the following:

$$\sum_{j \text{ over all } B_j} \left[\int_{r_{\min}}^{r_{\max}} \sum_{k=0}^{k_{\max}} p_k(B_j) K_k(r, t) f(r) dr - S^j(t) \right]^2 + \eta \left[\sum_{j \text{ over all } B_j} \sum_{k=0}^{k_{\max}} (p_k(B_j))^2 \right] \rightarrow \min. \quad (34)$$

Here, η is a regularization parameter required for restraining the overall amplitude of all the coefficients $p_k(B_j)$, because without regularization the iterative procedure diverges with $p_k(B_j) \rightarrow \infty$ and $f(r) \rightarrow 0$. This happens because the products of $p_k(B_j) \times f(r)$ contribute to the first term in both Eqs. (33) and (34). Note, that in contrast to **step 1**, values of $p_k(B_j)$

coefficients are allowed to be negative. The new values of $p_k(B_j)$ are then used as an input for **step 1** and the calculation is repeated iteratively until convergence. The convergence always proceeds towards the same result regardless of the weight values chosen at the first step

Note that this iterative processing algorithm does not take into account the magnitude of modulation depth values, because of a rather large uncertainty in determining those from the experimental data. However, when the quality of data is better, the algorithm can also use non-normalized DEER data as an input.

The best regularization parameter ζ was determined using an L-curve criterion, shown in SFig. 3a. The second regularization parameter η can be set to an arbitrary value ($\eta = 1$ was chosen for definiteness), because it only affects the scaling of solutions for $f(r)$ and $p_k(B_j)$ emerging from the processing.

Fig. 5a,b shows a distance distribution and $p_k(B_j)$ values obtained using the processing algorithm. The distance distribution shows a single peak centred at 3.75 nm with a full width and half height (FWHM) of about 0.3 nm, which provides a fairly good agreement with the expected values. Since these experimental DEER traces were collected up to $t = 2.3$ us, the largest meaningful distances are limited to ~ 5 nm. The small intensity feature appearing at ~ 5 nm is most likely a signature of an imperfect background correction. Confidence bounds for the distance distribution were obtained using a version of a bootstrap method[41]. There, resampling with replacement is used to produce an ensemble of a 1000 resampled DEER datasets. For convenience, the distance distribution and the orientation encoding coefficients minimizing both functionals in Eqs. (33) and (34) are labelled as $f^{\min}(r)$ and $p_k^{\min}(B_j)$. For each “ q ”-th dataset of the resampled ensemble the corresponding distance distribution $f^{(q)}(r)$ is produced using a procedure described in **step 1**, where coefficients $p_k(B_j)$ are assigned to $p_k^{\min}(B_j)$. The shaded areas in Fig. 5a represent 95% bounds for the ensemble of all $f^{(q)}(r)$

distance distributions generated in this manner. As can be seen from the graph, the intensity variation of small peaks around 4.4 nm and 5 nm is very large, meaning a very low likelihood that they are physically meaningful. In a similar manner, the bootstrap method was used to determine the uncertainties in the coefficients $p_k(B_j)$ using a procedure described in **step 2**, where $f(r)$ is assigned to $f^{\min}(r)$. The error bars in Fig. 5b represent 95% confidence bounds for $p_k(B_j)$. The DEER traces simulated using the best fit $f(r)$ and $p_k(B_j)$ are shown in Fig. S3b. It should be emphasized, that the presented uncertainties in $f(r)$ and $p_k(B_j)$ have a limited meaning of variance of one quantity given a certain value of the other quantity. At the moment only a full treatment using Bayesian approach was shown to provide all variances independent of one another, including those in a regularization parameter[42]. Such approach however, would be computationally too demanding here, given that the iterative processing for obtaining $f(r)$ and $p_k(B_j)$ takes about a 0.5-1 minute to converge on a regular desktop PC.

The coefficients $p_k(B_j)$ obtained using the processing algorithm can be compared with the corresponding weights $w_k^{\text{MPP}}(B_j)$ predicted for various values of the magnetic field B_j for the simple geometric biradical model shown in Fig. 3a. The comparison shown in Fig. S3c. provides the coefficient of determination $R^2 \approx 0.53$, which means that only 53% of the variation in $p_k(B_j)$ values can be explained by this model. This is not surprising because the simple geometric model does not fully account for the complexity of the conformational space. Furthermore, the iterative processing algorithm takes into account only MPP components up to a degree $k_{\text{max}} = 2$, so the orientation information contained in the MPP weights $w_k^{\text{MPP}}(B_j)$ with $k \geq 3$ is folded into the coefficients $p_k(B_j)$.

Overall, the results in Fig. 5 demonstrate that iterative processing algorithm is capable of disentangling the distance distribution $f(r)$ and the orientation information in a model-free manner. While the physical meaning of the distance distribution is straightforward, the

orientation information encoded in $p_k(B_j)$ has no obvious interpretation and such interpretation most likely cannot be obtained without some appropriate molecular models.

Model protein

MMM model-based simulation of spin-labelled protein DEER traces. The model-based approach employing spherical harmonics can also be applied to simulating DEER data obtained for a spin-labelled protein. The molecular model of τ_{C14} protein, shown in Fig. 6a, was previously obtained by solution NMR spectroscopy (PDB ID: 2AYA)[43]. The MMM software[26] was used to calculate the cryogenic (temperature 175 K) conformations and relative populations of MTSL rotamers in τ_{C14} , where R1 spin-labels, shown in the bottom of Fig. 6a, are introduced at Cys56 and Ser31 replaced with a cysteine via site-directed mutagenesis. Some of the R1 residue rotamers are shown attached to the structure of τ_{C14} in Fig. 6a. An ensemble of conformers is composed of all proteins models (20 models in the 2AYA structure) with associated MTSL rotamers at C56R1 and Ser31R1 (216 rotamers for each site). Each one out of $216 \times 216 \times 20$ conformers is characterized by its probability and Euler angles connecting nitroxide g -frames and the dipolar frame. The distance distribution between the two nitroxides is assumed to be a Gaussian with a centre at 2.9 nm and FWHH of 0.6 nm, as arises from both the MMM modelling and the molecular dynamics simulations done previously[35].

Using an ensemble of conformers generated in this manner, the calculation of DEER traces and modulation depths at various magnetic fields can then be carried out using a procedure similar to the one used above to obtain the results for the model biradical.

The EPR spectrum of the spin-labelled τ_{C14} and its EasySpin simulation is shown in Fig. 6b. The spin Hamiltonian parameters, such as a g -anisotropy tensor and a nitrogen hyperfine anisotropy tensor, were obtained by an automated fitting routine in the EasySpin software[28].

As seen in Fig. 6b, the simulation with these best fit parameters provides a very good agreement with the experimental spectrum. The spin Hamiltonian parameters thus obtained, are then used to calculate the PDFs $f_{\text{pump}}^{(A/B,g)}(\theta, \phi)$ and $f_{\text{obs}}^{(B/A,g)}(\theta, \phi)$, which in turn are used to produce the PDFs $f_{\text{pump}}^{(A/B,\text{dip})}(\theta, \phi)$ and $f_{\text{obs}}^{(B/A,\text{dip})}(\theta, \phi)$ by a weighted summation over an ensemble of conformers. Finally, those are used to obtain $\lambda(\theta, \phi)$ and its SHCs, enabling the calculation of the modulation depths and DEER traces, which are shown in Fig. 6b and c. The DEER traces calculated in this manner show a fairly good agreement with the experiment as seen in Fig. 6c. The orientation selection in these is evident from the dependence of the trace shape on the experimental magnetic field. In addition, the theoretical modulation depths for the protein model obtained by MMM and for randomly oriented nitroxides are somewhat different as shown in Fig. 6b, which is another signature of orientation selection presence. The MMM-based theoretical modulation depth values mostly agree with the experiment, except at the field 3378.7 mT, which may stem from the following three factors, listed below in the order of decreasing effect:

1. The conformational space produced by MMM may be somewhat different from what is actually present in the sample. Previously, it was shown that an ensemble of rotamers produced by MMM is significantly wider than what is observed in X-ray protein structures[44]. If not properly accounted, the excitation PDFs $f_{\text{pump}}^{(A/B,\text{dip})}(\theta, \phi)$ and $f_{\text{obs}}^{(B/A,\text{dip})}(\theta, \phi)$ calculated by MMM will be distinct from the experimental ones, thereby affecting the modulation depth and MPP weights.
2. Inaccurate g -anisotropy and hyperfine tensors may be an additional source of error, because they were determined solely based on the EPR spectrum fitting, i.e. without additional restraints from other experiments, such as done for example by Savitsky et al.[29]. Furthermore, the two MTSL labels in τ_{C14} may be located in slightly

different polar environments, which affects their spin Hamiltonian parameters[30], whereas the EPR spectrum simulation assumes the same parameters for both radicals in the pair. This factor affects only the PDFs $f_{\text{pump}}^{(\text{A/B,g})}(\theta, \phi)$ and $f_{\text{obs}}^{(\text{B/A,g})}(\theta, \phi)$. For example, the EasySpin fits of the EPR spectrum are not entirely unique, because about 5% different $(g_{xx} - g_{yy})$ with correspondingly adjusted values of A_{xx} still provide a reasonably good agreement with the experimental data. Such error translates into about 5% uncertainty in the values of SHCs. However, the PDFs $f_{\text{pump}}^{(\text{A/B,dip})}(\theta, \phi)$ and $f_{\text{obs}}^{(\text{B/B,dip})}(\theta, \phi)$, determining $\lambda(\theta, \phi)$ and consequently the modulation depth and the MPP weights, are produced by averaging, which should significantly mask this error. For that reason, an inaccuracy in the input spin Hamiltonian parameters should be much less significant than inaccurate accounting of the conformational space (factor 1).

3. Various experimental parameters, such as inaccurately set magnetic field or errors in setting MW pulse durations and a cavity tuning in a DEER experiment may also lead to somewhat different PDFs $f_{\text{pump}}^{(\text{A/B,g})}(\theta, \phi)$ and $f_{\text{obs}}^{(\text{B/A,g})}(\theta, \phi)$. Here, as in factor 2, the same argument about the masking effect of averaging applies, making it less likely to affect the results than inaccurate accounting of the conformational space (factor 1).

Fig. 6d shows the weights of MPP components contributing to the τ_{c14} DEER data. Similar to the model biradical example discussed above, the rather significant conformational distribution effectively acts here as a low-pass filter retaining only the MPP components with degrees $k \leq 2$. The principal component analysis, shown in Fig. S4a,b, also demonstrates that only a single principal component needs to be added to the average to describe all the traces in

the τ_{C14} DEER dataset. All this points to the fact that higher degree MPP components have rather small contributions.

Model-free disentangling of the distance and orientation information. Since only MPP components with degrees $k \leq 2$ contribute, a model-free disentangling of the orientation and the distance information can be performed by applying the Tikhonov regularization-based iterative processing algorithm outlined in the “Model biradical” section. Analogously, the highest degree components are truncated to $k_{\max} = 2$. The optimal regularization parameter ζ in Eq. (33) was obtained using an L-curve criterion, as shown in Fig. S5a, while the regularization parameter η in Eq. (34) is kept constant $\eta = 1$. The distance distribution and the $p_k(B_j)$ coefficients resulting from the iterative algorithm are shown in Fig. 7a,b. The DEER data fits show a rather good agreement with the experiment, with the best-fit DEER traces shown Fig. S5b. The distance distribution is centred at 2.8 nm with a spread at the half height of $r = 2.6 \dots 3.2$ nm, which is in fairly good agreement with the distribution obtained using molecular dynamics simulations[35]. The uncertainties for the distance distributions and MPP weights were found via a bootstrap method using resampling with replacement as described above in the “Model biradical” section. The values of orientation encoding coefficients $p_k(B_j)$ were also compared with the MPP weights $w_k^{\text{MPP}}(B_j)$ produced by the MMM model of τ_{C14} . The comparison shown in Fig.S5c provides the coefficient of determination $R^2 \approx 0.54$, which means that about 54% of variation in coefficients $p_k(B_j)$ can be explained by the MPP weights $w_k^{\text{MPP}}(B_j)$. This disagreement most likely stems from a somewhat inaccurate accounting of the rotamer conformational space by the MMM software[44] as already discussed above.

Discussion

As demonstrated in this paper, the analysis of DEER data using the spherical harmonics-based theory provides a number of very important insights. The theory shows that DEER spectra can be described using a linear combination of the modified Pake pattern components. The weights of these MPP components encode orientation information in a spin system. The distribution of orientations in an ensemble of conformers acts as a low-pass filter retaining only the lower degree components, as was illustrated using a nitroxide biradical and a spin-labelled protein.

The theory using spherical harmonics formalism enables model-based simulations of the nitroxide biradical and the spin-labelled τ_{C14} protein DEER data, where a simple geometric model and a MMM-based model were used respectively. In addition, in the case of the nitroxide biradical the orientation distributions can be described by analytical functions with parameters, describing the NO-bond tilt angle and the rigid linker bending angle. This helps with analytical calculation of random rotations, and enables a search in a parameter space for a model biradical. The calculations using SHCs are rather quick, even though no special effort was taken to optimize the Python code. In particular, calculations of DEER traces for Fig. 4, where the parameter space was scanned, employed harmonics up to $l_{\max} = 50$. They were performed using a regular desktop personal computer, where the runtime takes about 33 ms per DEER trace, while for all 6 positions of the magnetic field, it takes ~200 ms. The home-written routines used to calculate the effect of SHC rotations can be further optimized using some lower-level programming language. Furthermore, the calculations can be significantly sped up, if some *a priori* knowledge of the nature of conformational disorder suggests that spherical harmonics series can be truncated to even smaller degrees.

As pointed out earlier, a direct approach for simulating DEER traces uses summation over many conformers to produce a complete DEER trace[18–21,38]. In some situations, such summation may be somewhat long computationally. For example, in order to calculate the DEER traces for a nitroxide biradical with distributed conformers Abdullin et al.[19] employs a Monte-Carlo-based integration using 10^5 conformers. A significant level of redundancy in the number of used conformers is essential, because a large portion of these conformers are not excited by the MW pulses and therefore their contribution to the DEER trace is null. In contrast, the spherical harmonics-based approach calculates the effect of orientation distributions analytically, i.e. using a method that is potentially quicker, and that may be further improved when spherical harmonics of higher degrees can be truncated. The spherical harmonics-based simulation may therefore be a computationally more advantageous approach for searching the best fits in a space of parameters defining a conformational model using techniques such as simulated annealing or genetic algorithms[19]. However, a side-by-side comparison of the direct and spherical harmonics-based methods goes beyond the scope of this paper.

The findings of the spherical harmonics-based theory were also applied for model-free analysis of DEER datasets recorded at various locations within an EPR spectrum. The approach is based on a Tikhonov regularization-based iterative processing algorithm, which can disentangle the distance and orientation information from the data. It is important for the datasets to be recorded at many positions of an experimental magnetic field, such that substantially different sets of orientations are being selected by the pump and observer pulses. Application of this method however, requires some *a priori* knowledge about the system under study:

1. The orientation distribution needs to be “moderate” enough to produce a substantial low-pass filtering, thereby allowing to truncate the largest MPP degree included in the calculation. A significant suppression of MPP components with degrees $k \geq 3$

was demonstrated in the model-based simulations of the nitroxide biradical and the spin-labelled τ_{C14} protein. However, if no model pointing to such suppression is available, experimentally, it can be tested by carrying out the principal component analysis of the collected dataset. Small number of contributing MPP components may also be the case when MW pulses excite a large portion of an EPR spectrum, e.g. in DEER experiments with nitroxide radicals carried out using X-band and Q-band EPR spectrometers. Under such conditions this model-free analysis may also be applicable, even for spin systems with a fixed orientation with respect to one another, such as a rigid biradical.

2. No significant correlation between the distance and orientation distributions is assumed to be present in a molecule. However, as will be shown further, even if this condition is not satisfied exactly, the model-free analysis can still provide meaningful results in the presented examples. If the correlation between distances and orientations is non-negligible, it means that a function $f(r, \theta, \phi)$ replaces $f(r)$ in Eq. (5). The former, however, can still be expanded as a series of spherical harmonics:

$$f(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm}(r) Y_{lm}(\theta, \phi), \quad (35)$$

where coefficients $f_{lm}(r)$ are not constants but some functions of the distance r . The iterative procedure for finding the weights can be adjusted accordingly to find all $f_{lm}(r)$. However, such approach requires validation with proper model systems, which goes beyond the scope on this paper.

As an output the iterative algorithm provides a distance distribution $f(r)$, which has a straightforward interpretation, and orientation encoding coefficients $p_k(B_j)$, which interpretation is less obvious. While the orientation distribution maps into the values of these

encoding coefficients, reconstruction of orientation distributions from these coefficients cannot be done without a good physical model. As demonstrated in the biradical example, variation of the distribution model parameters, shown in Fig. 4, can produce rather similar DEER traces.

Applicability of the main assumptions. The notion of a “moderate” conformational disorder was used so far without a reference to any numerical value. The outlined theory allows to put this on a more quantitative basis. Based on how it is constructed, the iterative processing algorithm for determining the distances produces the most reliable results when MPPs with degrees $k \geq 3$ do not contribute significantly. In the language of spherical harmonics, it implies that spherical harmonics with a degree greater or equal to $l = 2k = 6$ have very small coefficients in the expansion of $\lambda(\theta, \phi)$, shown in Eq. (25). These coefficients are suppressed due to a conformational disorder, implying some distribution in Euler angles determining the orientation of a nitroxide. The intervals between the "latitudinal" zeros of functions $Y_{lm}(\theta, \phi)$ are separated by about $\delta\theta \sim 180^\circ/l$, which for $l = 2k = 6$ gives about $\delta\theta \sim 180^\circ/6 \sim 30^\circ$. The angle distributions broaden the excitation PDFs, therefore, if the Euler angles $(\alpha_i^{(A/B)}, \beta_i^{(A/B)}, \gamma_i^{(A/B)})$ have Gaussian distribution widths of about $(\sigma_\alpha^{(A/B)}, \sigma_\beta^{(A/B)}, \sigma_\gamma^{(A/B)}) = (30^\circ, 30^\circ, 30^\circ)$, the averaging should produce $\lambda(\theta, \phi)$ with significantly suppressed harmonics $Y_{6m}(\theta, \phi)$. This provides a crude estimate from above for the required distribution widths of the Euler angle values.

When the MW pulse has a wide enough excitation bandwidth, i.e. its excitation PDFs $f_{\text{pump/obs}}^{(A/B,g)}(\theta, \phi)$ do not contain large degree spherical harmonics, this requirement may be relaxed. If a MW pulse excites a total fraction of the spectrum Δ (where a full excitation has $\Delta = 1$), the excitation PDF will contain spherical harmonics up to a degree of about $l_{\text{max}} \sim 1/\Delta$. Therefore, if the pulses excite a fraction $\Delta > 1/6^{\text{th}}$ of the spectrum, then $\lambda(\theta, \phi)$ should in principle have significantly suppressed harmonics $Y_{6m}(\theta, \phi)$. When the bandwidth of the

excitation pulses is so wide, the model-free approach can be applied even to an entirely rigid system.

Another important underlying assumption concerns a lack of correlation between orientations and distances. In the presented examples with the nitroxide biradical and the spin-labelled τ_C14 protein, the orientations and distances of the nitroxides are actually correlated, i.e. the assumption that the distance distribution function $f(r, \theta, \phi)$ is independent on the angles θ and ϕ is not valid in the strict physical sense. It is important therefore, to explain why ignoring this fact in the model-free and model-based analysis still provides reasonably good answers. The expression for the orientation-entangled distance distribution $f(r, \theta, \phi)$ in Eq. (35) can be rewritten as:

$$f(r, \theta, \phi) = f(r) + \sum_{l=1}^{\infty} \sum_{m=-l}^l f_{lm}(r) Y_{lm}(\theta, \phi), \quad (36)$$

where $f(r) = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi f(r, \theta, \phi) \sin \theta d\theta$ is the actual distance distribution, and $f_{lm}(r) = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi f(r, \theta, \phi) Y_{lm}^*(\theta, \phi) \sin \theta d\theta$.

If a real physical distance distribution $f(r)$ has non-zero values in an interval between r_{\min} and r_{\max} , then functions $f_{lm}(r)$ should also be limited to the same interval. They can be rewritten as $f_{lm}(r) = f(r)B_{lm}(r)$, where $B_{lm}(r)$ is a smooth function defined in the interval $r_{\min} \dots r_{\max}$. Its values may be positive or negative, and it may also have maxima and minima in the interval $r_{\min} \dots r_{\max}$. The product of $\lambda(\theta, \phi)$ and $f(r, \theta, \phi)$ in Eq. (5) can be calculated using expansions of spherical harmonics products into linear combinations with Clebsch-Gordan coefficients. After collecting all terms related to the same spherical harmonics, one obtains:

$$f(r)[1 + B_{10}(r)Y_{10}(\theta, \phi) + \dots][\lambda_{00} + \lambda_{10}Y_{10}(\theta, \phi) + \dots] =$$

$$= f(r) \left[\lambda_{00} + \sum_{l=1}^{\infty} \sum_{l=-m}^m (\lambda_{lm} + \lambda'_{lm}(r)) Y_{lm}(\theta, \phi) \right],$$

where functions $\lambda'_{lm}(r)$ are some linear combinations of $B_{lm}(r)$, which weights depend on λ_{lm} and Clebsch-Gordan coefficients. In other words, in a general case of the correlated distance and orientation distributions, the coefficients in the spherical harmonics expansion in Eq. (25) for $l \geq 1$ do depend on r .

At the same time, the recorded DEER trace has a finite length T . Therefore, if the dipolar frequencies differ by a value smaller than $1/T$ they cannot be distinguished. As a result,

distances differing smaller than $\delta r = \sqrt[3]{\frac{1}{52 \text{ MHz} \times T (\mu\text{s})}} \text{ nm}$, cannot be resolved. For the

nitroxide biradical, the actual physical distance distribution is rather narrow $r = 3.75 \pm 0.13 \text{ nm}$.

At the same time, the DEER trace is truncated at 2.3 μs , giving a resolution of $\delta r = 0.2 \text{ nm}$.

Such resolution of course would mask any effect arising due to the distance dependence of

$\lambda'_{lm}(r)$. Similarly, for the spin-labelled τ_{C14} protein, the distance distribution is about

$r = 2.9 \pm 0.3 \text{ nm}$. At the same time, the resolution due to a finite DEER trace is of $\delta r = 0.25$

MHz, which again masks any effect of $\lambda'_{lm}(r)$. As a result, the approximation of DEER spectra

with MPP components still applies, and the model-based and model-free simulations provide

reasonably good results in the presented examples. Effectively, this is a consequence of the

truncated DEER traces and sufficiently narrow distance distributions.

Comparison with Marko et al.'s approach[14]. The model-free analysis presented in the

current work and the approach previously proposed by Marko et al. are very closely related. In

the latter, the distance distribution is obtained from a synthetic dataset made by a summation

of many DEER traces, which were collected with an observer pulse set to many possible

spectral positions within an X-band EPR line. In the language of the spherical harmonics

theory, this means that upon summation, all higher degree components ($k > 0$) cancel each

other out, because they have MPP weights of opposing signs, while the only remaining MPP component is $k = 0$, because its MPP weight always stays positive. The datasets required for this summation need to be collected using an EPR instrument where the pulse frequencies can be set far enough apart to cover the width of an EPR line. For the nitroxide radicals at W-band this can be achieved using a bimodal[45] or a broadband[46] cavity.

However, the distance can be obtained from a synthetic dataset even when this requirement is not satisfied. For example, Jarvi et al.[47] summed DEER data collected with the pump and observer pulses set at various positions of a wide Cu^{2+} EPR line at Q-band, while the pulse frequencies being only 100 MHz apart. The obtained distance distribution agreed well with the results of their X-band experiments, where orientation selection is not so strong. The spherical harmonics-based theory suggests why the analysis of a summed dataset is a meaningful shortcut. Most likely, the summation still helps to suppress the MPP components with degrees $k > 0$ strongly enough to become undetectable with a given experimental noise level.

The data in the current report also falls into this category as shown in Fig. S6 of the Supplementary Information. The distance distributions obtained by Tikhonov regularization of the summed datasets of the nitroxide biradical and the protein, are quite similar to the ones obtained by the iterative processing algorithm. From the point of view of an experimentalist who wishes to get a rough estimate of a distance, the analysis of a summed dataset may therefore in many situations provide a quick yet accurate enough answer, even though the instrumental conditions needed for Marko's approach are not satisfied exactly. However, it would still be prudent to compare that answer with the one obtained using the iterative processing procedure, because it is unknown *a priori* whether the synthetic dataset contains any significant contributions of higher degree MPP components or not.

If the data quality is very high, the instrumental requirement of setting the pulse frequencies far apart from one another becomes essential. The datasets for the iterative processing algorithm, however, do not even need to be collected in this manner. The iterative nature of the processing helps to produce a reliable distance distribution, because it involves a functional minimization by searching for both the optimal distance and the orientation distributions simultaneously, while in Marko et.al's approach only orientation encoding functions $\tilde{\lambda}(\cos \theta) = \int_0^{2\pi} \lambda(\theta, \phi) d\phi$ are used for minimization. In other words, iterations can be viewed as a way of refining the initial guess obtained from a synthetic summed dataset.

In addition, Marko et al.'s application of Tikhonov regularization to obtain orientation encoding functions $\tilde{\lambda}(\cos \theta)$, in principle, can be viewed as a generalization of **step 2** of the iterative processing algorithm. However, in the case of a moderate conformational distribution such generalization is unnecessary, because the meaningful information on the orientations is contained only in a limited number of MPP components. Indeed, the visual inspection of $\tilde{\lambda}(\cos \theta)$ graphs for a model biradical presented in ref.[14], reveals that they can be described by some small degree polynomials of $\cos \theta$. It means that spherical harmonics series in Eq.(25) has only a few contributing terms, and therefore only small degree MPPs have contributions to the DEER traces under such experimental conditions.

To conclude, the presented spherical harmonics-based theory demonstrates applicability for a model-based and a model-free analysis of DEER data with a moderate conformational distribution (or potentially for rigid systems with large pulse excitation bandwidth). Therefore, the approaches presented here are very promising for studying a rather wide class of systems, in particular those at high magnetic fields where orientation selection effects become prominent.

Acknowledgement

The author would like to thank Prof. Daniella Goldfarb (Weizmann Institute of Science, Israel) for the helpful suggestions regarding the manuscript.

References.

- [1] A.D. Milov, K.M. Salikhov, M.D. Shchirov, Application of the double resonance method to electron spin echo in a study of the spatial distribution of paramagnetic centers in solids, *Sov. Phys. Solid. State.* 23 (1981) 565–569 .
- [2] M. Pannier, S. Veit, A. Godt, G. Jeschke, H.W. Spiess, Dead-time free measurement of dipole-dipole interactions between electron spins, *J. Magn. Reson.* 142 (2000) 331–340 doi:10.1006/jmre.1999.1944.
- [3] R. Ward, A. Bowman, E. Sozudogru, H. El-Mkami, T. Owen-Hughes, D.G. Norman, EPR distance measurements in deuterated proteins, *J. Magn. Reson.* 207 (2010) 164–167 doi:10.1016/j.jmr.2010.08.002.
- [4] G. Jeschke, DEER Distance Measurements on Proteins, *Annu. Rev. Phys. Chem.* 63 (2012) 419–446 doi:10.1146/annurev-physchem-032511-143716.
- [5] Y.D. Tsvetkov, A.D. Milov, A.G. Maryasov, Pulsed electron–electron double resonance (PELDOR) as EPR spectroscopy in nanometre range, *Russ. Chem. Rev.* 77 (2008) 487 doi:10.1070/rc2008v077n06abeh003782.
- [6] O. Schiemann, T.F. Prisner, Long-range distance determinations in biomacromolecules by EPR spectroscopy, *Q. Rev. Biophys.* 40 (2007) 1–53 doi:10.1017/S003358350700460X.
- [7] R.G. Larsen, D.J. Singel, Double electron-electron resonance spin-echo modulation:

- Spectroscopic measurement of electron spin pair separations in orientationally disordered solids, *J. Chem. Phys.* 98 (1993) 5134–5146 doi:10.1063/1.464916.
- [8] A. Schweiger, G. Jeschke, *Principles of Pulse Electron Paramagnetic Resonance*, 1st ed., Oxford University Press, USA, 2001.
- [9] A.M. Bowen, C.E. Tait, C.R. Timmel, J.R. Harmer, Orientation-Selective DEER Using Rigid Spin Labels, Cofactors, Metals, and Clusters, in: C.R. Timmel, J.R. Harmer (Eds.), *Struct. Inf. from Spin-Labels Intrinsic Paramagn. Centres Biosci. Struct. Bond.*, Springer Berlin Heidelberg, 2013: pp. 283–327 doi:10.1007/430_2013_115.
- [10] Y.-W. Chiang, P.P. Borbat, J.H. Freed, The determination of pair distance distributions by pulsed ESR using Tikhonov regularization, *J. Magn. Reson.* 172 (2005) 279–295 doi:10.1016/j.jmr.2004.10.012.
- [11] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C.R. Timmel, D. Hilger, H. Jung, DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data, *Appl. Magn. Reson.* 30 (2006) 473–498 doi:10.1007/BF03166213.
- [12] K.I. Sen, T.M. Logan, P.G. Fajer, Protein dynamics and monomer-monomer interactions in AntR activation by electron paramagnetic resonance and double electron-electron resonance, *Biochemistry.* 46 (2007) 11639–11649 doi:10.1021/bi700859p.
- [13] S.G. Worswick, J.A. Spencer, G. Jeschke, I. Kuprov, Deep neural network processing of DEER data, *Sci. Adv.* 4 (2018) 1–18 doi:10.1126/sciadv.aat5218.
- [14] A. Marko, D. Margraf, H. Yu, Y. Mu, G. Stock, T. Prisner, Molecular orientation studies by pulsed electron-electron double resonance experiments, *J. Chem. Phys.* 130 (2009) 064102 doi:10.1063/1.3073040.

- [15] A.M. Bowen, M.W. Jones, J.E. Lovett, T.G. Gaule, M.J. McPherson, J.R. Dilworth, C.R. Timmel, J.R. Harmer, Exploiting orientation-selective DEER: Determining molecular structure in systems containing Cu(II) centres, *Phys. Chem. Chem. Phys.* 18 (2016) 5981–5994 doi:10.1039/c5cp06096f.
- [16] Y. Polyhach, A. Godt, C. Bauer, G. Jeschke, Ye, Spin pair geometry revealed by high-field DEER in the presence of conformational distributions, *J. Magn. Reson.* 185 (2007) 118–129 doi:10.1016/j.jmr.2006.11.012.
- [17] J.E. Lovett, A.M. Bowen, C.R. Timmel, M.W. Jones, J.R. Dilworth, D. Caprotti, S.G. Bell, L.L. Wong, J. Harmer, Structural information from orientationally selective DEER spectroscopy, *Phys. Chem. Chem. Phys.* 11 (2009) 6840 doi:10.1039/b907010a.
- [18] A. Marko, T.F. Prisner, An algorithm to analyze PELDOR data of rigid spin label pairs, *Phys. Chem. Chem. Phys.* 15 (2013) 619–627 doi:10.1039/C2CP42942J.
- [19] D. Abdullin, G. Hagelueken, R.I. Hunter, G.M. Smith, O. Schiemann, Geometric model-based fitting algorithm for orientation-selective PELDOR data, *Mol. Phys.* 113 (2015) 544–560 doi:10.1080/00268976.2014.960494.
- [20] C. Abé, D. Klose, F. Dietrich, W.H. Ziegler, Y. Polyhach, G. Jeschke, H.-J. Steinhoff, Orientation selective DEER measurements on vinculin tail at X-band frequencies reveal spin label orientations, *J. Magn. Reson.* 216 (2012) 53–61 doi:10.1016/j.jmr.2011.12.024.
- [21] D. Margraf, B.E. Bode, A. Marko, O. Schiemann, T.F. Prisner, Conformational flexibility of nitroxide biradicals determined by X-band PELDOR experiments, *Mol. Phys.* 105 (2007) 2153–2160 doi:10.1080/00268970701724982.
- [22] M.A. Stevens, J.E. McKay, J.L.S. Robinson, H. El Mkami, G.M. Smith, D.G. Norman,

- The use of the Rx spin label in orientation measurement on proteins, by EPR, *Phys. Chem. Chem. Phys.* 18 (2016) 5799–5806 doi:10.1039/c5cp04753f.
- [23] A. Savitsky, A.A. Dubinskii, M. Flores, W. Lubitz, K. Möbius, Orientation-Resolving Pulsed Electron Dipolar High-Field EPR Spectroscopy on Disordered Solids: I. Structure of Spin-Correlated Radical Pairs in Bacterial Photosynthetic Reaction Centers, *J. Phys. Chem. B.* 111 (2007) 6245–6262 doi:10.1021/jp070016c.
- [24] V.P. Denysenkov, T.F. Prisner, J. Stubbe, M. Bennati, High-field pulsed electron-electron double resonance spectroscopy to determine the orientation of the tyrosyl radicals in ribonucleotide reductase, *Proc. Natl. Acad. Sci.* 103 (2006) 13386–13390 doi:10.1073/pnas.0605851103.
- [25] A. Marko, D. Margraf, P. Cekan, S.T. Sigurdsson, O. Schiemann, T.F. Prisner, Analytical method to determine the orientation of rigid spin labels in DNA, *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* 81 (2010) 1–9 doi:10.1103/PhysRevE.81.021911.
- [26] G. Jeschke, MMM: A toolbox for integrative structure modeling, *Protein Sci.* 27 (2018) 76–85 doi:10.1002/pro.3269.
- [27] A.D. Milov, Y.D. Tsvetkov, Double electron-electron resonance in electron spin echo: Conformations of spin-labeled poly-4-vinylpyridine in glassy solutions, *Appl. Magn. Reson.* 12 (1997) 495–504 doi:10.1007/BF03164129.
- [28] S. Stoll, A. Schweiger, EasySpin, a comprehensive software package for spectral simulation and analysis in EPR, *J. Magn. Reson.* 178 (2006) 42–55 doi:10.1016/j.jmr.2005.08.013.
- [29] A. Savitsky, A.A. Dubinskii, M. Plato, Y.A. Grishin, H. Zimmermann, K. Möbius, High-Field EPR and ESEEM Investigation of the Nitrogen Quadrupole Interaction of

- Nitroxide Spin Labels in Disordered Solids: Toward Differentiation between Polarity and Proticity Matrix Effects on Protein Function, *J. Phys. Chem. B.* 112 (2008) 9079–9090 doi:10.1021/jp711640p.
- [30] E. Bordignon, Spin-Labeling EPR of Proteins: Dynamics and Water Accessibility of Spin-Label Side Chains, *Encycl. Biophys.* (2013) 2439–2447 doi:10.1007/978-3-642-16712-6.
- [31] A.R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton University Press, 1957 doi:10.2307/j.ctt1cx3v9p.
- [32] D.A. Varshalovich, A.N. Moskalev, V.K. Khersonskii, *Quantum theory of angular momentum : irreducible tensors, spherical harmonics, vector coupling coefficients, 3nj symbols*, Singapore ; Philadelphia : World Scientific Pub, 1986.
- [33] A. Godt, C. Franzen, S. Veit, V. Enkelmann, M. Pannier, G. Jeschke, EPR Probes with Well-Defined, Long Distances between Two or Three Unpaired Electrons, *J. Org. Chem.* 65 (2000) 7575–7582 doi:10.1021/jo005556j.
- [34] A. Gennaro, A. Karabanov, A. Potapov, W. Köckenberger, Heteronuclear DNP of ^1H and ^{19}F nuclei using BDPA as a polarizing agent, *Phys. Chem. Chem. Phys.* 22 (2020) 7803–7816 doi:10.1039/D0CP00892C.
- [35] A. Potapov, H. Yagi, T. Huber, S. Jergic, N.E. Dixon, G. Otting, D. Goldfarb, Nanometer-Scale Distance Measurements in Proteins Using Gd^{3+} Spin Labeling, *J. Am. Chem. Soc.* 132 (2010) 9040–9048 doi:10.1021/ja1015662.
- [36] J.R. Driscoll, D.M. Healy, Computing Fourier Transforms and Convolutions on the 2-Sphere, *Adv. Appl. Math.* 15 (1994) 202–250 doi:10.1006/aama.1994.1008.
- [37] M.A. Wieczorek, M. Meschede, *SHTools: Tools for Working with Spherical*

- Harmonics, *Geochemistry, Geophys. Geosystems.* 19 (2018) 2574–2592
doi:10.1029/2018GC007529.
- [38] G. Jeschke, Y. Polyhach, Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance, *Phys. Chem. Chem. Phys.* 9 (2007) 1895
doi:10.1039/b614920k.
- [39] J.R. Taylor, *An introduction to error analysis : the study of uncertainties in physical measurements*, 2nd ed., University Science Books, Sausalito, Calif., 1997.
- [40] J.E. Jackson, *A User's Guide to Principal Components*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 1991 doi:10.1002/0471725331.
- [41] M.R. Chernick, R.A. LaBudde, *An Introduction to Bootstrap Methods with Applications to R*, John Wiley & Sons, Inc., Somerset, United States, 2011.
- [42] T.H. Edwards, S. Stoll, A Bayesian approach to quantifying uncertainty from experimental noise in DEER spectroscopy, *J. Magn. Reson.* 270 (2016) 87–97
doi:10.1016/j.jmr.2016.06.021.
- [43] X.-C. Su, S. Jergic, M.A. Keniry, N.E. Dixon, G. Otting, Solution structure of Domains IVa and V of the τ subunit of Escherichia coli DNA polymerase III and interaction with the α subunit, *Nucl. Acids Res.* 35 (2007) 2825–2832 doi:10.1093/nar/gkm080.
- [44] D. Abdullin, G. Hagelueken, O. Schiemann, Determination of nitroxide spin label conformations via PELDOR and X-ray crystallography, *Phys. Chem. Chem. Phys.* 18 (2016) 10428–10437 doi:10.1039/C6CP01307D.
- [45] I. Tkach, S. Pornsuwan, C. Höbartner, F. Wachowius, S.T. Sigurdsson, T.Y. Baranova, U. Diederichsen, G. Sicoli, M. Bennati, Orientation selection in distance measurements between nitroxide spin labels at 94 GHz EPR with variable dual frequency irradiation,

Phys. Chem. Chem. Phys. 15 (2013) 3433 doi:10.1039/c3cp44415e.

- [46] C.L. Motion, S.L. Cassidy, P.A.S. Cruickshank, R.I. Hunter, D.R. Bolton, H. El Mkami, S. Van Doorslaer, J.E. Lovett, G.M. Smith, The use of composite pulses for improving DEER signal at 94 GHz, *J. Magn. Reson.* 278 (2017) 122–133 doi:10.1016/j.jmr.2017.03.018.
- [47] A. Gamble Jarvi, K. Rangelova, S. Ghosh, R.T. Weber, S. Saxena, On the Use of Q-Band Double Electron-Electron Resonance to Resolve the Relative Orientations of Two Double Histidine-Bound Cu²⁺ Ions in a Protein, *J. Phys. Chem. B.* 122 (2018) 10669–10677 doi:10.1021/acs.jpcc.8b07727.

Figure Captions

Fig. 1 Illustration for excitation PDFs and their power spectra using a fictitious biradical. (a) A fictitious biradical, consisting of two flexibly linked nitroxides. A transformation from a spin system g -frame to a dipolar frame is done by three Euler angles $(\alpha_i^{(A)}, \beta_i^{(A)}, \gamma_i^{(A)})$ and $(\alpha_i^{(B)}, \beta_i^{(B)}, \gamma_i^{(B)})$ for the spin systems A and B respectively. The *red* and *green arrows* represent average orientations of the principal axes of the two frames, such that $(\bar{\alpha}^{(A)}, \bar{\beta}^{(A)}, \bar{\gamma}^{(A)}) = (0^\circ, 90^\circ, 180^\circ)$ and $(\bar{\alpha}^{(B)}, \bar{\beta}^{(B)}, \bar{\gamma}^{(B)}) = (90^\circ, 90^\circ, 0^\circ)$. The *light red* and *green arrows* represent a distribution of principal axis orientations, such that Euler angles $(\Delta\alpha_i^{(A)}, \Delta\beta_i^{(A)}, \Delta\gamma_i^{(A)})$ and $(\Delta\alpha_i^{(B)}, \Delta\beta_i^{(B)}, \Delta\gamma_i^{(B)})$ are distributed according to zero-centred Gaussian distributions with widths $(\sigma_\alpha^{(A,B)}, \sigma_\beta^{(A,B)}, \sigma_\gamma^{(A,B)}) = (30^\circ, 30^\circ, 30^\circ)$ for both A and B .

(b) Simulated EPR spectrum of the biradical shown in (a). The *Red* and *green arrows* mark the positions of the observer and pump pulses respectively, which frequencies are separated by 20 MHz.

(c)(i) Excitation PDFs $f_{\text{obs}}^{(A,g)}(\theta, \phi)$ and $f_{\text{pump}}^{(B,g)}(\theta, \phi)$, corresponding to the observer pulse excitation of the spin system A and the pump pulse excitation of the spin system B respectively

(ii) The result of the g -frame coordinate systems transformation by the average rotation with angles $(\bar{\alpha}^{(A)}, \bar{\beta}^{(A)}, \bar{\gamma}^{(A)})$ and $(\bar{\alpha}^{(B)}, \bar{\beta}^{(B)}, \bar{\gamma}^{(B)})$.

(iii) Excitation PDFs $f_{\text{obs}}^{(A,\text{dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(B,\text{dip})}(\theta, \phi)$ resulting from a rotation by the random Euler angles followed by a summation over all conformers, giving $f_{\text{obs}}^{(A,\text{dip})}(\theta, \phi)$ and $f_{\text{pump}}^{(B,\text{dip})}(\theta, \phi)$.

(iv) Pair excitation PDF $\lambda(\theta, \phi)$.

Surface plots in (c)(i), (c)(ii), (c)(iii), (c)(iv) are normalized by their own specific factors to provide a better visual colour gradient.

(d) Power spectrum of the PDFs $f_{\text{obs}}^{(A,g)}(\theta, \phi)$ (*red*) and $f_{\text{pump}}^{(B,g)}(\theta, \phi)$ (*green*) as a function of a spherical harmonic degree l .

(e) Power spectrum of the

PDFs $f_{\text{obs}}^{(\text{A,dip})}(\theta, \phi)$ (*red*), $f_{\text{pump}}^{(\text{B,dip})}(\theta, \phi)$ (*green*) and $\lambda(\theta, \phi)$ (*blue*) as a function of a spherical harmonic degree l . Total power in (d) and (e) is normalized to unity.

Fig. 2 DEER spectra consist of modified Pake pattern components. (a) Modified Pake pattern spectra for the degrees $k = 0$ (*blue*), $k = 1$ (*orange*), $k = 2$ (*green*), $k = 3$ (*red*), $k = 4$ (*purple*), $k = 5$ (*brown*) and a fixed distance $r = 3.73$ nm. (b) Time-domain DEER traces corresponding to the spectra shown in (a). Colour legend is same as (a). (c) The modified Pake pattern spectra corresponding to the degrees $k = 0, \dots, 5$ with a Gaussian distributed distance of an average value $\bar{r} = 3.73$ nm, and a variance $\sqrt{\langle (r - \bar{r})^2 \rangle} = 0.11$ nm.

Fig. 3 Model-based simulation of the nitroxide biradical DEER datasets shows a good agreement with the experiment. (a) Structure of the nitroxide biradical molecule. A simple geometric model is used to produce an ensemble of conformers. Unrestrained rotation around acetylene bond makes the NO bond to trace a surface of a cone with an angle $\Psi = 25^\circ$ (*blue dashed lines*). Bending of the rigid linker (*red dashed lines*) produces a Gaussian distribution of angles with a width $\Delta\Phi = 5^\circ$. (b) Experimental (*blue*) and simulated (*orange*) biradical EPR spectra, overlaid with modulations depth values corresponding to specific magnetic fields. The EPR spectra are normalized, the left vertical axis represents the modulation depth values. The *blue crosses (with error bars)* represent experimental modulation depth values, the *green circles* – theoretical values calculated using the simple geometrical biradical model, illustrated in (a), the *red squares* – theoretical values obtained using a model with randomly oriented nitroxides. (c) Experimental background corrected and normalized DEER traces recorded at the magnetic fields marked in (b) and overlaid with simulations using the model in (a). The DEER traces were measured at the magnetic fields B_j : 3374.7 mT (*black*), 3376.6 mT (*red*), 3378.5 mT (*green*), 3380.2 mT (*blue*), 3382.0 mT (*purple*), 3385.2 mT (*yellow*). (d) MPP component weights $w_k^{\text{MPP}}(B_j)$ as a function of a degree k , according to the model in (a). The

magnetic fields B_j and the corresponding colour legend are the same as in (c). The weights $w_k^{\text{MPP}}(B_j)$ are normalized so that $w_0^{\text{MPP}}(B_j) = 1$ for ease of comparison.

Fig. 4 Exploring the parameter space for model-based simulations of the nitroxide biradical DEER data. Reduced chi-square χ_r^2 as a function of angles Ψ and $\Delta\Phi$, shows a minimum marked with a *white arrow*. The angles are defined for the simple geometrical biradical model shown in Fig. 3a.

Fig. 5 Model-free iterative processing of the nitroxide biradical DEER datasets provides (a) a distance distribution and (b) orientation encoding coefficients $p_k(B_j)$. The *light blue* shaded areas in (a) correspond to 95% confidence bounds. The orientation encoding coefficients $p_k(B_j)$ correspond to the DEER traces measured at the magnetic fields: 3374.7 mT (*black*), 3376.6 mT (*red*), 3378.5 mT (*green*), 3380.2 mT (*blue*), 3382.0 (*purple*), 3385.2 mT (*yellow*). The error bars in (b) mark the 95% confidence bounds.

Fig. 6 MMM model-based simulation of the spin-labelled τ_{C14} protein DEER datasets shows a good agreement with the experiment. (a) (top) τ_{C14} structural model (model #1 from PDB: 2AYA) with attached R1 rotamers calculated using MMM-software. (bottom) The structure of MTSL-labelled cysteine R1. (b) Experimental (*blue*) and simulated (*orange*) τ_{C14} EPR spectra, overlaid with the modulations depth values corresponding to specific magnetic fields. The EPR spectra are normalized, the left vertical axis represents the modulation depth values. The *blue crosses (with error bars)* represent experimental modulation depth values, the *green circles* – theoretical values calculated using the MMM-model in (a), the *red squares* – theoretical values calculated using a model with randomly oriented nitroxides. (c) Experimental background corrected and normalized DEER traces recorded at the magnetic fields marked in (b) and overlaid with simulations using the MMM-model in (a). The DEER traces were measured at the magnetic fields B_j : 3378.7 mT (*black*), 3380.4 mT (*red*), 3383.0 mT (*green*), 3385.5 mT

(*blue*), 3386.5 (*purple*). (d) MPP component weights w_k^{MPP} as a function of a degree k , according to the model in (a). The magnetic fields B_j and the corresponding colour legend are the same as in (c). The weights $w_k^{\text{MPP}}(B_j)$ are normalized so that $w_0^{\text{MPP}}(B_j) = 1$ for ease of comparison.

Fig. 7 Model-free iterative processing of the spin-labelled τ_C14 protein DEER datasets provides (a) a distance distribution and (b) orientation encoding coefficients $p_k(B_j)$. The *light blue* shaded areas in (a) correspond to 95% confidence bounds. The orientation encoding coefficients $p_k(B_j)$ correspond to the traces measured at the magnetic fields B_j : 3378.7 mT (*black*), 3380.4 mT (*red*), 3383.0 mT (*green*), 3385.5 mT (*blue*), 3386.5 (*purple*). The error bars in (b) mark the 95% confidence bounds.

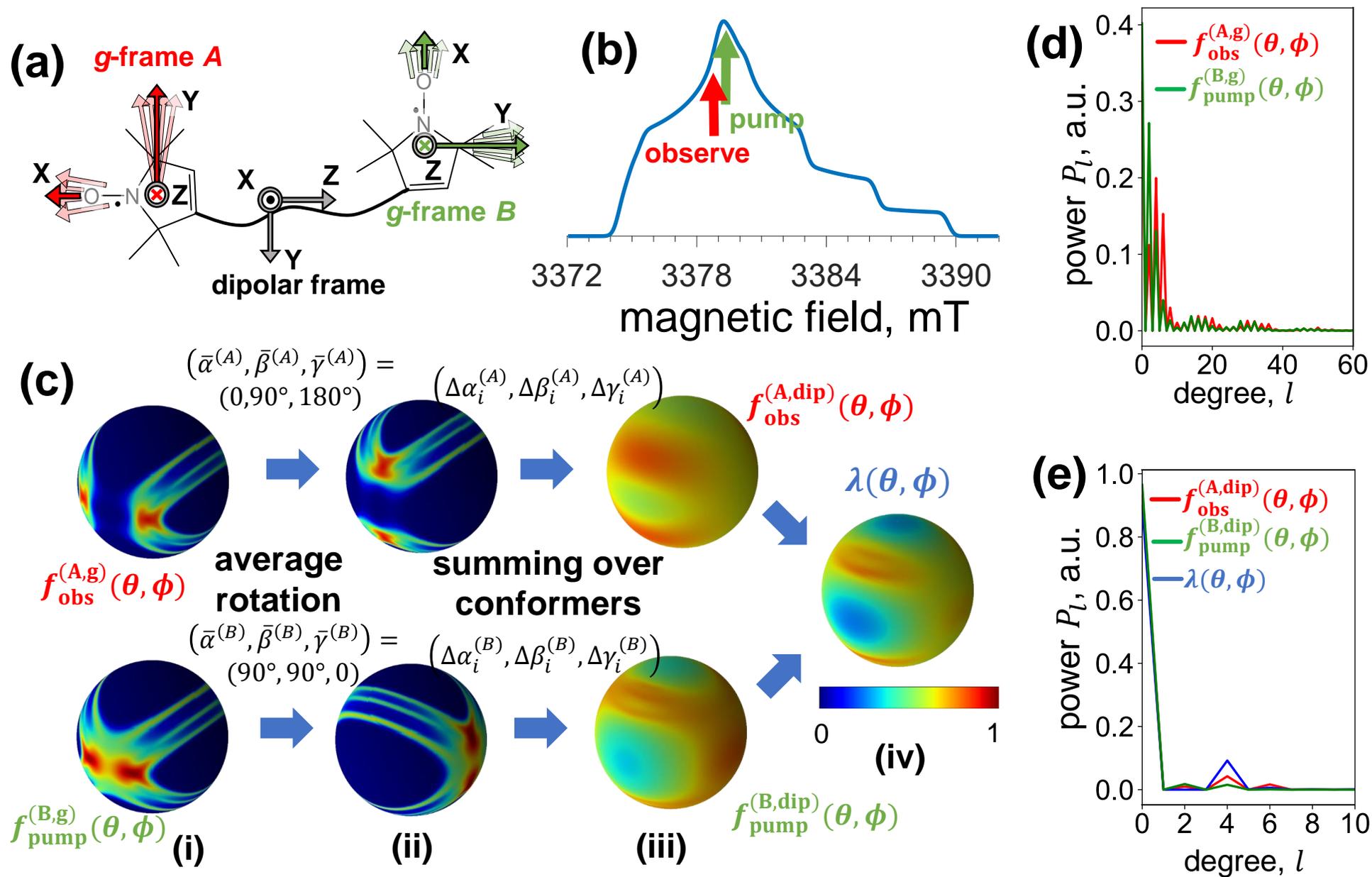


Figure 1

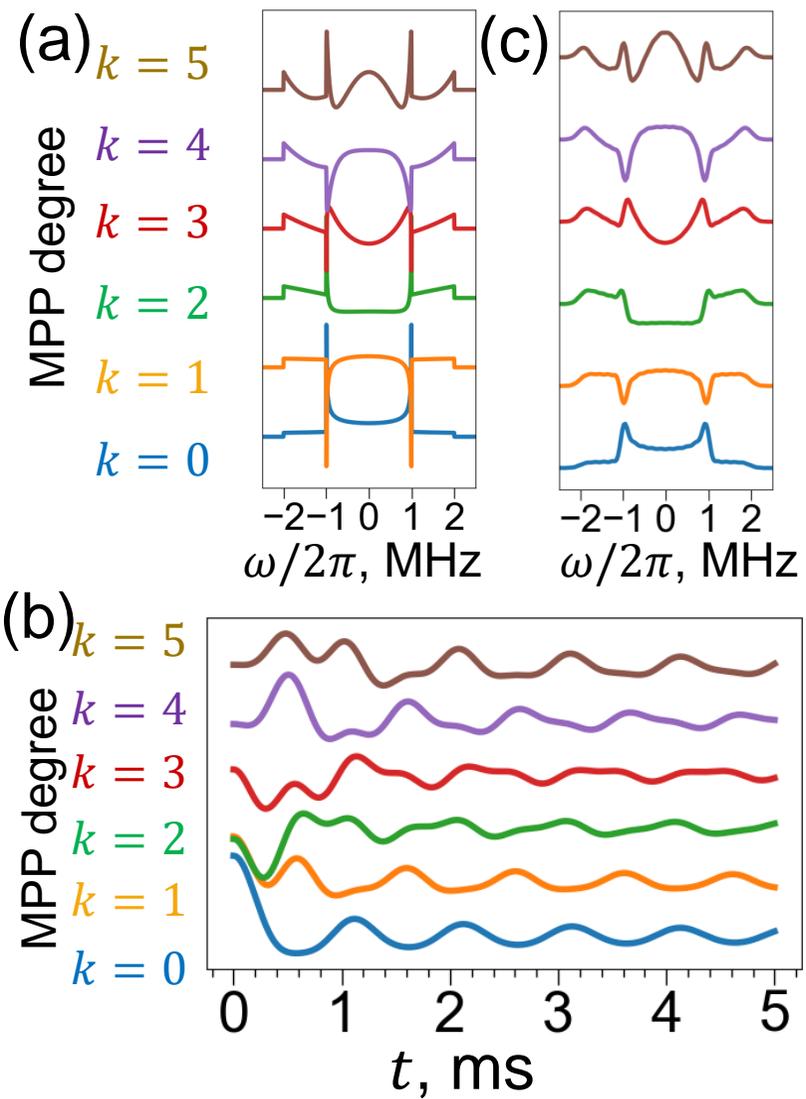


Figure 2

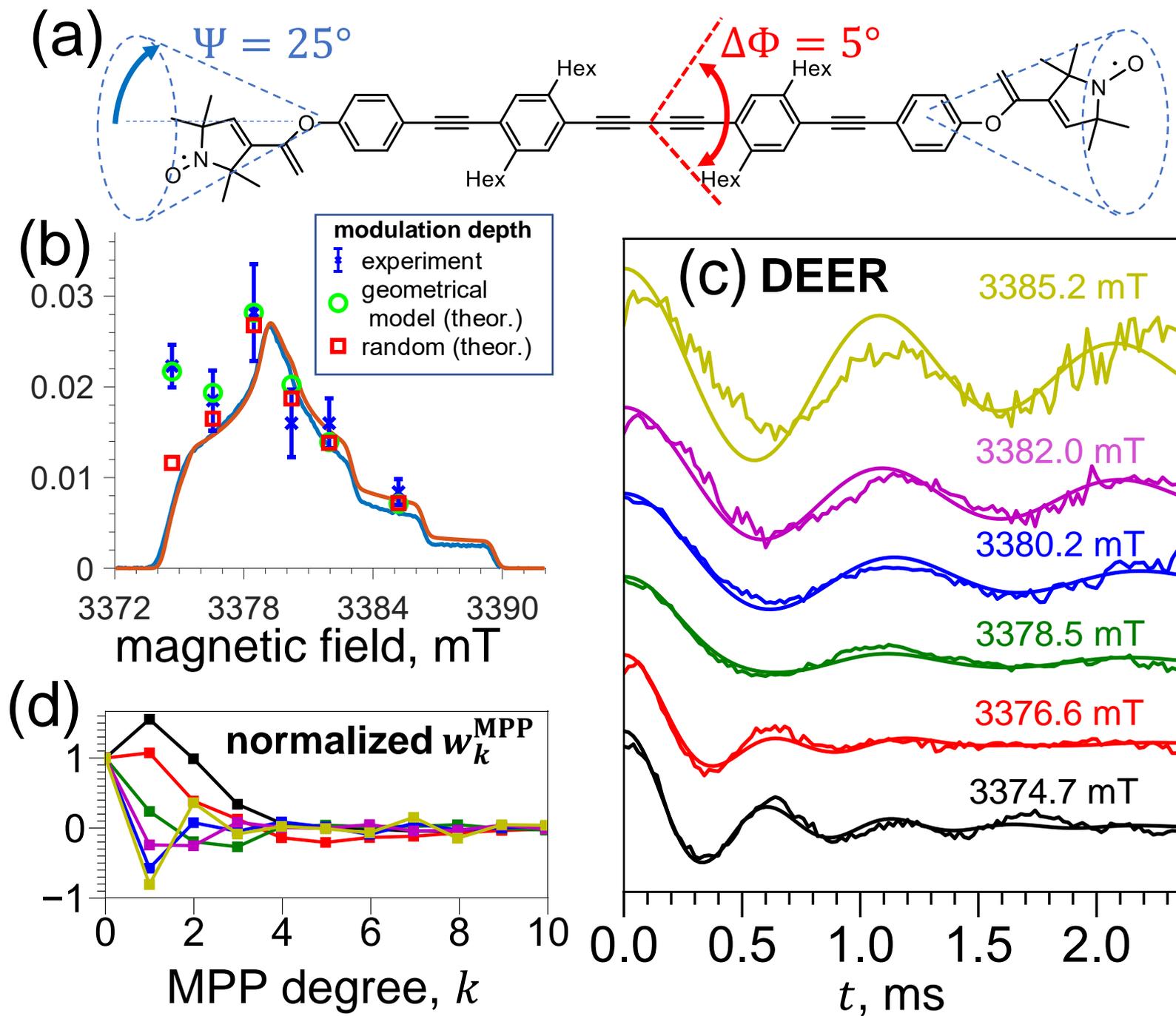


Figure 3

Figure 4

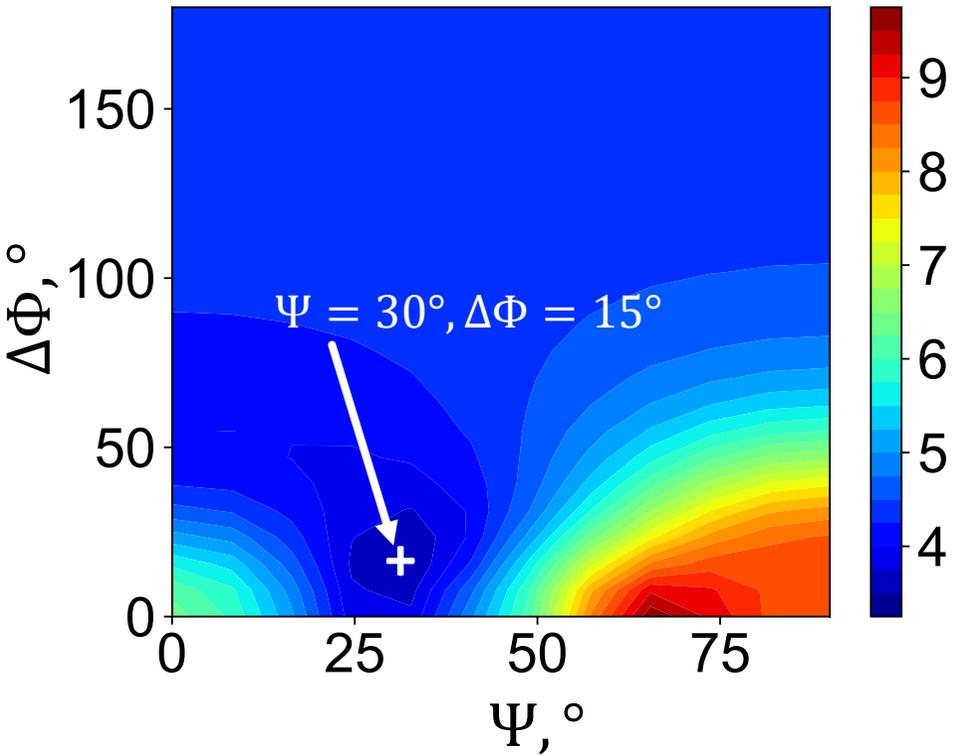


Figure 5

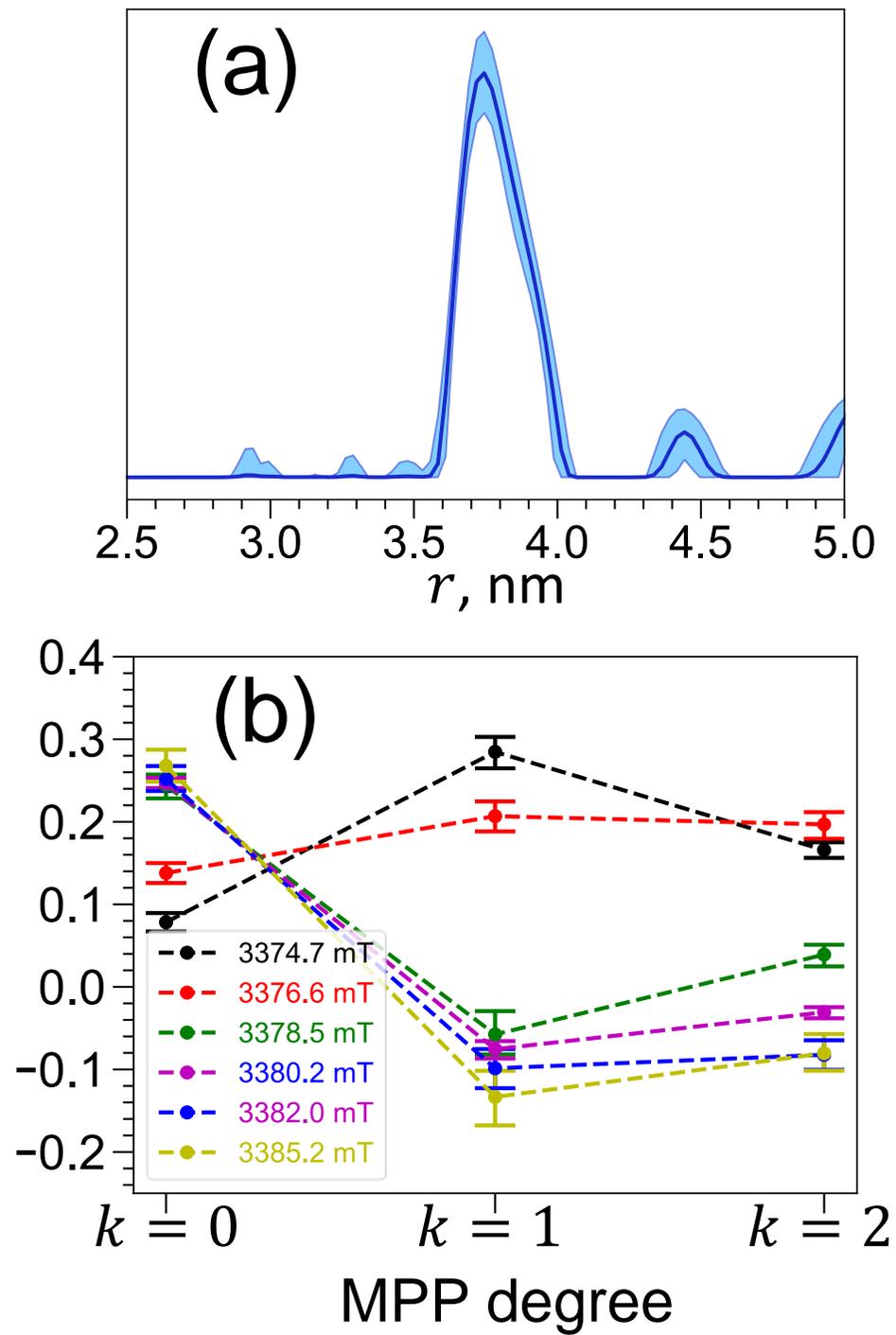


Figure 6

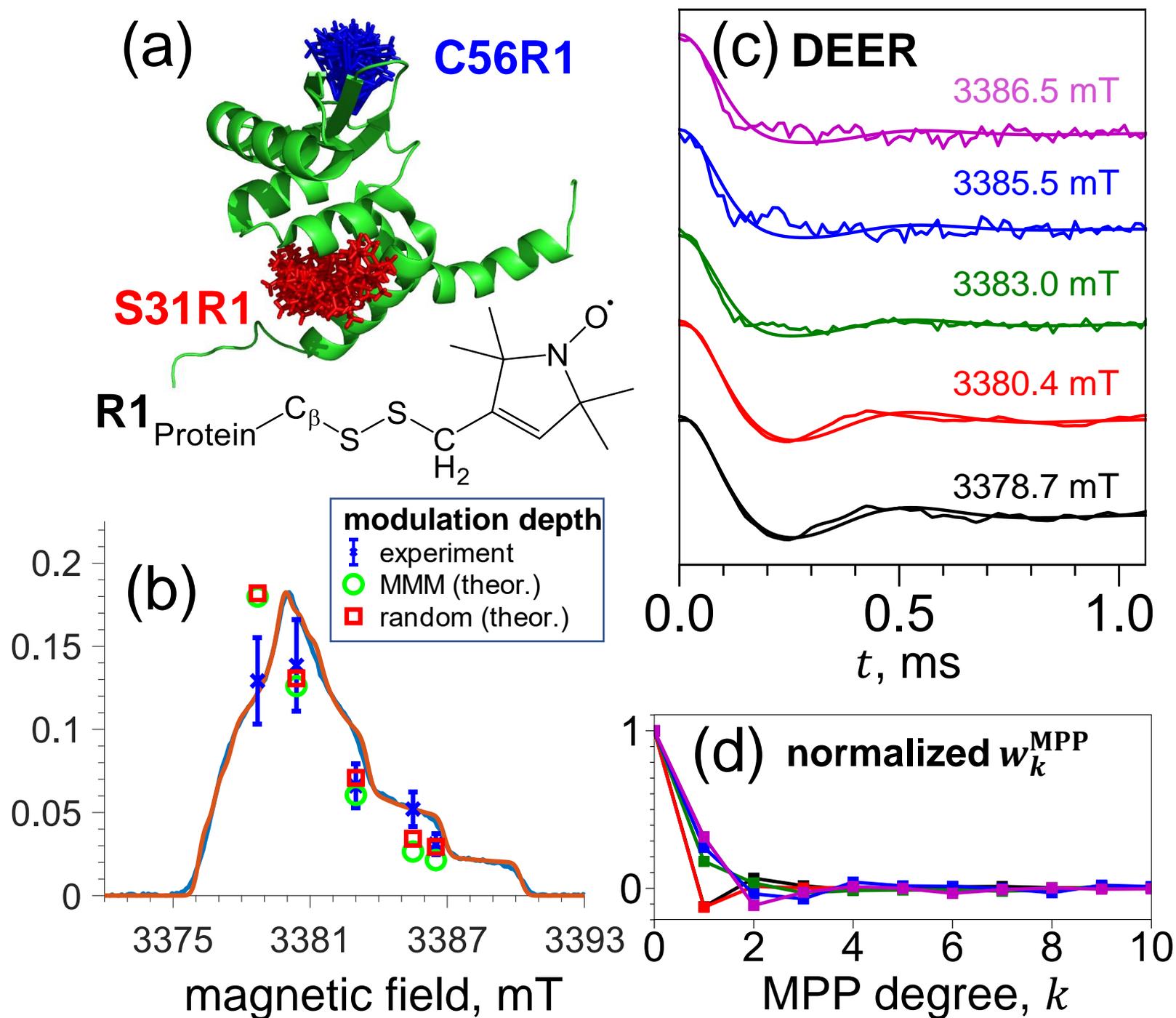
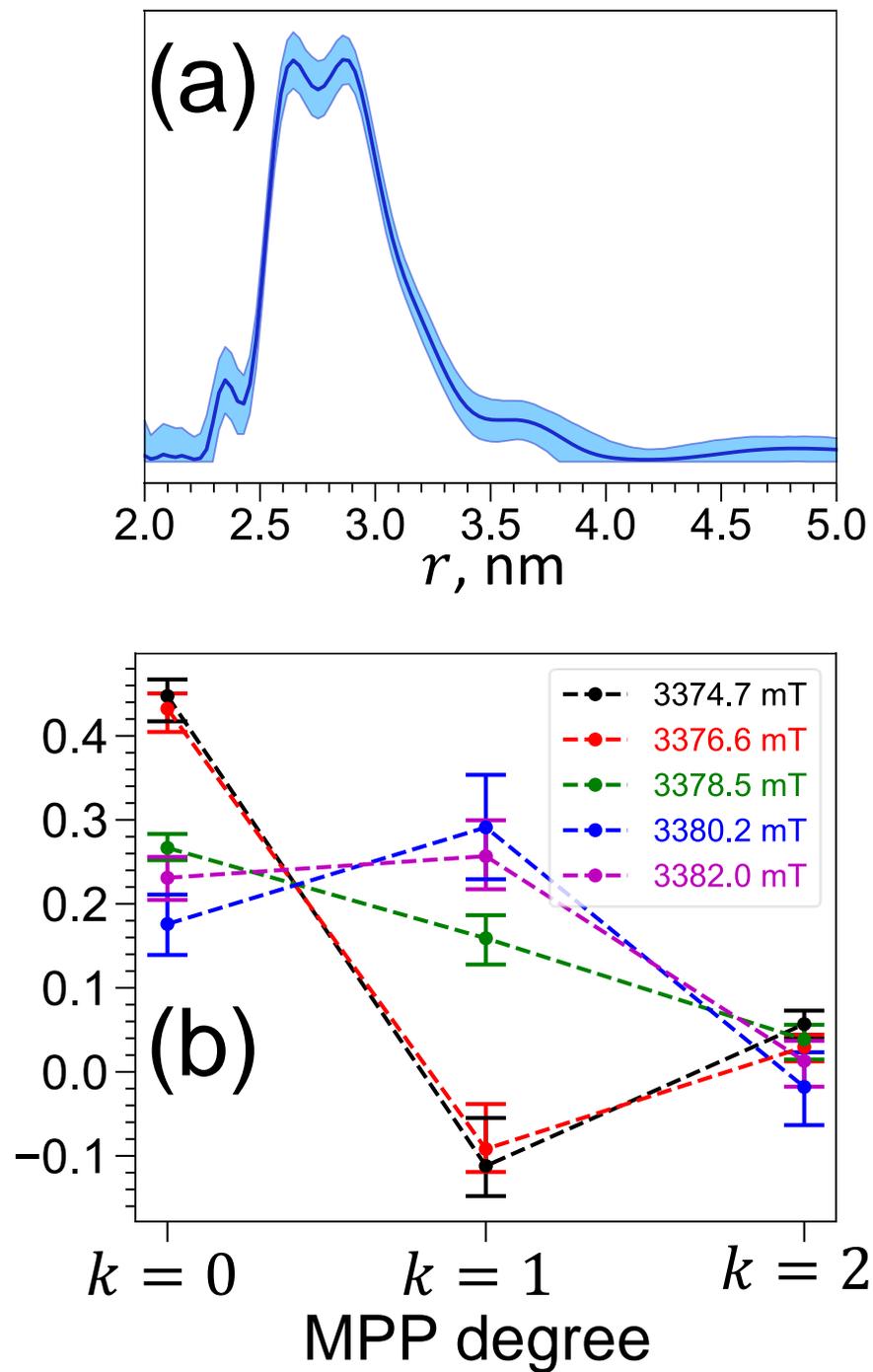


Figure 7



Supplementary Information: Application of spherical harmonics for DEER data analysis in systems with a conformational distribution.

Alexey Potapov*

School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

Contents

Spherical harmonics basics	S2
Notes on averaging due to random rotations	S3
DEER spectra	S4
References	S5
Supplementary figures	S6

*Corresponding author

Email address: alexey.potapov@nottingham.ac.uk (Alexey Potapov)

Spherical harmonics basics

The definitions of complex spherical harmonics $Y_{lm}(\theta, \phi)$ and their main properties are described in detail elsewhere[1]. Below is a brief summary of the most important aspects relevant to the theory presented in the main text.

Spherical harmonics series. A complex function $f(\theta, \phi)$ defined on a spherical surface (i.e. $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$) and satisfying a condition:

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta |f(\theta, \phi)|^2 < \infty, \quad (1)$$

can be represented as a series of spherical harmonics $Y_{lm}(\theta, \phi)$ as follows:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_{lm}(\theta, \phi). \quad (2)$$

The complex coefficients f_{lm} can be found using:

$$f_{lm} = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_{lm}^*(\theta, \phi) f(\theta, \phi). \quad (3)$$

Normalization. Note, that in contrast to the convention typically used in many quantum mechanics textbooks, this work uses 4π -normalization of the spherical harmonics, because calculations using SHTools package use such definition by default[2]. The orthogonality of spherical harmonics can be written then as:

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_{lm}^*(\theta, \phi) Y_{l'm'}(\theta, \phi) = 4\pi \delta_{l',l} \delta_{m',m}, \quad (4)$$

where $\delta_{i,j}$ is a Kronecker symbol.

Other properties. The spherical harmonics can also be expressed using other analytic functions, in particular:

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{(2l+1)(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi}, \quad (5)$$

where $P_l^m(x)$ denotes an associated Legendre polynomial of a degree l and an order m . In addition, the complex conjugates of spherical harmonics are:

$$Y_{lm}^*(\theta, \phi) = (-1)^m Y_{l,-m}(\theta, \phi). \quad (6)$$

Upon inversion of a coordinate system, spherical harmonics transform as:

$$Y_{lm}(-\theta, -\phi) = (-1)^l Y_{lm}(\theta, \phi), \quad (7)$$

which means that only harmonics with even degree l are invariant upon such transformation.

Power. Defining the power of a function $f(\theta, \phi)$ as an integral over its modulus squared, the generalized Parseval's theorem allows expressing this total power using the coefficients in its spherical harmonics expansion:

$$\frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta |f(\theta, \phi)|^2 = \sum_{l=0}^{\infty} P_l, \quad (8)$$

where P_l is the power carried by the spherical harmonics with a degree l , given by:

$$P_l = \sum_{m=-l}^l |f_{lm}|^2. \quad (9)$$

Real spherical harmonics. While the main text uses *complex* spherical harmonics, which produce more compact mathematical expressions, the SHTools package carries out calculations using closely related *real* spherical harmonics $y_{lm}^\pm(\theta, \phi)$, which are defined as:

$$\begin{aligned} y_{lm}^+(\theta, \phi) &= \frac{1}{\sqrt{2}} (Y_{lm}(\theta, \phi) + Y_{lm}^*(\theta, \phi)) = \sqrt{2} (-1)^m \operatorname{Re}[Y_{lm}(\theta, \phi)], \\ y_{lm}^-(\theta, \phi) &= \frac{i}{\sqrt{2}} (Y_{lm}(\theta, \phi) - Y_{lm}^*(\theta, \phi)) = \sqrt{2} (-1)^m \operatorname{Im}[Y_{lm}(\theta, \phi)], \\ y_{l0}(\theta, \phi) &= Y_{l0}(\theta, \phi). \end{aligned} \quad (10)$$

Since $y_{l,-m}^+(\theta, \phi) = y_{lm}^+(\theta, \phi)$ and $y_{l,-m}^-(\theta, \phi) = -y_{lm}^-(\theta, \phi)$, the real spherical harmonics expansion of a function $f(\theta, \phi)$ becomes:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=0}^l [c_{lm} y_{lm}^+(\theta, \phi) + s_{lm} y_{lm}^-(\theta, \phi)], \quad (11)$$

where the coefficients c_{lm} and s_{lm} can be calculated using complex coefficients f_{lm} :

$$\begin{aligned} c_{lm} &= \frac{(-1)^m f_{lm} + f_{l,-m}}{\sqrt{2}}, \\ s_{lm} &= i \frac{(-1)^m f_{lm} - f_{l,-m}}{\sqrt{2}}. \end{aligned} \quad (12)$$

Note, that in contrast to Eq.(2), the summation over m in Eq.(11) runs from 0 to l .

By default SHTools package uses real spherical harmonics[2], therefore the spherical harmonics series expansion of any function is represented with large arrays, containing c_{lm} and s_{lm} coefficients, which in general are complex numbers. For a real function $f(\theta, \phi) = f^*(\theta, \phi)$, the coefficients c_{lm} and s_{lm} become real and can be simplified as:

$$\begin{aligned} c_{lm} &= \sqrt{2} \operatorname{Re}[f_{lm}] (-1)^m, \\ s_{lm} &= -\sqrt{2} \operatorname{Im}[f_{lm}] (-1)^m. \end{aligned} \quad (13)$$

Notes on averaging due to random rotations

This section details some of the expressions used to calculate the effect of averaging under random rotations, and provides pathways for implementing those in computer code.

Rotation around z-axis. Using Eq.(12) and (15) of the main text, it is possible to show that a coordinate system rotation by an Euler angle α around z -axis, transforms a function $f(\theta, \phi)$ with real spherical harmonic coefficients c_{lm} and s_{lm} into a function $f'(\theta, \phi)$ which coefficients given by:

$$\begin{aligned} c'_{lm} &= c_{lm} \cos m\alpha + s_{lm} \sin m\alpha, \\ s'_{lm} &= -c_{lm} \sin m\alpha + s_{lm} \cos m\alpha \end{aligned} \quad (14)$$

Uniformly random rotation. If an Euler angle α has a uniform random distribution in the interval $-\pi \leq \alpha \leq \pi$, the averaging of Eq.(14) over α produces zeros for all c_{lm} and s_{lm} , except $c'_{l0} = c_{l0} = f_{l0}$. From the computing point of view the effect of such averaging is trivial because it is simply equivalent to setting most values of c'_{lm} and all values of s'_{lm} to null.

Gaussian distributed random rotation. If an Euler angle α is distributed according to a zero-centred Gaussian distribution $g(\alpha) = e^{-\alpha^2/\sigma_\alpha^2}$, with $-\pi \leq \alpha \leq \pi$ chosen for convenience, the averaging of Eq.(14) in the limit $\sigma_\alpha \ll 2\pi$ produces:

$$\begin{aligned} c'_{lm} &= \int_{-\pi}^{\pi} g(\alpha) [c_{lm} \cos m\alpha + s_{lm} \sin m\alpha] d\alpha \approx \\ &\approx \int_{-\infty}^{\infty} e^{-\frac{\alpha^2}{\sigma_\alpha^2}} \frac{e^{im\alpha} + e^{-im\alpha}}{2} d\alpha = \sqrt{\pi}\sigma_\alpha e^{-\frac{m^2\sigma_\alpha^2}{4}} c_{lm}, \text{ and similarly:} \\ s'_{lm} &\approx \sqrt{\pi}\sigma_\alpha e^{-\frac{m^2\sigma_\alpha^2}{4}} s_{lm} \end{aligned} \quad (15)$$

This result demonstrates, that a distributed angle α produces an averaged function which SHCs c'_{lm} and s'_{lm} dependent on a factor $\sim e^{-\frac{m^2\sigma_\alpha^2}{4}}$, which quickly decays with an increase of m . Therefore as follows from Eq.(12), f'_{lm} coefficients are suppressed by a factor of $\sim e^{-\frac{m^2\sigma_\alpha^2}{4}}$ compared to f_{lm} .

Averaging over a Gaussian distributed Euler angle γ done in a similar manner produces the same result.

The effect of averaging over an Euler angle β is easier to discuss as a result of averaging the matrix elements $\langle D_{m,m'}^{(l)}(0, \beta, 0) \rangle$. This can be qualitatively evaluated using a rotation property in Eq.(23) of the main text. Averaging the matrix $\langle D_{m,m'}^{(l)}(\beta, 0, 0) \rangle$ over angles β distributed according to $g(\beta) = e^{-\beta^2/\sigma_\beta^2}$, with $-\pi \leq \beta \leq \pi$, produces a diagonal matrix with elements $\langle D_{m,m'}^{(l)}(\beta, 0, 0) \rangle = \delta_{m,m'} \sqrt{\pi}\sigma_\beta e^{-\frac{m^2\sigma_\beta^2}{4}}$. The rotation under $\hat{D}(0, \frac{\pi}{2}, 0) \hat{D}(\frac{\pi}{2}, 0, 0)$ in Eq.(23) of the main text scrambles the matrix elements with the same degree l , meaning that l diagonal values of $\langle D_{m,m'}^{(l)}(\beta, 0, 0) \rangle$ are "spread" over the entire $l \times l$ matrix. In practice, the matrices $D_{m,m'}^{(l)}(0, \frac{\pi}{2}, 0)$ and $D_{m,m'}^{(l)}(\frac{\pi}{2}, 0, 0)$ are precalculated and recalled to be combined with $\langle D_{m,m'}^{(l)}(\beta, 0, 0) \rangle$ to provide a full matrix determining f'_{lm} coefficients from the coefficients f_{lm} . Those in turn can be transformed into c'_{lm} and s'_{lm} using Eq.(12). The net effect of such averaging on the coefficients c'_{lm} and s'_{lm} is their suppression by factors $\sim \frac{1}{l}$ and $\sim e^{-\frac{m^2\sigma_\beta^2}{4}}$ compared to c_{lm} and s_{lm} . In other words, averaging is equivalent to applying a low-pass filter suppressing spherical harmonics with large values of l and m in the function $f(\theta, \phi)$.

DEER spectra

The individual meaningful DEER spectrum components $\tilde{S}_{2k,0}(\omega)$, shown in Eq.(28) of the main text can be calculated as follows:

$$\tilde{S}_{2k,0}(\omega) = \frac{1}{2} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta [\delta(\omega - \omega_{\text{dd}}(r, \theta)) + \delta(\omega - \omega_{\text{dd}}(r, \theta))] Y_{2k,0}(\theta, \phi), \quad (16)$$

where $\delta(\omega)$ is a Dirac δ -function. After integration over angle ϕ , replacement of $Y_{2k,0}(\theta, \phi)$ with its representation using associated Legendre polynomials, shown in Eq.(5), and taking into account the symmetry with respect to a sign change in Eq.(7), $\tilde{S}_{2k,0}(\omega)$ becomes:

$$\tilde{S}_{2k,0}(\omega) = 2\pi\sqrt{4k+1} \int_{\theta=\pi/2}^{\theta=0} [\delta(\omega - \omega_{\text{dd},0}(3\cos^2\theta - 1)) + \delta(\omega - \omega_{\text{dd},0}(3\cos^2\theta - 1))] P_{2k}^0(\cos\theta) d\cos\theta. \quad (17)$$

After replacing the variable $z = \cos^2 \theta$, a careful integration leads to the following result:

$$\tilde{S}_{2k,0}(\omega) = \begin{cases} \pi\sqrt{4k+1} \sqrt{\frac{3}{1-\frac{\omega}{\omega_{dd,0}}}} P_{2k}^0 \left(\sqrt{\frac{1-\frac{\omega}{\omega_{dd,0}}}{3}} \right), & \text{for } -2 \leq \frac{\omega}{\omega_{dd,0}} < -1; \\ \pi\sqrt{4k+1} \sqrt{\frac{3}{1+\frac{\omega}{\omega_{dd,0}}}} P_{2k}^0 \left(\sqrt{\frac{1+\frac{\omega}{\omega_{dd,0}}}{3}} \right) + \sqrt{\frac{3}{1-\frac{\omega}{\omega_{dd,0}}}} P_{2k}^0 \left(\sqrt{\frac{1-\frac{\omega}{\omega_{dd,0}}}{3}} \right), & \text{for } -1 \leq \frac{\omega}{\omega_{dd,0}} \leq -1; \\ \pi\sqrt{4k+1} \sqrt{\frac{3}{1+\frac{\omega}{\omega_{dd,0}}}} P_{2k}^0 \left(\sqrt{\frac{1+\frac{\omega}{\omega_{dd,0}}}{3}} \right), & \text{for } 1 < \frac{\omega}{\omega_{dd,0}} \leq 2. \end{cases} \quad (18)$$

References

- [1] D. A. Varshalovich, A. N. Moskalev, V. K. Khersonskii, Quantum theory of angular momentum : irreducible tensors, spherical harmonics, vector coupling coefficients, 3nj symbols , Singapore ; Philadelphia : World Scientific Pub., 1986.
- [2] M. A. Wieczorek, M. Meschede, SHTools: Tools for Working with Spherical Harmonics, Geochemistry, Geophysics, Geosystems 19 (8) (2018) 2574–2592. doi:10.1029/2018GC007529.

Supplementary figures

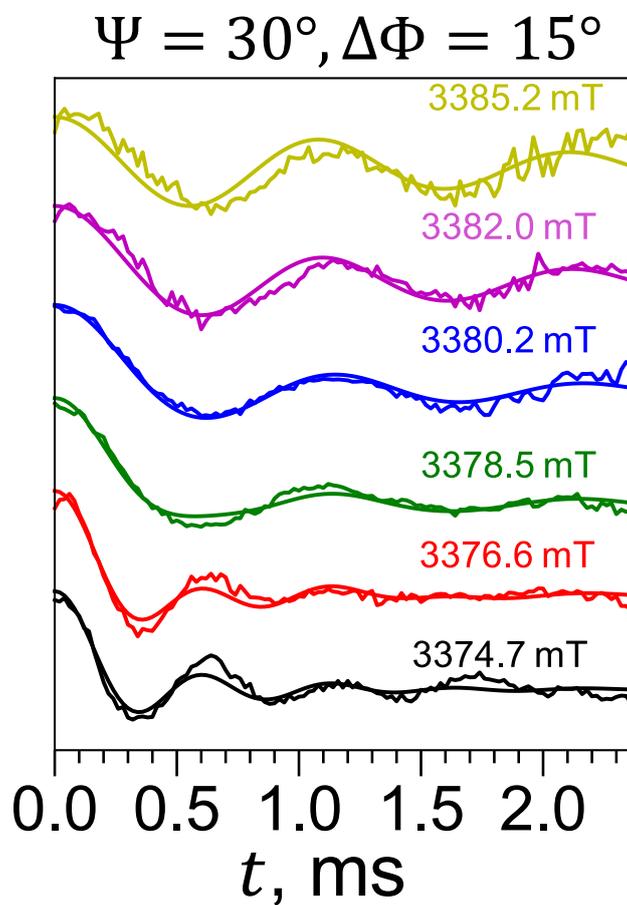


Figure S1: Experimental background corrected and normalized DEER traces of the nitroxide biradical, shown also in the main text Fig. 3c, overlaid with the fits obtained using parameters $\Psi = 30^\circ$ and $\Delta\Phi = 15^\circ$ as explained in the main text. The DEER traces were measured at the magnetic fields: 3374.7 mT (*black*), 3376.6 mT (*red*), 3378.5 mT (*green*), 3380.2 mT (*blue*), 3382.0 mT (*purple*), 3385.2 mT (*yellow*).

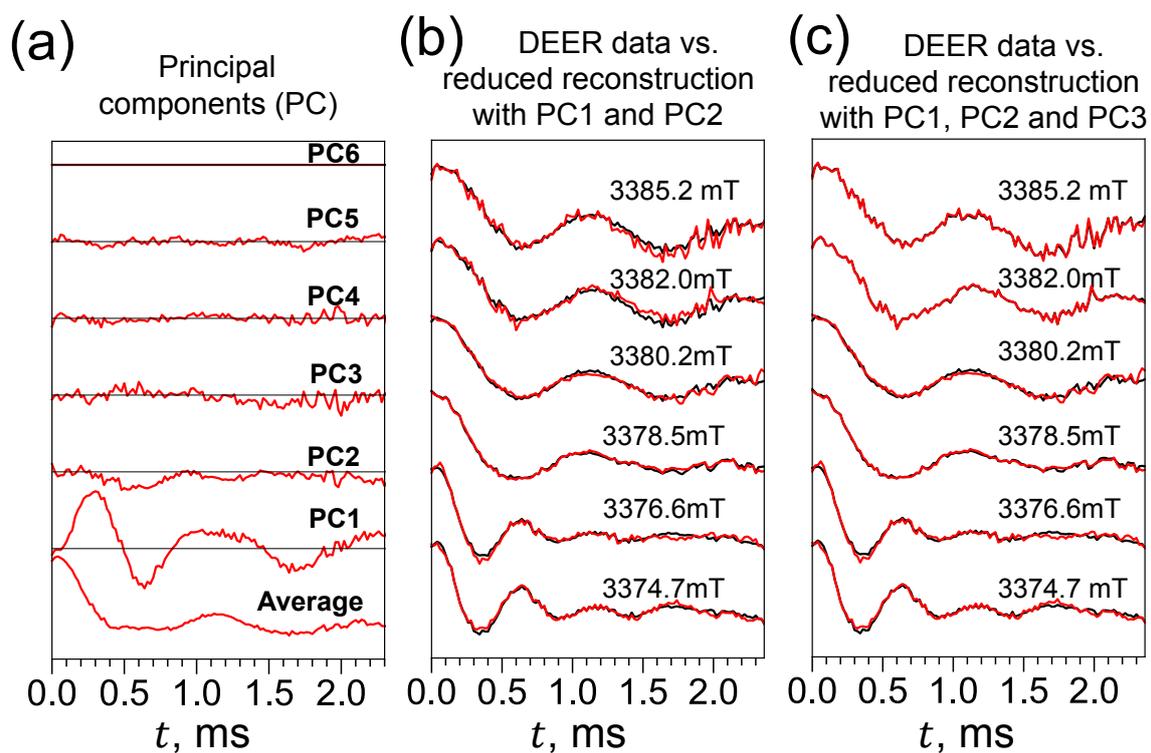


Figure S2: Principal components analysis of the nitroxide biradical DEER data. (a) The average trace and the principal components (*red*) shown on the same scale, with the zero offsets lines marked in *black*. (b) The experimental biradical DEER traces (*red*) overlaid with reconstructed traces (*black*), obtained using only the first two principal components. (c) The experimental DEER traces (*red*) overlaid with reconstructed traces (*black*), obtained using only the first three principal components.

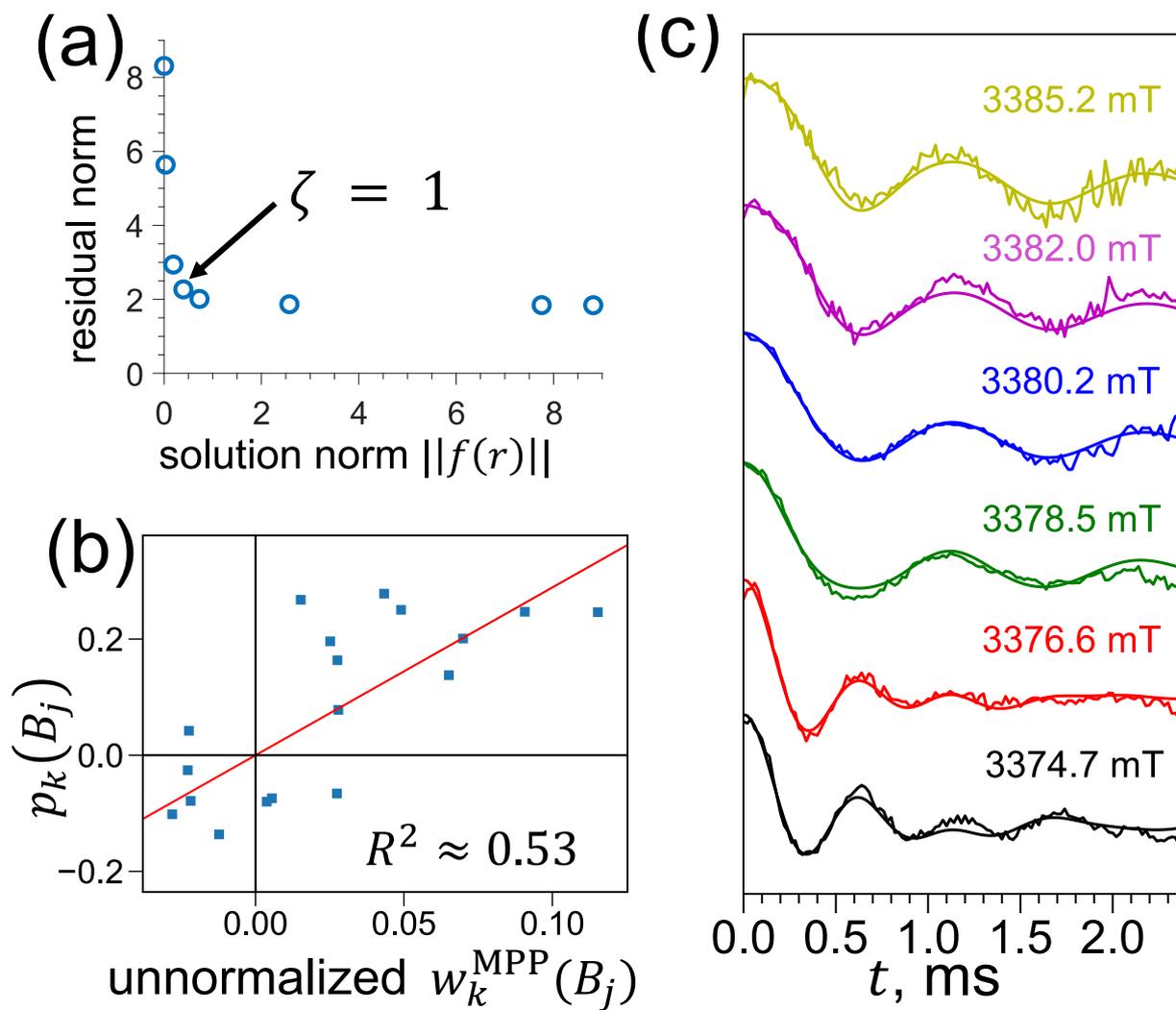


Figure S3: Details for processing the nitroxide biradical DEER data using the Tikhonov regularization-based iterative processing algorithm explained in the main text. (a) L-curve: residual norm against solution norm calculated for various values of regularization parameter ζ . (b) Experimental background corrected and normalized DEER traces recorded at various magnetic fields and overlaid with fits obtained using the iterative processing algorithm. The DEER traces were measured at the magnetic fields: 3374.7 mT (*black*), 3376.6 mT (*red*), 3378.5 mT (*green*), 3380.2 mT (*blue*), 3382.0 mT (*purple*), 3385.2 mT (*yellow*). (c) orientation encoding coefficients $p_k(B_j)$ against MPP weights $w_k^{\text{MPP}}(B_j)$ calculated using a simple geometric model in Fig. 3a. Red line represents a linear fit with zero-intercept.

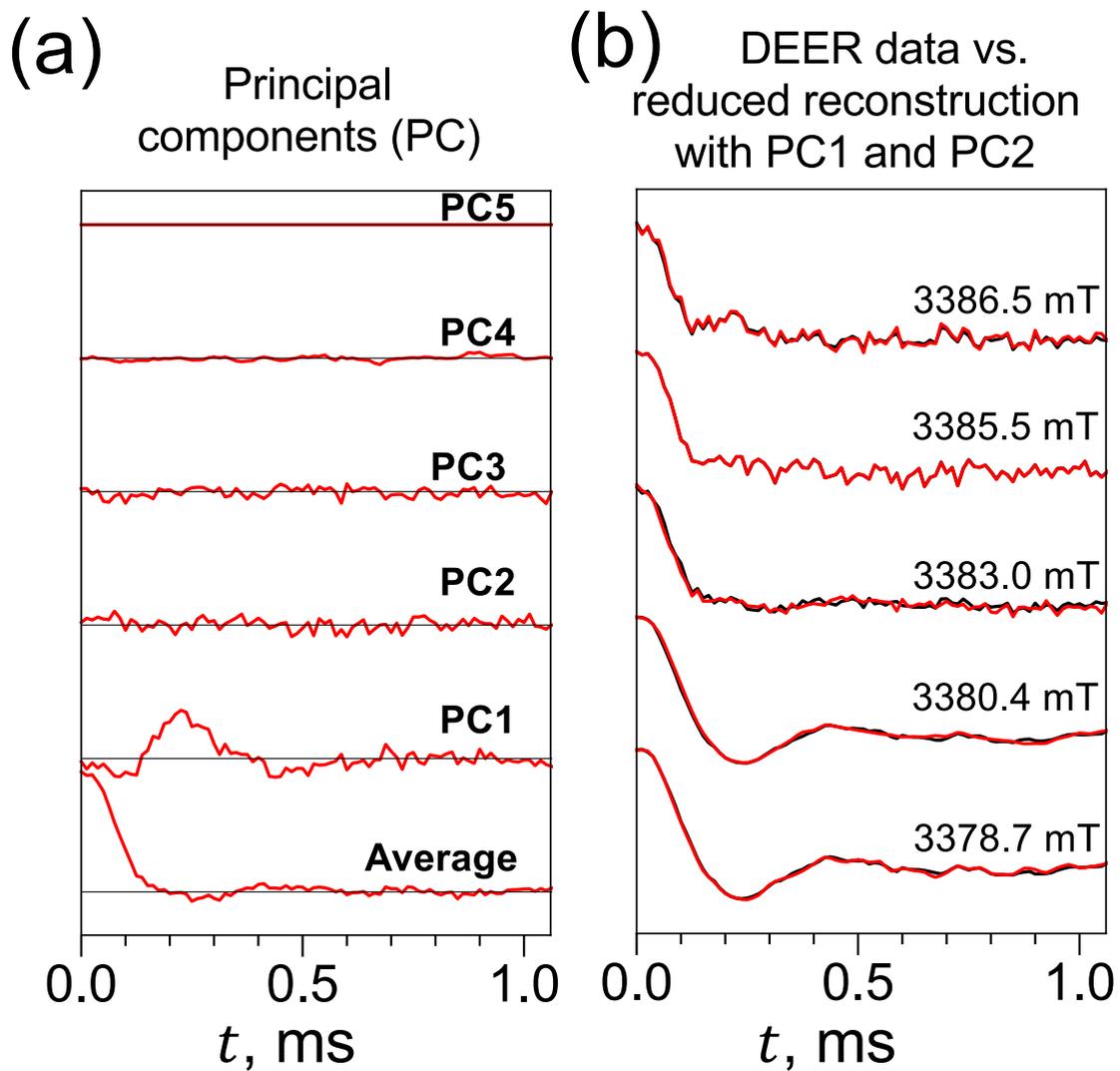


Figure S4: Principal components analysis of the spin-labelled τ_C14 protein DEER data. (a) The average trace and the principal components (*red*) shown on the same scale, with the zero offsets lines marked in *black*. (b) The experimental biradical DEER traces (*red*) overlaid with reconstructed traces (*black*), obtained using only the first two principal components.

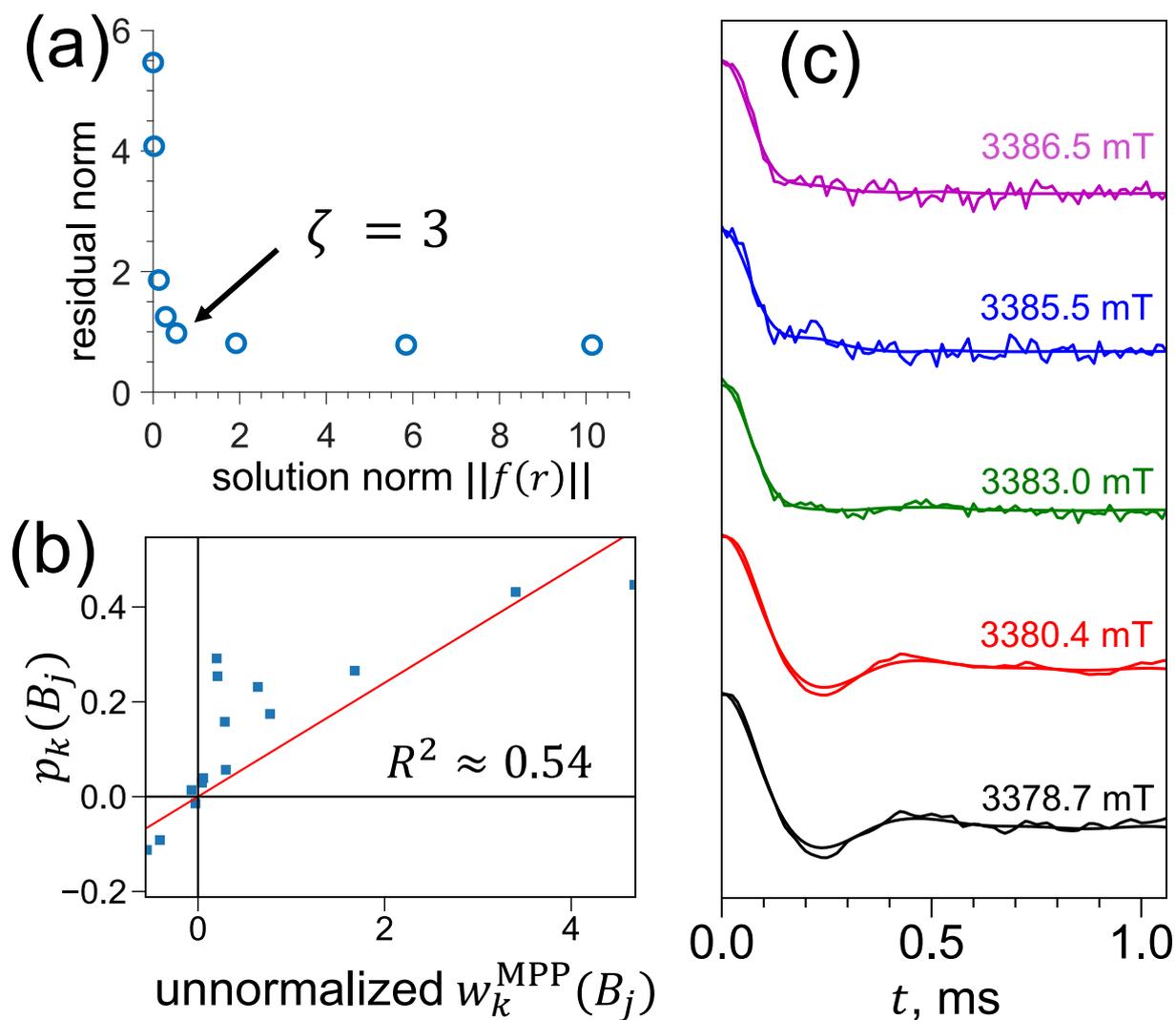


Figure S5: Details for processing the spin-labelled τ_C14 protein DEER data using the Tikhonov regularization-based iterative processing algorithm explained in the main text. (a) L-curve: residual norm versus solution norm calculated for various values of regularization parameter ζ . (b) Experimental background corrected and normalized DEER traces recorded at various magnetic fields and overlaid with fits obtained using the iterative processing algorithm. The DEER traces were measured at the magnetic fields: 3378.7 mT (*black*), 3380.4 mT (*red*), 3383.0 mT (*green*), 3385.5 mT (*blue*), 3386.5 (*purple*). (c) orientation encoding coefficients $p_k(B_j)$ against MPP weights $w_k^{\text{MPP}}(B_j)$ calculated using a simple geometric model in Fig. 3a. Red line represents a linear fit with zero-intercept.

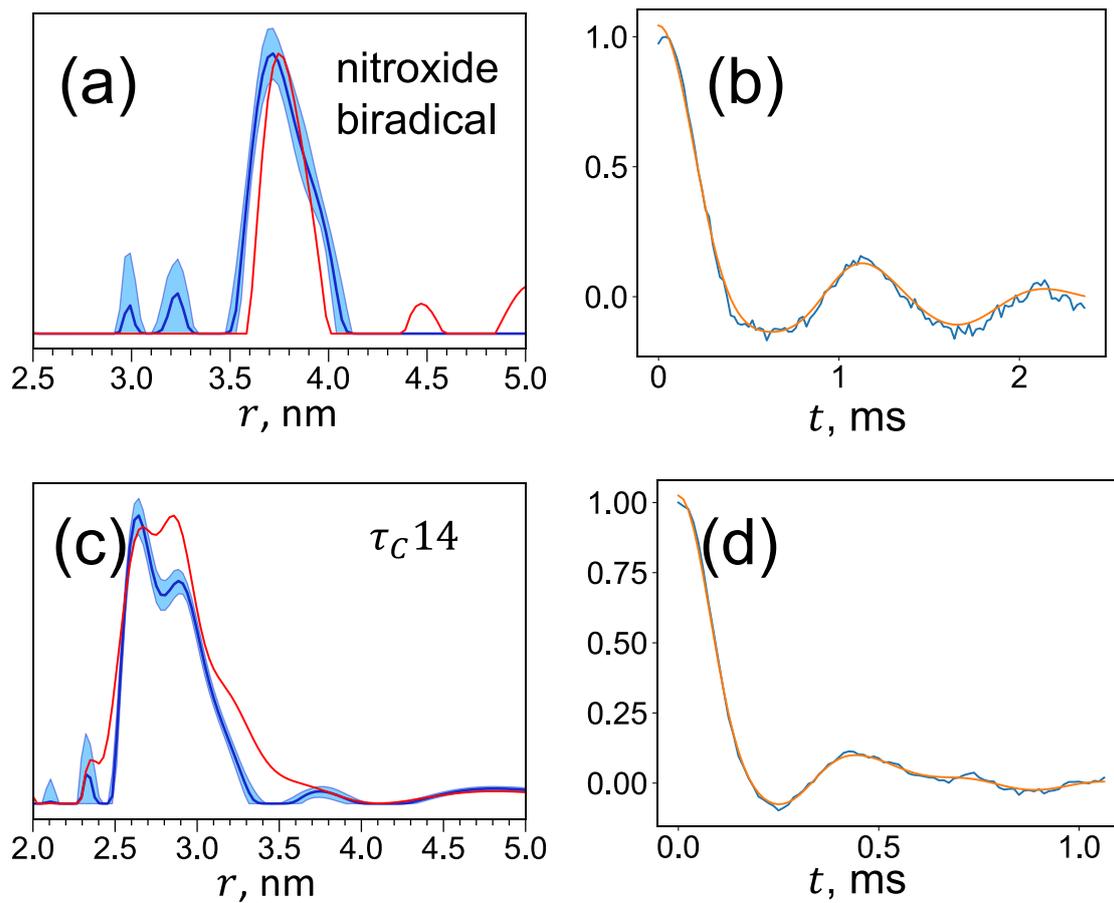


Figure S6: Distance distributions (*blue*) obtained by the Tikhonov regularization of a synthetic trace, produced by a summation of individual DEER traces of (a) the nitroxide biradical and (c) the τ_{C14} protein. The *light blue* shaded areas in (a) and (c) correspond to the 95% confidence bounds obtained using a bootstrap method with a resampling by replacement. The *red* line in (a) and (c) corresponds to the best fit results for the nitroxide and τ_{C14} protein respectively, and is also shown in Figs. 5 and 7 of the main text. (b) and (d) The synthetic DEER traces (*blue*) obtained by a summation of the nitroxide biradical and τ_{C14} protein datasets respectively. *Yellow* line is a DEER trace calculated using distributions in panels (a) and (c).