



Cite this: *Phys. Chem. Chem. Phys.*,
2022, 24, 22645

Dipolar pathways in multi-spin and multi-dimensional dipolar EPR spectroscopy†

Luis Fábregas-Ibáñez,^a Valerie Mertens,^a Irina Ritsch,^a Tona von Hagens,^a Stefan Stoll^b and Gunnar Jeschke^a

Dipolar electron paramagnetic resonance (EPR) experiments, such as double electron–electron resonance (DEER), measure distributions of nanometer-scale distances between unpaired electrons, which provide valuable information for structural characterization of proteins and other macromolecular systems. We present an extension to our previously published general model based on dipolar pathways valid for multi-dimensional dipolar EPR experiments with more than two spin-1/2 labels. We examine the 4-pulse DEER and TRIER experiments in terms of dipolar pathways and show experimental results confirming the theoretical predictions. This extension to the dipolar pathways model allows the analysis of previously challenging datasets and the extraction of multivariate distance distributions.

Received 5th July 2022,
Accepted 7th September 2022

DOI: 10.1039/d2cp03048a

rsc.li/pccp

1 Introduction

Measuring distributions of nanometer-scale distances between unpaired electrons in naturally occurring or engineered paramagnetic systems is the primary focus of dipolar EPR spectroscopy.^{1,2} These distributions contain valuable information for the structural characterization of proteins and other macromolecular systems,^{3,4} even those lacking long-range order in a sample.⁵ For this reason, dipolar EPR spectroscopy provides unique information about structurally disordered or highly complex systems.

The most broadly applied type of dipolar EPR experiment is double electron–electron resonance (DEER, also known as pulsed electron double resonance, PELDOR). Since the introduction of the 3-pulse DEER experiment,¹ the 4-pulse DEER experiment has become the most common,^{6–8} but many other variations have been developed.^{9–20} All these experiments use a sequence of microwave pulses to generate a spin echo and record the echo amplitude as a function of pulse timings. From the resulting modulated spin echo signal, the underlying spin–spin distance distribution can be inferred, provided one has an adequate theoretical model to describe the experimental data.

In a previous work,²¹ we introduced a uniform theoretical model of dipolar EPR spectroscopy experiments on two-spin systems based on dipolar pathways.^{9,10,22–24} In general, a

dipolar signal is a linear combination of oscillatory contributions from individual dipolar pathways. While single-pathway models have found widespread application in the past, we showed that a multi-pathway model is essential to accurately describe certain experimental dipolar signals that are typically encountered in experiments such as 4-pulse and 5-pulse DEER in two-spin systems.²¹ Multi-spin systems, *i.e.* systems with more than two spins, have been studied by DEER spectroscopy in the past both in model compounds^{25–27} and protein systems.^{28–35} Multi-spin systems can be used to simultaneously infer distances between multiple pairs of spins and help elucidate geometries that would be challenging based solely on single distances obtained from two-spin systems.

In this work, we extend our previous multi-pathway model to multi-spin systems and multi-dimensional experiments, where one or more time delays are shifted separately. This considerably expands upon the previously published models for multi-spin systems measured by 4-pulse DEER.^{26,27} The new model can be used to extract multivariate distance distributions from multi-spin systems and allows the analysis of complex multi-dimensional experiments such as the triple electron resonance (TRIER) experiment,¹⁵ which had previously remained a challenge to model and analyze.

This paper is structured as follows. In Section 2, we present the theoretical expansion of the current dipolar pathways model to the general multi-spin case. In Section 3, we describe the 4-pulse DEER and TRIER experiments applied to multi-spin systems in terms of dipolar pathways. In Section 4, we analyze a selection of experimental 4-pulse DEER and TRIER datasets acquired on different oligoradical and protein systems to validate our theoretical model and predictions.

^a ETH Zurich, Laboratory of Physical Chemistry, Vladimir-Prelog-Weg 2, Zurich, Switzerland

^b Department of Chemistry, University of Washington, Seattle, WA 98195, Washington, USA

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d2cp03048a>



2 Theoretical model

A specific dipolar EPR experiment is characterized by a pulse sequence with N microwave pulses and the time intervals T_n separating them. These timings are chosen such that a spin echo is generated at a time T_N after the N -th pulse. The index $n = (1, 2, \dots, N)$ indicates a time interval. To report on the coupling between electron spins within a molecular system, a modulation of the spin echo is recorded by varying one or more intervals T_n simultaneously or separately. Writing the sequence of delays as a vector $\mathbf{T} = (T_1, T_2, \dots, T_N)$, the sequence timings of a D -dimensional experiment are given by^{21,36}

$$\mathbf{T}(\mathbf{t}) = \mathbf{T}(t_1, \dots, t_D) = \mathbf{T}_0 + \sum_{d=1}^D t_d \mathbf{D}_d \quad (1)$$

where \mathbf{T}_0 is the vector of initial delays and $\mathbf{t} = (t_1, t_2, \dots, t_D)$ are the time coordinates that are varied during the experiment (see Fig. 1). The unitless vectors Δ_d specify the incrementation schemes of the corresponding time coordinates. The elements of Δ_d indicate whether the interval T_n is incremented ($\Delta_{d,n} = 1$), decremented ($\Delta_{d,n} = -1$) or kept constant ($\Delta_{d,n} = 0$) along the d -th dimension. If $\sum_n \Delta_{d,n} = 0$ for all d , then the overall length of the pulse sequence is independent of all t_d , thus representing a constant-length experiment.

In a highly diluted sample, where the multi-spin-labeled molecules or proteins in the sample are, on average, far from each other, the spin echo amplitude $V(\mathbf{t})$ is described by the following product model as originally introduced by Milov *et al.*¹

$$V(\mathbf{t}) = V_0 \cdot V_{\text{intra}}(\mathbf{t}) \cdot V_{\text{inter}}(\mathbf{t}) \quad (2)$$

Here, the prefactor V_0 is the echo amplitude if all dipolar interactions were refocused, and $V_{\text{intra}}(\mathbf{t})$ and $V_{\text{inter}}(\mathbf{t})$ are the echo modulations from intramolecular and intermolecular spin-spin interactions, respectively (also referred to as foreground and background). The intramolecular contribution depends on the distribution of intramolecular spin-spin distances. The intermolecular contribution depends on the overall spin concentration in the sample and their spatial distribution.

In the rest of this section, we derive the expression for the modulation of the spin echo arising from an isolated multi-spin system with fixed geometry and orientation for an arbitrary multi-dimensional pulse sequence, and from there, the expressions for $V_{\text{intra}}(\mathbf{t})$ and $V_{\text{inter}}(\mathbf{t})$ for a disordered dilute sample.

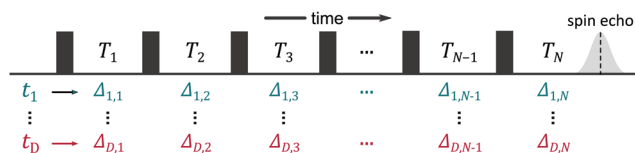


Fig. 1 General multi-dimensional pulse sequence. The sequence consists of N pulses (black boxes) separated by the time intervals T_n . The spin echo is detected after the last time interval T_N . Several time intervals can be incremented/decremented (as indicated by the vectors $\Delta_d = (\Delta_{d,1}, \dots, \Delta_{d,N})$) according to the different experimental time axes t_d . The number of different time axes t_d defines the dimensionality D of the experiment.

We limit all our theoretical treatments to $S = 1/2$ spins under weak dipolar coupling and in the high-temperature regime.^{21,37} We assume the absence of large g -factor anisotropies,^{38,39} exchange coupling,^{40–42} spectral diffusion,⁴³ multi-spin multi-quantum coherences,^{44,45} dynamical decoupling effects, and conformer-dependent relaxation rates.⁴⁶ If a sample or dipolar EPR experiment does not satisfy these assumptions, the model is not applicable.

2.1 Multi-spin dipolar echo modulation

We start by considering a group of M electron spins. This system has $Q = M(M - 1)/2$ spin pairs (see Fig. 2) with the corresponding dipolar frequencies

$$\omega_{\text{dip},q} = \frac{\mu_0}{4\pi} \frac{g^2 \mu_B^2}{\hbar} \frac{1}{r_q^3} (1 - 3 \cos^2 \theta_q) = \frac{\mathcal{D}}{r_q^3} (1 - 3 \cos^2 \theta_q) \quad (3)$$

where θ_q is the angle between the q -th interspin vector and the applied magnetic field, r_q is the q -th interspin distance and the constant \mathcal{D} combines the g -values of the two electrons (close to the value of the free electron, g_e), the square of the Bohr magneton μ_B , the magnetic constant μ_0 and the reduced Planck constant \hbar . In addition to the dipole-dipole interaction, the spins have resonance frequency offsets Ω_q .

For an N -pulse sequence, when starting at thermal equilibrium (*i.e.* any of the 2^M states are populated) there are $4^{M(N+1)}$ possible single-element transfer pathways, of which only those ending in single-quantum coherence after the last pulse are detectable as an FID or echo. In addition, only those lead to a stationary spin echo whose resonance offset phase is refocused at time T_N after the last pulse, independent of \mathbf{t} .²¹ This allows us to describe the echo modulation as a function of \mathbf{t} solely in terms of the dipolar phase of the individual pathways. Many of these single-element transfer pathways result in the same dipolar phase evolution for each of the dipolar interactions,²¹ and they can be collected into K distinct multi-spin dipolar pathways. The total spin echo amplitude E is a sum of the

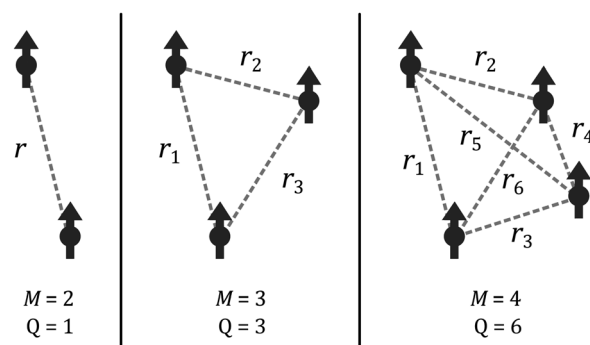


Fig. 2 Schematic representation of dipolar interactions in different multi-spin systems. The spins are shown as black arrows and all possible interspin distances r_q are shown as dashed lines for a two-spin system ($M = 2$, $Q = 1$), a three-spin system ($M = 3$, $Q = 3$), and a four-spin system ($M = 4$, $Q = 6$).



contributions E_k from all these multi-spin dipolar pathways

$$E(\mathbf{t}) = \sum_{k=1}^K E_k(\mathbf{t}) = \sum_{k=1}^K A_k \exp(-i\phi_{\text{dip},k}(\mathbf{t})) \quad (4)$$

Here, $\phi_{\text{dip},k}(\mathbf{t})$ is the net dipolar phase accumulated along pathway k , i.e. the sum of the accumulated dipolar phases over all spin pairs q and all free-evolution intervals n ,

$$\phi_{\text{dip},k}(\mathbf{t}) = \sum_{n=1}^N \sum_{q=1}^Q s_{k,q,n} \cdot \frac{1}{2} \omega_{\text{dip},q} T_n(\mathbf{t}) \quad (5)$$

where the dipolar phase accumulation factor $s_{k,q,n}$ indicates whether the dipolar phase due to spin pair q increases (+1), remains constant (0), or decreases (−1) during interval n for pathway k . The complex-valued dipolar pathway amplitude A_k is the product of the transfer amplitudes of the individual single-element transfer pathways contained in dipolar pathway k . Pathway amplitudes are determined by the configuration of the individual pulses such as pulse length, shape, phase and microwave power.^{9,47,48} As in our previous work, we use the scaling convention $\sum |A_k| = 1$.

To separate the structure of the spin system from the details of the pulse sequence, we rewrite eqn (5) as

$$\phi_{\text{dip},k}(\mathbf{t}) = \sum_{q=1}^Q \omega_{\text{dip},q} \mathbf{t}_{\text{dip},k,q}(\mathbf{t}) \quad (6)$$

where $\mathbf{t}_{\text{dip},k,q}(\mathbf{t})$ is the D -dimensional vector of effective dipolar evolution times of spin pair q for dipolar pathway k given by

$$\mathbf{t}_{\text{dip},k,q}(\mathbf{t}) = \frac{1}{2} \sum_{n=1}^N s_{k,q,n} T_n(\mathbf{t}) = \frac{1}{2} \mathbf{s}_{k,q} \cdot \mathbf{T}(\mathbf{t}) \quad (7)$$

The vectors $\mathbf{s}_{k,q}$ collect the sequence of $s_{k,q,n}$ for a given pathway and spin pair, constituting a pair dipolar pathway. This is illustrated in Fig. 3 for a three-spin system. The set of all

pair dipolar accumulation factors $\mathbf{s}_{k,q}$ for a given k gives the multi-spin dipolar accumulation factors S_k .

We now consider the modulation of a single dipolar pathway echo contribution $E_k(\mathbf{t})$ along an arbitrary dimension d . A dipolar pathway echo contribution $E_k(\mathbf{t})$ is modulated along the time coordinate t_d with respect to pair q only if $t_{\text{dip},k,q,d}(\mathbf{t})$ is not constant. If $t_{\text{dip},k,q,d}(\mathbf{t})$ is constant for all q , the whole pathway contribution $E_k(\mathbf{t})$ will be constant along dimension d . If the contribution is modulated, the spins along a pair dipolar pathway refocus at a time when the associated dipolar phase is zero, or, equivalently, when the effective dipolar evolution time is zero ($\mathbf{t}_{\text{dip},k,q}(\mathbf{t}) = 0$). Thus, each modulated pair dipolar pathway has a characteristic D -dimensional refocusing time point $\mathbf{t}_{\text{ref},k,q}$ with elements given by

$$t_{\text{ref},k,q,d} = -\frac{\mathbf{s}_{k,q} \cdot \mathbf{T}_0}{\mathbf{s}_{k,q} \cdot \Delta_d} \quad (8)$$

where the dipolar phase of the pathway is zero and its contribution to the total echo amplitude is at its largest.

With eqn (1) and (7), the above equation gives

$$\mathbf{t}_{\text{dip},k,q}(\mathbf{t}) = \delta_{k,q} \odot (\mathbf{t} - \mathbf{t}_{\text{ref},k,q}) \quad (9)$$

where denotes the Hadamard (element-wise) product and the D -dimensional prefactor vector $\delta_{k,q}$ has elements

$$\delta_{k,q,d} = \frac{1}{2} \mathbf{s}_{k,q} \cdot \Delta_d \quad (10)$$

which characterize whether the contribution from spin pair q for pathway k is modulated ($\delta_{k,q,d} \neq 0$) or unmodulated ($\delta_{k,q,d} = 0$) along the d -th time dimension. The magnitude of the prefactor also characterizes whether the pathway evolves as a harmonic ($|\delta_{k,q,d}| > 1$), subharmonic ($|\delta_{k,q,d}| < 1$), or equal ($|\delta_{k,q,d}| = 1$) of the d -th experimental time vector t_d . Finally, we rewrite the total echo amplitude as

$$E(\mathbf{t}) = \sum_{k=1}^K A_k \exp\left(-i \sum_{q=1}^Q \omega_{\text{dip},q} \mathbf{t}_{\text{dip},k,q}(\mathbf{t})\right) \quad (11)$$

The full set of multi-spin pathways contains pairs k and k' with opposite-sign phase accumulation factors $S_k = (\mathbf{s}_{k,1}, \dots, \mathbf{s}_{k,Q})$ and $S_{k'} = (-\mathbf{s}_{k,1}, \dots, -\mathbf{s}_{k,Q})$. They accumulate a dipolar phase of opposite sign $\phi_{\text{dip},k}(\mathbf{t}) = -\phi_{\text{dip},k'}(\mathbf{t})$ and, in the high-temperature limit, they present the same amplitude $A_k = A_{k'}$. This results in their echo contributions $E_k(\mathbf{t})$ and $E_{k'}(\mathbf{t})$ being complex-conjugate, leading to an echo that is composed purely by real-valued contributions

$$E(\mathbf{t}) = \sum_{k=1}^K 2A_k \cos\left(\sum_{q=1}^Q \omega_{\text{dip},q} \mathbf{t}_{\text{dip},k,q}(\mathbf{t})\right) \quad (12)$$

where k now runs over the subset of multi-spin dipolar pathways pairs. For simplicity, from now on we will refer to the combination of S_k and $-S_k$ as just S_k .

Now, if all spins in the system have the same spectral distribution (a common situation, e.g., when all paramagnetic centers are nitroxides) then any of the spins is affected by a given pulse in the sequence with the same probability. As a

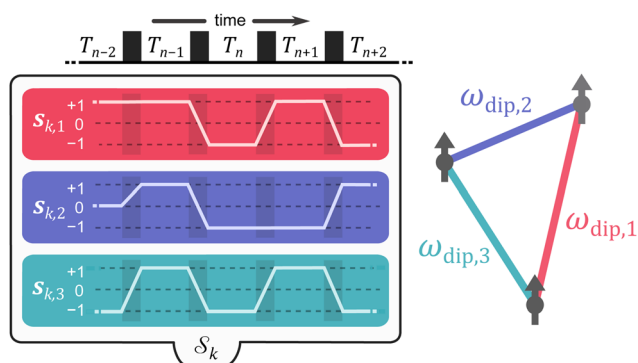


Fig. 3 Schematic representation of a multi-spin dipolar pathway, exemplified in a three-spin system. During a sequence of pulses (shown as black boxes) and time intervals T_n , a multi-spin dipolar pathway with multi-spin dipolar phase accumulation factors S_k is composed of Q pair dipolar phase accumulation factors $\mathbf{s}_{k,q}$ (shown as white lines) that describe whether the dipolar phase corresponding to the q -th dipolar interaction frequency $\omega_{\text{dip},q}$ increases (+1), decreases (−1), or remains constant (0). Each pair dipolar pathway is colored as its corresponding dipolar interaction.



consequence, all multi-spin dipolar pathways that are related by permutations of spin pairs have identical amplitude. This assumption allows us to account for fewer dipolar pathway amplitudes, considerably reducing the complexity of the model. For multi-spin systems consisting of different spin species (e.g., a combination of nitroxide and trityl radicals), this approximation may not be valid and the full set of dipolar pathways must be considered separately. Likewise, if the sites have strongly different labelling efficiencies, the approximation becomes invalid.

The model derived thus far is exact within our set of assumptions, yet its complexity is cumbersome for an intuitive interpretation or for its practical application. Furthermore, the number of detectable and modulated dipolar pathways increases significantly for each additional spin in a molecule and for each pulse in an experiment.

To simplify the model, we re-organize the multi-spin dipolar pathways based on their modulation of the spin echo. All multi-spin pathways with the same modulated ($|\delta_{k,q,d}| > 0$) pair dipolar pathways but different unmodulated ($|\delta_{k,q,d}| = 0$) pair dipolar pathways will make the same contribution to the echo modulation. We can combine them into what we call a set of unique multi-spin dipolar pathways with multi-spin accumulation factors \tilde{S}_k , which consist of combinations of modulated pair dipolar pathways and a single unmodulated pathway that combines all unmodulated pair dipolar pathways. The amplitudes of these unique multi-spin pathways are given by the sum of the amplitudes of the multi-spin pathways described by it. This is illustrated in Fig. 4 for a three-spin system, where five different multi-spin dipolar pathways (① to ⑤) with identical modulated pair dipolar pathways and different unmodulated pair dipolar pathways can be combined into a unique multi-spin dipolar pathway \tilde{S}_1 with amplitude $\tilde{A}_1 = A_1 + \dots + A_5$.

This description provides us with an opportunity to simplify the model even further. The set of unique multi-spin dipolar pathways can be constructed from a generic unmodulated pair

dipolar pathway and a significantly reduced set of modulated pair dipolar pathways $\mathbf{s}_p = (\mathbf{s}_{p,1}, \dots, \mathbf{s}_{p,N})$ and $-\mathbf{s}_p$ (see Fig. 4). Each of these pair dipolar pathways will have an associated refocusing time $\mathbf{t}_{\text{ref},p} = (\mathbf{t}_{\text{ref},p,1}, \dots, \mathbf{t}_{\text{ref},p,D})$, and factor $\delta_p = (\delta_{p,1}, \dots, \delta_{p,D})$. The pair dipolar pathways in a multi-spin system for a given pulse sequence consist at least of the modulated pathways found in a two-spin system for that pulse sequence. Each unique modulated multi-spin dipolar pathway can in turn be constructed by combination of $Q' < Q$ modulated pair dipolar pathways. This also significantly simplifies the amplitudes of the unique multi-spin dipolar pathways. The amplitude of unique multi-spin pathway k composed of Q' modulated pair dipolar pathways can be well-approximately re-parameterized in terms of the amplitudes λ_p associated with the corresponding pair dipolar pathway and the amplitude λ_u associated with the unmodulated pair dipolar pathway

$$\tilde{A}_k \approx (Q - Q')! \lambda_u^{Q-Q'} \prod_p \lambda_p \quad (13)$$

This approximation assumes the amplitude of a pair dipolar pathway to be independent of any other pairs in the system. This is again illustrated in Fig. 4 for three-pulse DEER on a three-spin system, where we can see that the accumulation factors $\tilde{S}_1 \dots \tilde{S}_8$ of the unique multi-spin pathways ① to ⑧ can be fully constructed from one unmodulated pair dipolar pathway and two modulated pair dipolar pathways' accumulation factors \mathbf{s}_1 and \mathbf{s}_2 .

The formalism above has two important consequences. First, it allows us to describe the dipolar signals just in terms of the subset of pair dipolar pathways with accumulation factors \mathbf{s}_p instead of the (significantly larger) full set of multi-spin pathways with accumulation factors \tilde{S}_k . Second, due to the normalization of the pathway amplitudes and the permutation of the multi-spin pathways it follows that all $|\lambda_p| \leq 1/Q!$. Therefore, the amplitude of a multi-spin pathway \tilde{A}_k decreases

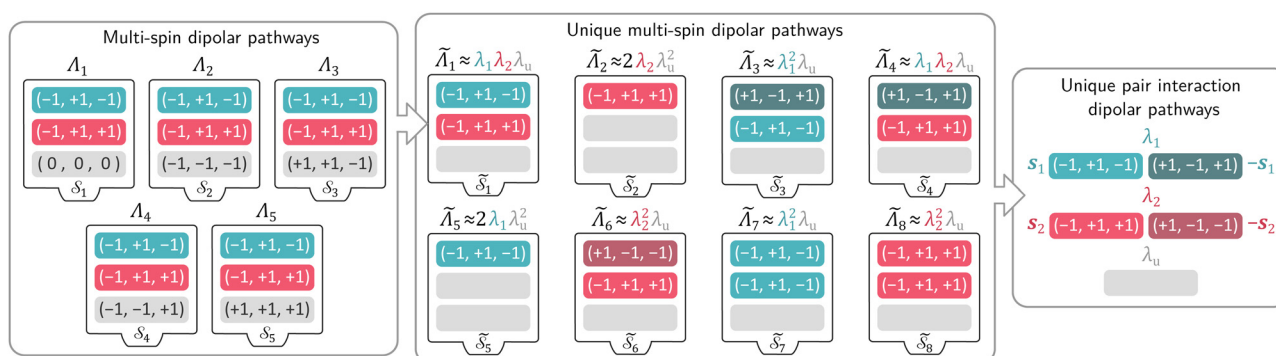


Fig. 4 Schematic representation of the sequential simplification of the multi-spin dipolar pathways models. An example subset of five different multi-spin dipolar pathways with multi-spin accumulation factors \tilde{S}_k for the one-dimensional three-pulse DEER experiment on a three-spin system is shown as boxes containing the individual modulated ($|\delta_p| > 0$, colored) and unmodulated ($|\delta_p| = 0$, grey) pair dipolar pathways $\mathbf{s}_{k,q}$, which can be collected into a unique multi-spin dipolar pathway 1 with multi-spin accumulation factors \tilde{S}_1 where all unmodulated pair pathways are represented by a generic pair dipolar pathway. Other sets of multi-spin dipolar pathways can be collected into unique multi-spin dipolar pathways 2 to 8 with accumulation factors $\tilde{S}_2 \dots \tilde{S}_8$. The set of eight different unique multi-spin dipolar pathways can be described by one unmodulated pathway (grey) and two unique pair dipolar pathways with pair accumulation factors \mathbf{s}_1 (turquoise) and \mathbf{s}_2 (magenta). The amplitudes of the different pathways are given on top of each box.

quickly with the number Q' of modulated pair dipolar pathways contained in it. The presence of three or more modulated pair dipolar pathways results in negligible amplitudes. Therefore, any multi-spin system can be approximately described by just its two-spin and three-spin interactions (*i.e.* $Q' < 3$) without relevant loss of accuracy. In this decisive simplification, we can thus approximate the multi-spin echo modulation as

$$E(\mathbf{t}) \approx E_0 + \sum_p \sum_q^Q (Q-1)! \lambda_u^2 \lambda_p \cos(\omega_{\text{dip},q} \mathbf{t}_{\text{dip},p}) \\ + \sum_{p \leq p'} \sum_{\substack{q \neq q' \\ c_{qq'}=1}}^Q (Q-2)! \lambda_u \lambda_p \lambda_{p'} [\cos(\omega_{\text{dip},q} \mathbf{t}_{\text{dip},p} + \omega_{\text{dip},q'} \mathbf{t}_{\text{dip},p'}) \\ + \cos(\omega_{\text{dip},q} \mathbf{t}_{\text{dip},p} - \omega_{\text{dip},q'} \mathbf{t}_{\text{dip},p'})] \quad (14)$$

where E_0 collects the constant echo amplitude contributed by all unmodulated dipolar pathways, the first term describes the two-spin contributions, and the second term the three-spin contributions to the echo modulation. The two terms of the three-spin contributions represent the different combinations of pathways with \mathbf{s}_p and $-\mathbf{s}_p$, respectively, which have identical amplitudes (see Fig. 4) within our approximation. In both terms, the first sum accounts for all combinations of Q' pair dipolar pathways, and the second sum accounts for the permutation symmetry of the spin system. Note that for the three-spin contributions we only need to sum over the cases where the q -th and q' -th interspin vectors are joined by a common spin ($c_{qq'} = 1$). If they are disjoint ($c_{qq'} = 0$) they will not contribute to the signal. The expression above can be further simplified to

$$E(\mathbf{t}) \approx E_0 + \sum_p \sum_q^Q (Q-1)! \lambda_u^2 \lambda_p \cos(\omega_{\text{dip},q} \mathbf{t}_{\text{dip},p}) \\ + \sum_{p \leq p'} \sum_{\substack{q \neq q' \\ c_{qq'}=1}}^Q 2(Q-2)! \lambda_u \lambda_p \lambda_{p'} \cos(\omega_{\text{dip},q} \mathbf{t}_{\text{dip},p}) \cos(\omega_{\text{dip},q'} \mathbf{t}_{\text{dip},p'}) \quad (15)$$

This approximate model represents a significant extension to the similar model previously published by Jeschke *et al.*²⁶ Also note that this definition of the term three-spin contributions is broader than the one used by Pribitzer *et al.*,¹⁵ where it was used for signal contributions modulated by two dipolar frequencies along a single dimension.

2.2 Intramolecular contributions

To obtain an expression for the intramolecular contribution $V_{\text{intra}}(\mathbf{t})$ in a disordered sample, such as a frozen solution, we average the echo amplitude $E(\mathbf{t})$ over the uniform orientational distribution of the spin cluster. If the molecule's conformations are distributed, we additionally average over the conformational distribution. Different choices of internal coordinates are possible to represent conformations. Here, we use the set of

inter-spin distances $\mathbf{r} = (r_1, \dots, r_Q)$, which forms a set of non-redundant internal coordinates for $M \leq 4$ (for larger systems, or in the presence of symmetry, other representations are more efficient). The conformational distribution is described by the Q -variate distance distribution $P(\mathbf{r})$, which is only defined over the region of \mathbf{r} that fulfills the generalized triangle inequality

$$r_v \leq \sum_{u \neq v} r_u \quad \text{for } v = 1, \dots, Q \quad (16)$$

Together, the orientational and conformational averaging is

$$V_{\text{intra}}(\mathbf{t}) = \int d\mathbf{r} P(\mathbf{r}) \int d\cos\theta d\phi E(\mathbf{t}) \quad (17)$$

where θ and ϕ describe the orientation of the spin cluster. Inserting $E(\mathbf{t})$ from eqn (15) into this gives

$$V_{\text{intra}}(\mathbf{t}) \approx A_0 + \sum_p \sum_q^Q (Q-1)! \lambda_u^2 \lambda_p \int_0^\infty dr_q P_q(r_q) K_0^{(2)}(\mathbf{t}_{\text{dip},p}, r_q) \\ + \sum_{p \leq p'} \sum_{\substack{q \neq q' \\ c_{qq'}=1}}^Q 2(Q-2)! \lambda_u \lambda_p \lambda_{p'} \int_0^\infty dr_q \int_0^\infty dr_{q'} \int_0^\infty dr_{q''} \\ \times P_{qq'q''}(r_q, r_{q'}, r_{q''}) K_0^{(3)}(\mathbf{t}_{\text{dip},p}, \mathbf{t}_{\text{dip},p'}, r_q, r_{q'}, r_{q''}) \quad (18)$$

where A_0 is the unmodulated contribution that results from integration of E_0 , $r_{q''}$ is the distance of the interspin vector that forms a joint triangle with the q -th and q' -th interspin vectors, and $P_q(r_q)$ and $P_{qq'q''}(r_q, r_{q'}, r_{q''})$ are the univariate and trivariate marginal distance distributions

$$P_q(r_q) = \int_0^\infty \dots \int_0^\infty P(\mathbf{r}) \prod_{i \neq q} dr_i \quad (19)$$

$$P_{qq'q''}(r_q, r_{q'}, r_{q''}) = \int_0^\infty \dots \int_0^\infty P(\mathbf{r}) \prod_{i \neq q, q', q''} dr_i \quad (20)$$

The two-spin dipolar kernel in eqn (18) is given by

$$K_0^{(2)}(\mathbf{t}_{\text{dip},p}, r_q) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^\pi d\theta_q \sin\theta_q \cos(\zeta_{p,q}(1 - 3\cos^2\theta_q)) \quad (21)$$

where $\zeta_{p,q} = \mathcal{D}\mathbf{t}_{\text{dip},p}/r_q^3$. This can be evaluated analytically²¹ to give

$$K_0^{(2)}(\mathbf{t}_{\text{dip},p}, r_q) = \frac{1}{z_{p,q}} [F_C(z_{p,q}) \cos\zeta_{p,q} + F_S(z_{p,q}) \sin\zeta_{p,q}] \quad (22)$$

where $z_{p,q} = \sqrt{6\zeta_{p,q}/\pi}$ and F_C and F_S are the cosine and sine Fresnel integral functions, respectively. The three-spin dipolar kernel in eqn (18) is given by^{26,27,49}

$$K_0^{(3)}(\mathbf{t}_{\text{dip},p}, \mathbf{t}_{\text{dip},p'}, r_q, r_{q'}, r_{q''}) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^\pi d\theta_q \sin\theta_q \\ \times \cos(\zeta_{p,q}(1 - 3\cos^2\theta_q)) \cos(\zeta_{p',q'}(1 - 3\cos^2\theta_{q'})) \quad (23)$$



where

$$\cos \theta_{q'} = \sin \chi_{qq'} \cos \varphi \sin \theta_q + \cos \chi_{qq'} \cos \theta_q \quad (24)$$

and $\chi_{qq'}$ is the angle between the interspin vectors of the q' -th and q -th spin pair, given by the law of cosines

$$\chi_{qq'} = \arccos\left(\frac{r_q^2 + r_{q'}^2 - r_{qq'}^2}{2r_q r_{q'}}\right) \quad (25)$$

The three-spin dipolar kernel has no closed-form solution and must be evaluated numerically.

It is important to note that while the two-spin intramolecular contributions depend solely on the individual univariate marginal distance distributions, the three-spin contributions depend on the trivariate distance distributions, which encode the correlation information between triads of interspin distances. These correlations provide a new layer of valuable information for the structural characterization of the molecular systems under study.

2.3 Intermolecular contributions

To derive an expression for the echo modulation V_{inter} due to intermolecular dipolar interactions we must consider the full system with N_s spins residing in different spin clusters. We assume a random uniform distribution of spin clusters in the sample^{1,50} and disregard any volume exclusion effects.⁵¹ Under a high dilution of the spin clusters we can reduce the problem to pair interactions between spins and neglect any multi-spin intermolecular interactions between the spins residing in different spin clusters. The total echo arising from these interactions is the product over the individual pair contributions to the echo

$$E_{N_s}(\mathbf{t}) = \prod_{b=1}^{N_s} E_b(\mathbf{t}) \quad (26)$$

where the index b runs over all spins residing in other spin clusters. The total intermolecular contribution can be computed by averaging over all possible orientations and configurations of the spins in other molecules in space. The averaging situation is analogous to the two-spin case for which we already provided a detailed derivation in ref. 21. Inserting $E(\mathbf{t})$ from eqn (15) into the expression above gives

$$\begin{aligned} V_{\text{inter}}(\mathbf{t}) &= V_{\text{inter},0}(\mathbf{t}) \prod_p^P \prod_q^Q \exp(-\kappa c_s (Q-1)! \lambda_u^2 \lambda_p |\mathbf{t}_{\text{dip},p}|) \\ &\times \prod_{p \leq p'}^P \prod_{\substack{q \neq q' \\ c_{qq'}=1}}^Q \exp(-\kappa c_s 2(Q-2)! \lambda_u \lambda_p \lambda_{p'} |\mathbf{t}_{\text{dip},p} + \mathbf{t}_{\text{dip},p'}|) \end{aligned} \quad (27)$$

where c_s is the spin concentration, $\kappa = 8\pi^2/(9\sqrt{3})\mathcal{D}$, and $V_{\text{inter},0}$ is a constant scaling factor that can be merged into V_0 . The intermolecular contribution is thus a product of all the exponential decays arising from each two-spin and three-spin interaction in the system. The different refocusing times of the individual dipolar pathways can result in kinks of the total

intermolecular contribution, which were already observed and studied in our previous work.²¹

3 Specific experiments

In this section, we discuss the 4-pulse DEER and TRIER experiments in terms of their pair dipolar pathways to understand the different possible contributions to their respective dipolar signals. In the ESI,† we provide a general script to compute all possible modulated dipolar pathways contributing to the dipolar signal (within the assumptions of this work) for arbitrary pulse sequences and multi-spin systems.

3.1 4-Pulse DEER

The 4-pulse DEER experiment is the most common dipolar EPR experiment, it is one-dimensional ($\mathbf{t} = t$), and its dipolar pathways in two-spin systems have been discussed extensively.^{21–23} Dipolar signals acquired by this experiment for multi-spin systems are composed of contributions from different multi-spin pathways that can be constructed from the four unique pair dipolar pathways ① to ④ summarized in Fig. 5. The main contribution ① refocuses at $t = \tau_1$ and the contributions ② and ③ refocus at the edges of the accessible time range. ② is often referred to as the “2+1” contribution.^{4,48}

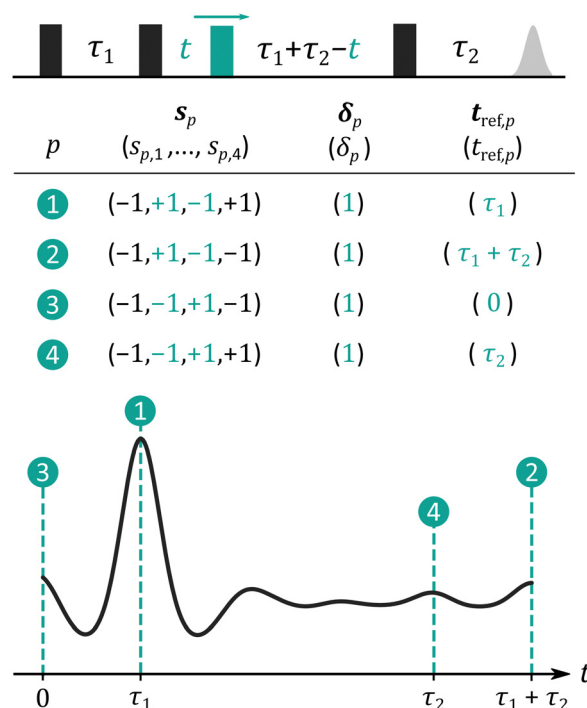


Fig. 5 Pair dipolar pathways of the multi-spin 4-pulse DEER experiment. The 4-pulse DEER pulse sequence is shown on top, with every probe pulse represented as a black box and the pump pulse as a green box. The table contains a subset of the most commonly encountered pair dipolar pathways s_p along with their harmonics δ_p and refocusing times $t_{\text{ref},p}$. The pathways are ordered in decreasing estimated amplitude for commonly reported experimental conditions. The refocusing times are illustrated as dashed turquoise lines on top of a schematic 4-pulse DEER dipolar signal shown as a black line.



In the absence of any pulse excitation overlap and using π -pulses of high inversion efficiency, the 4-pulse DEER signal consists of just pathway ①. In the presence of excitation overlap, pathways ② to ④ become significant.²¹

In terms of three-spin contributions, we can expect non-negligible contributions from pathways where one or two dipolar interactions evolve according to pair pathway ①. Since the 4-pulse DEER is set up such that ① has the largest amplitude, its three-spin counterpart will also have a significant amplitude. For this same reason, we can expect very small or negligible amplitude from three-spin contributions arising from combinations of other pathways such as ② or ③. However, if their amplitude becomes significant due to considerable pulse excitation overlap or broadband or shaped pulses, their three-spin contributions can also become non-negligible. For example, the four-pulse version of the nDEER experiment,¹¹ which uses such broadband pulses and records ② as its main contribution, would have a significant three-spin contribution from pathways where two dipolar interactions evolve according to ② instead of ①.

3.2 TRIER

The triple electron resonance (TRIER) experiment is a two-dimensional form of a six-pulse DEER experiment where the two pump pulses are shifted according to $\mathbf{t} = (t_1, t_2)$ (see Fig. 6).^{15,52} The experiment was originally designed for the study of frequency-domain correlations between the different dipolar frequencies in a multi-spin system. Its complexity is reflected in the large number of multi-spin pathways that can contribute to its signal. The different multi-spin pathways that can be constructed from the 24 pair pathways are summarized in Fig. 6 ordered by decreasing expected amplitude based on typical experimental conditions.

The TRIER experiment aims to record the contributions arising from the t_1 -modulated pathway ① refocusing at $t_1 = \tau_1$ and the t_2 -modulated pathway ② refocusing at $t_2 = \tau_3$. The TRIER model proposed originally^{15,52} contained both the two-spin contributions (modulated along each time dimension) as well the three-spin contributions (modulated along both dimensions) arising from pathways ① and ②. Under typical experimental conditions, both pathways will provide the largest contributions to the TRIER signal. Thus, we can expect the most intensive signal at a time $\mathbf{t} = (\tau_1, \tau_3)$.

In addition, the TRIER signal can easily present contributions from pathways where one of the probe pulses fails to invert the dipolar phase (due to insufficient inversion efficiency or pulse excitation bandwidth overlap). This happens due to imperfections of the fourth pulse in pathways ③ and ④, of the last pulse in pathways ⑤ and ⑥, or of the second pulse in ⑦ and ⑧. Pathways ③ and ④ refocus at $t_1 = \tau_2 + \tau_3$ and $t_2 = \tau_1 + \tau_2$, respectively, at the edges of the measurable signal, analogous to the ③ contribution in the 4-pulse DEER experiment. Pathways ⑥ and ⑦ refocus at $t_1 = 0$ and $t_2 = 0$, respectively, at the very beginning of the experimental signal, again analogous to the ③ pathway of the 4-pulse DEER experiment. If the same pulse configuration is used for all probe pulses, one can expect all these additional pathways to contribute with essentially the same amplitude.

Similarly, if the inversion efficiency of the probe pulses is not perfect, pathways where two probe pulses fail to invert the dipolar phase can have non-negligible amplitudes. Pathways ⑨, and ⑩ occur due to the second and fourth, pathways ⑪ and ⑫ due to the fourth and last, pathways ⑬ and ⑭ due to the second and last pulses failing to invert the dipolar phase. Similarly, pathways ⑮ and ⑯ arise from the failure of all three probe pulses.

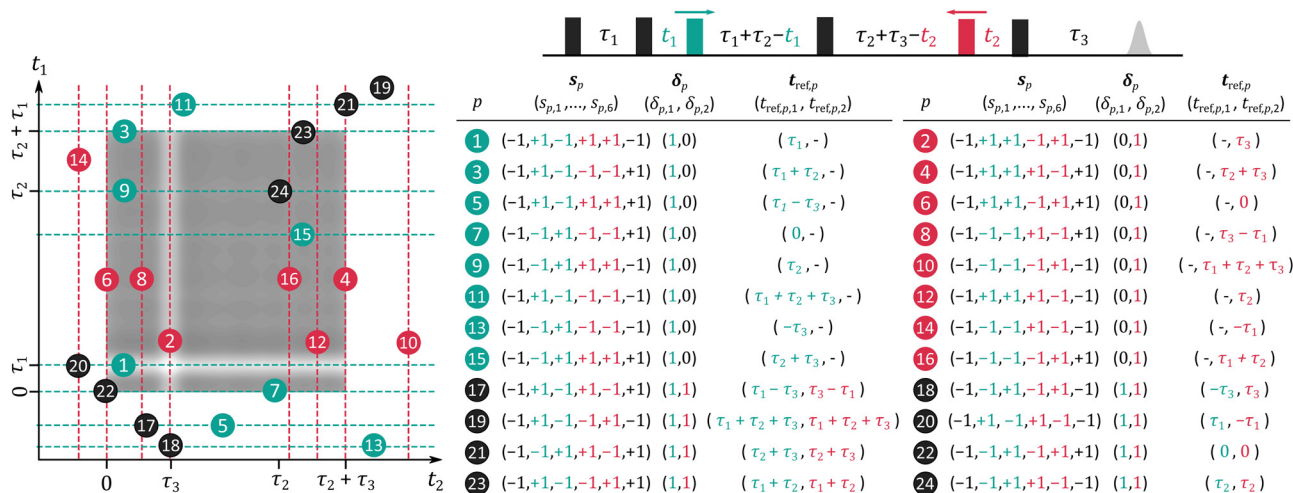


Fig. 6 Pair dipolar pathways of the multi-spin TRIER experiment. The TRIER pulse sequence is shown on top, every probe pulse represented as a black box, the t_1 -shifted pump pulse as a green box, and the t_2 -shifted pump pulse as a red box. The table lists all possible pair dipolar pathways s_p along with their harmonics δ_p and refocusing times $t_{ref,p}$ modulated along t_1 (green), t_2 (red), and both (black). The pathways are ordered in decreasing estimated amplitude for commonly reported experimental conditions. In the left panel, the refocusing times of the one-dimensional modulated contributions are illustrated as dashed green and red lines, and the two-dimensional modulated contributions as black circles on top of a schematic TRIER dipolar signal shown as greyscale contours.

All these pathways are modulated either along t_1 or t_2 . This is a consequence of the three-frequency nature of the experiment. The use of different frequencies for the two pump pulses results in higher amplitudes for those pathways where one pump pulse inverts the dipolar evolution and the other one does not, as in pathways ① to ⑩. In contrast, pathways ⑪ to ⑲ correspond to those cases where both pump pulses invert the dipolar phase evolution resulting in contributions that are modulated along both dimensions.

Another important detail is that, depending on the choice of pulse delays, the refocusing times of some dipolar pathways, *e.g.*, ⑤, ⑩ to ⑭, and ⑰ to ⑲ can fall outside the accessible range of t_1 and t_2 (from $t_1 = 0$ to $\tau_2 + \tau_3$ and $t_1 = 0$ to $\tau_1 + \tau_3$). It is important that their amplitudes are minimized to ensure that their contributions are negligible. Otherwise, it can pose a serious challenge to identify their presence and analyze the data.

In terms of three-spin contributions, we can expect non-negligible contributions from pathways where one or two dipolar interactions evolve according to a combination of pair pathways ① and ②. Since the TRIER experiment is set up such that ① and ② have the largest amplitudes, their one-dimensional modulated three-spin counterparts as well as the two-dimensional modulated three-spin contribution, where one dipolar interaction evolves according to ① and the other along ②, will also have a significant amplitude. For this same reason, we can expect very small or negligible amplitude from three-spin contributions arising from combinations of other pathways. However, if their amplitude becomes significant due to large pulse excitation overlap other three-spin contributions can also become non-negligible.

4 Experimental demonstration

To validate our theoretical model and demonstrate its utility, we analyzed a series of experiments on a variety of rigid oligo(*p*-phenyleneethynylene) (oligoPPE) tri- and tetra-radicals, as well as a triply labeled protein. All datasets were analyzed using the open-source DeerLab v0.14.3 package⁵³ with Python 3.8. The experimental details, as well as all experimental datasets and the corresponding analysis scripts, are provided in the ESI.†

One of the most challenging aspects of the analysis of dipolar signals arising from multi-spin systems is the proper modeling of the multivariate distance distribution $P(\mathbf{r})$. In contrast to the two-spin scenario, where the distance distribution is defined over a one-dimensional distance domain, for a multi-spin system it is defined over a Q -dimensional distance domain. The dimensionality makes non-parametric modeling of the distribution unfeasible, requiring a parametric model for its analysis. For our analyses, based on our choice of molecular systems, we use a Q -variate Gaussian distribution

$$P(\mathbf{r}) = \frac{1}{(2\pi)^{Q/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{r} - \bar{\mathbf{r}})^T \Sigma^{-1} (\mathbf{r} - \bar{\mathbf{r}})\right) \quad (28)$$

with a covariance matrix Σ parameterized by its Cholesky decomposition

$$\Sigma = \mathbf{L}\mathbf{L}^T \text{ with } \mathbf{L} = \begin{pmatrix} \ell_{11} & 0 & 0 & \dots & 0 \\ \ell_{21} & \ell_{22} & 0 & \dots & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{Q1} & \ell_{Q2} & \ell_{Q3} & \dots & \ell_{QQ} \end{pmatrix} \quad (29)$$

The standard deviations σ_q and correlation coefficients $\rho_{qq'}$ are related to $\ell_{qq'}$ via

$$\sigma_q = \sqrt{\ell_{qq}^2 + \sum_{u=1}^{q-1} \ell_{qu}^2} \quad (30)$$

$$\rho_{qq'} = \frac{\ell_{qq'}\ell_{q'q'} + \sum_{u=1}^{q'-1} \ell_{qu}\ell_{q'u}}{\sigma_q\sigma_{q'}} \quad (31)$$

The distribution is parameterized by the mean-distance vector $\bar{\mathbf{r}} = \{\bar{r}_1, \dots, \bar{r}_Q\}$, and the set of matrix elements $\ell_{qq'}$. The parameterization in terms of $\ell_{qq'}$ ensures the positive definiteness of the covariance matrix, a necessary condition for $P(\mathbf{r})$ to be a well-defined probability density function. Additionally, all elements of $P(\mathbf{r})$ not fulfilling the generalized triangle condition in eqn (16) are set to zero. Due to this condition, the final shape of the distance distribution no longer necessarily corresponds exactly to a multivariate Gaussian distribution. Due to the large parameter space, we expect a significant risk that the optimization of the model parameters ends up in a local minimum. In order to ensure the global minimum is found, it is recommended the use of global-optimization algorithms (such as multi-start optimization^{53,54}) or the use of prior information (*e.g.*, the mean inter-spin distances) to obtain a reasonable guess of the model start values. In our case, we estimated the parameter optimization start values based on information provided by structural simulations (see below). Nevertheless, we want to note that even without prior knowledge of a system's structure, the same global minimum will exist, the only difference being that the task of finding it will be computationally costlier. Another challenging aspect is the evaluation of the three-spin dipolar kernel in eqn (23). A full numerical evaluation is unfeasible, particularly for two-dimensional experiments such as TRIER. To efficiently evaluate the spherical integrals in eqn (23), we employ the SOPHE grid.^{55–57} The distance averaging in eqn (18) over the three-spin contributions was performed by repeated random sampling from the trivariate marginal distance distributions. The two-spin contributions are computed *via* the analytical solutions of the dipolar kernel in eqn (22) and *via* numerical evaluation over the univariate marginal distance distributions.



To estimate the ground truth multivariate distance distributions of the different multi-spin systems, we performed a series of structural simulations. For the oligoPPE radicals we performed simulations based on the harmonic-segmented-chains (HSC) model,^{58,59} which was tested in the past for such shape-persistent oligomers in the context of combined DEER and fluorescence resonance energy transfer (FRET) measurements.⁶⁰ For the proteins, we simulated the multivariate distance distribution based on the protein structure using MMMx.⁶¹ All HSC and MMMx simulation scripts are provided in the ESI,[†] as well.

4.1 4-Pulse DEER

We first analyzed the X-band 4-pulse DEER data from two solutions of different rigid oligoPPE triradicals in dOTP measured and previously published in the dissertation of von Hagens.⁶² For each system, several 4-pulse DEER traces were recorded using varying pump pulse powers, resulting in different relative amplitudes of the two-spin and three-spin contributions. The first oligoPPE system (T111) has three nitroxide-terminated chains of equal length, each with a terminal nitroxide radical. The second oligoPPE system (T011) has two chains of equal length and one that is shorter. We fitted the datasets using the model given in eqn (18), taking into account the pair dipolar

pathway ① only. Due to the orientation selectivity typically encountered in such systems, the high signal-to-noise ratio of the data, and our inability to accurately model the orientation weights, we included a stretch factor in the exponential function in eqn (27) as an additional fit parameter to accommodate slight deviations from an exponential decay. All the datasets were fitted globally, with a global trivariate distance distribution and background parameters. The refocusing times of the dipolar pathways were computed from the theoretical values based on the experimental pulse delays.

Global analysis results in an accurate description of all datasets by the model as seen in Fig. 7 for T111 and in Fig. 8 for T011. All experimental signals are accurately described by the global models for both the T111 and T011 datasets. Thanks to the relatively high number of high-quality datasets with varying fractions of the two-spin and three-spin contributions, the uncertainty of the results are small (see Tables S3 and S4 in the ESI[†]). With decreasing power (increasing attenuation) of the pump pulse power, the relative amplitude of the three-spin contributions decreases. In general, the contributions from three-spin interactions are non-negligible and significantly contribute to the overall signal. Even at lower powers (≥ 7 dB attenuation) the amplitude of the three-spin contributions is still non-negligible despite its considerably reduced amplitude.

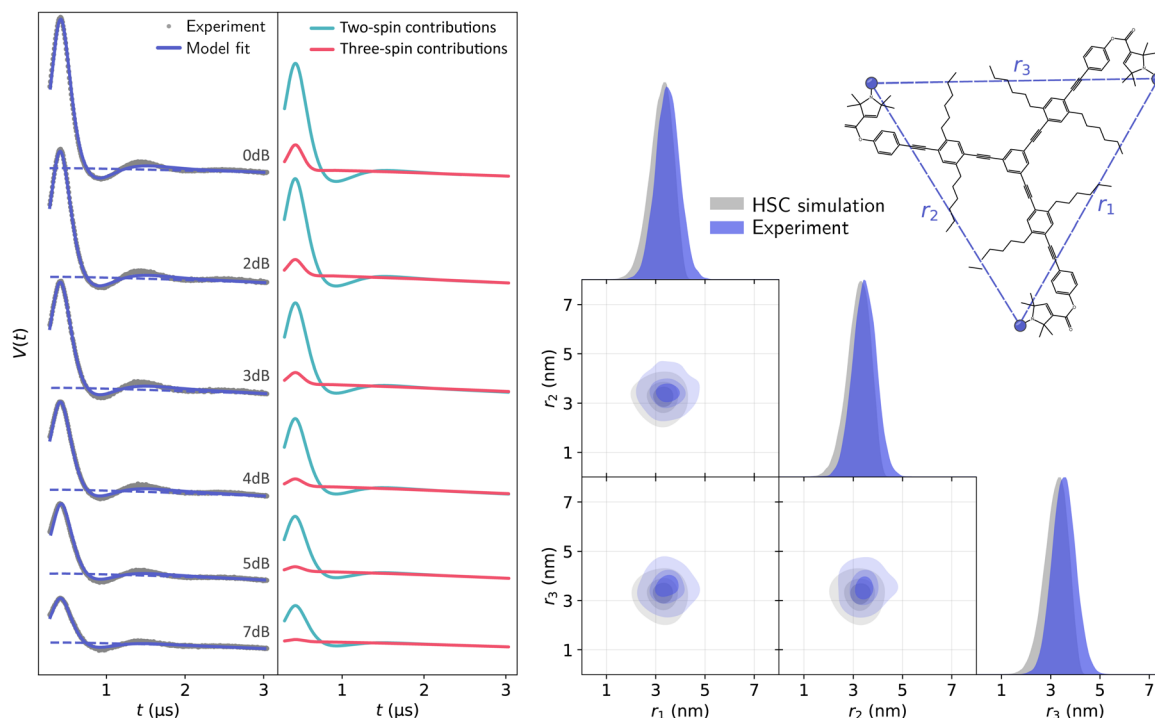


Fig. 7 Global analysis with DeerLab of a series of X-band 4-pulse DEER of oligoPPE triradical T111 (top right) acquired with $\tau_1 = 0.4 \mu\text{s}$ and $\tau_2 = 6 \mu\text{s}$, and different levels of microwave power attenuation. The experimental datasets are shown in the left panel as grey dots, along with the model fits and unmodulated contributions shown as solid and dashed blue lines, respectively. The contributions from two-spin interactions are shown as turquoise lines, and the contributions arising from three-spin interactions are shown as red lines. For clarity, only the first $3 \mu\text{s}$ out of the $6.4 \mu\text{s}$ of the recorded trace are shown (the full traces are shown in Fig. S1 in the ESI[†]). The right panel shows a globally fitted trivariate distance distribution. The univariate marginal distributions are shown as filled areas, and the bivariate marginal distributions are shown as colored contours. The fitted distribution is shown in blue, and the HSC simulation is shown in grey.

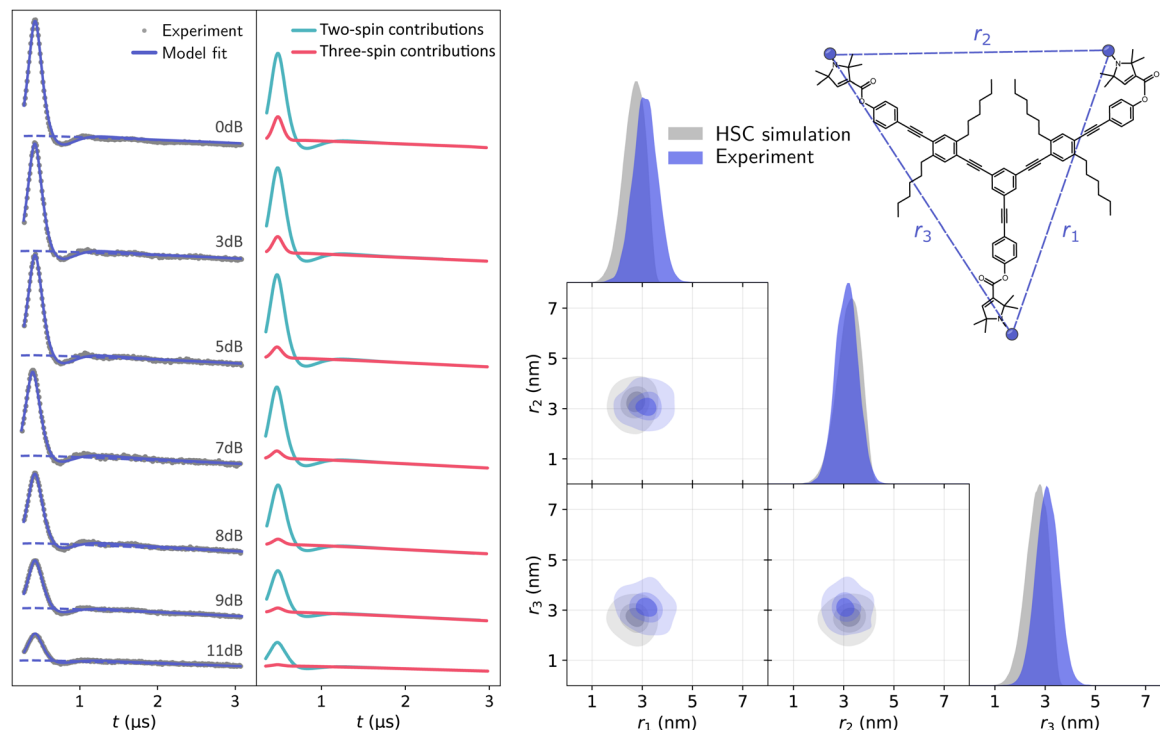


Fig. 8 Global analysis with DeerLab of a series of X-band 4-pulse DEER of oligoPPE triradical T011 (top right) acquired with $\tau_1 = 0.4 \mu\text{s}$ and $\tau_2 = 12 \mu\text{s}$, and different levels of microwave power attenuation. The experimental datasets are shown in the left panel as grey dots, along with the model fits and unmodulated contributions shown as solid and dashed blue lines, respectively. The contributions from two-spin interactions are shown as turquoise lines, and the contributions arising from three-spin interactions are shown as red lines. For clarity, only the first 3 μs out of the 12.1 μs of the recorded trace are shown (the full traces are shown in Fig. S2 in the ESI†). The right panel shows a globally fitted trivariate distance distribution. The univariate marginal distributions are shown as filled areas, and the bivariate marginal distributions are shown as colored contours. The fitted distribution is shown in blue, and the HSC simulation is shown in grey.

The globally fitted distance distribution is in good agreement with the HSC simulations for the T111 system and in reasonable agreement for the T011 systems, with some discrepancies visible in the width of the univariate marginal distributions and a slight overestimate of the shorter distance in T011. These can be caused by minor orientation selection effects, limited accuracy of the bending parameters of the HSC model at the central ring, and the presence of a certain degree of skewness in the HSC simulated distribution not captured by a Gaussian distribution.

Next, we analyzed the Q-band 4-pulse DEER data from a triple MTSL-labeled RNA polymerase protein complex (Rpo47) of Rpo4 (C36R1, G63R1) and Rpo7 (K123R1) measured and previously published in the dissertation of Ritsch.⁶³ Again, several 4-pulse DEER traces were recorded using varying pump pulse powers, resulting in different relative amplitudes of the two-spin and three-spin contributions. We fitted the dipolar signals using the same model as with the oligoPPE radicals. All datasets were globally fitted with a trivariate distance distribution and the three-dimensional homogeneous background in eqn (27). Again, the global analysis results in an accurate fit of all datasets by the model as shown in Fig. 9, where we can again see that the model globally fits all datasets accurately. As expected from visual inspection, the analysis confirms the non-negligible two-spin contribution from pathways

② and ③ at the edges of the signal. However, the three-spin contributions related to those pathways have negligible contributions. As in the previous examples, the total three-spin contribution in all cases is non-negligible and represents a significant part of the dipolar signal, particularly at early times. The fitted trivariate distance distribution agrees reasonably well with the MMMx simulations, with some minor discrepancies between the widths of the fitted distribution and those of the simulation. These discrepancies can be attributed to the higher noise levels in the datasets and to the limitations of the multivariate Gaussian model to accurately describe the distance distribution.

To examine a system with more than three spins, we analyzed another power series of X-band 4-pulse DEER measurements previously published by von Hagens⁶² on an oligoPPE tetradical in dOTP consisting of four equal polymer branches first published by Schiemann *et al.*²⁵ The presence of four spins requires a hexivariate distance distribution. As discussed above, we only need to account for two- and three-spin contributions and can safely neglect any four-spin contributions. We employed the same model for the dipolar signal as with the oligoPPE triradicals. Based on the rigidity of the system, to stabilize and make the analysis more computationally feasible, we assumed all off-diagonal elements of the Cholesky decomposition matrix in eqn (29) to be zero (effectively assuming correlation coefficients of zero between



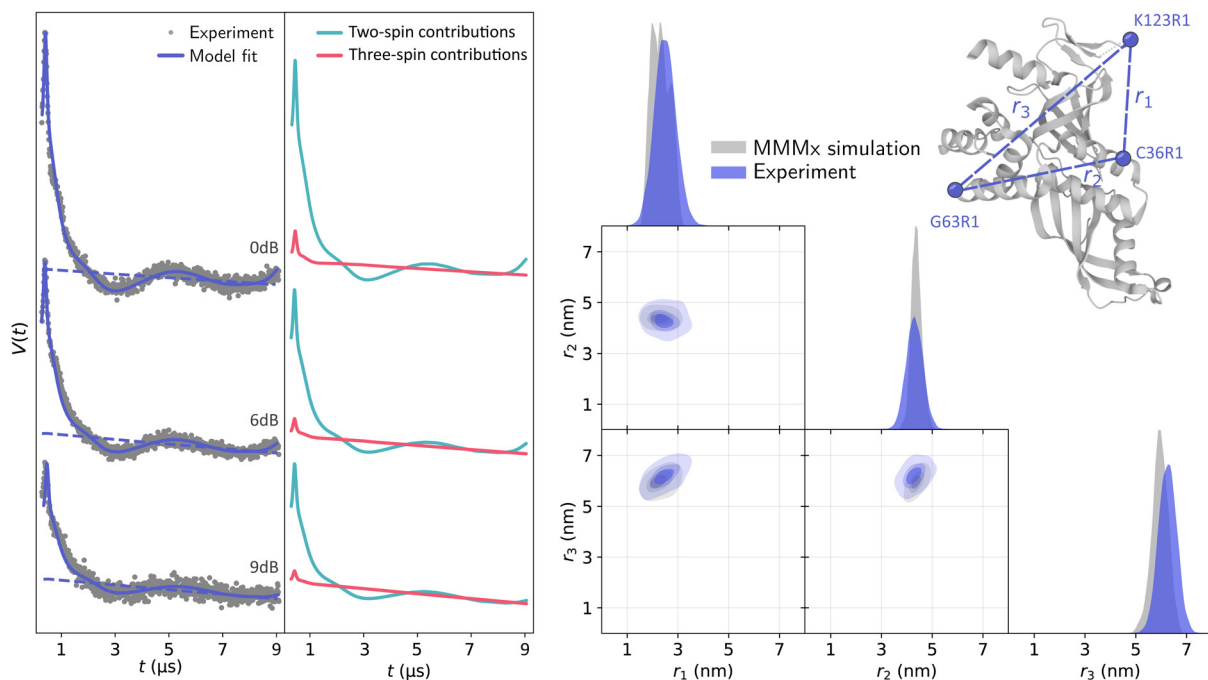


Fig. 9 Global analysis with DeerLab of a series of Q-band 4-pulse DEER of triply MTSL-labeled Rpo47 protein (ribbon model top right, PDB:1GO3) acquired with $\tau_1 = 0.4 \mu\text{s}$ and $\tau_2 = 9 \mu\text{s}$, and different levels of microwave power attenuation. The experimental datasets are shown in the left panel as grey dots, along the model fits and unmodulated contributions shown as solid and dashed blue lines, respectively. The contributions from two-spin interactions are shown as turquoise lines, and the contributions arising from three-spin interactions are shown as red lines. The right panel shows a globally fitted trivariate distance distribution. The univariate marginal distributions are shown as filled areas, and the bivariate marginal distributions are shown as colored contours. The fitted distribution is shown in blue, and the MMMx simulation is shown in grey.

different distances). The global fit results are shown in Fig. 10. The global model fits all datasets accurately at all power levels. In the fit we can see that the relative amplitudes of the total three-spin contributions are significantly larger than in the three-spin cases, particularly at early times, as expected. The globally fitted hexivariate distance distribution is in astonishingly good agreement with the HSC simulation, particularly in terms of the mean distances and with minor discrepancies in the widths of the marginal distributions. These discrepancies can have the same origins as in the oligoPPE triradicals, namely the limitations of a multivariate Gaussian distribution and of the HSC model. Note that trends in distribution width, such as the narrower distributions along the two PPE backbones and broader distributions between sites in different backbones, are reproduced in the analysis of the experimental data.

In general, we can conclude that a multi-pathway model including two- and three-spin contributions provide an accurate description of 4-pulse DEER dipolar signals arising from multi-spin systems. The multivariate distance distributions inferred from the experimental data agree with the ones obtained from structural simulations and avoid the presence of artifacts caused by the use of incomplete models, such as a two-spin model (see Fig. S4–S11 in the ESI†). Note that, while our analysis has primarily profited from a global analysis over multiple datasets acquired at different pulse power levels, this kind of analysis can be formally applied even to a single 4-pulse DEER trace, but we can expect robustness to suffer substantially.

4.2 TRIER

To validate the TRIER pathways model, we analyzed Q-band TRIER datasets acquired using a home-built spectrometer⁶⁴ on several three-spin systems. The detected echos were integrated and the two-dimensional signal was phase-corrected without any further signal modifications. The experiments were optimized with respect to minimal pulse excitation band overlap. At the same time, inversion efficiency with frequency-swept pulses in Q-band is known to be imperfect.⁴⁷ Therefore, we included pair pathways ① to ⑯ in the model, for which we expect non-negligible amplitudes due to the limited inversion efficiencies. We neglected any contributions from pathways ⑰ to ⑳ due to the well-separated three frequency bands of the different pulses. To avoid numerical issues during the data analysis we also neglect contributions from pathways ⑤, ⑧, ⑬, and ⑭ refocusing the detectable range of the signal for which we expect negligible contributions in the detected signal range. Nevertheless, the large number of parameters and complexity of the model makes the analysis of TRIER even more prone to convergence towards local minima during the fit optimization, making the use of global optimization or prior information more relevant.

First, we acquired and analyzed new TRIER data from oligoPPes T111 and T011. The results of the analyses are shown in Fig. 11 and 12. In the case of the T111 TRIER dataset, we can see an overall good agreement of the fitted model with the



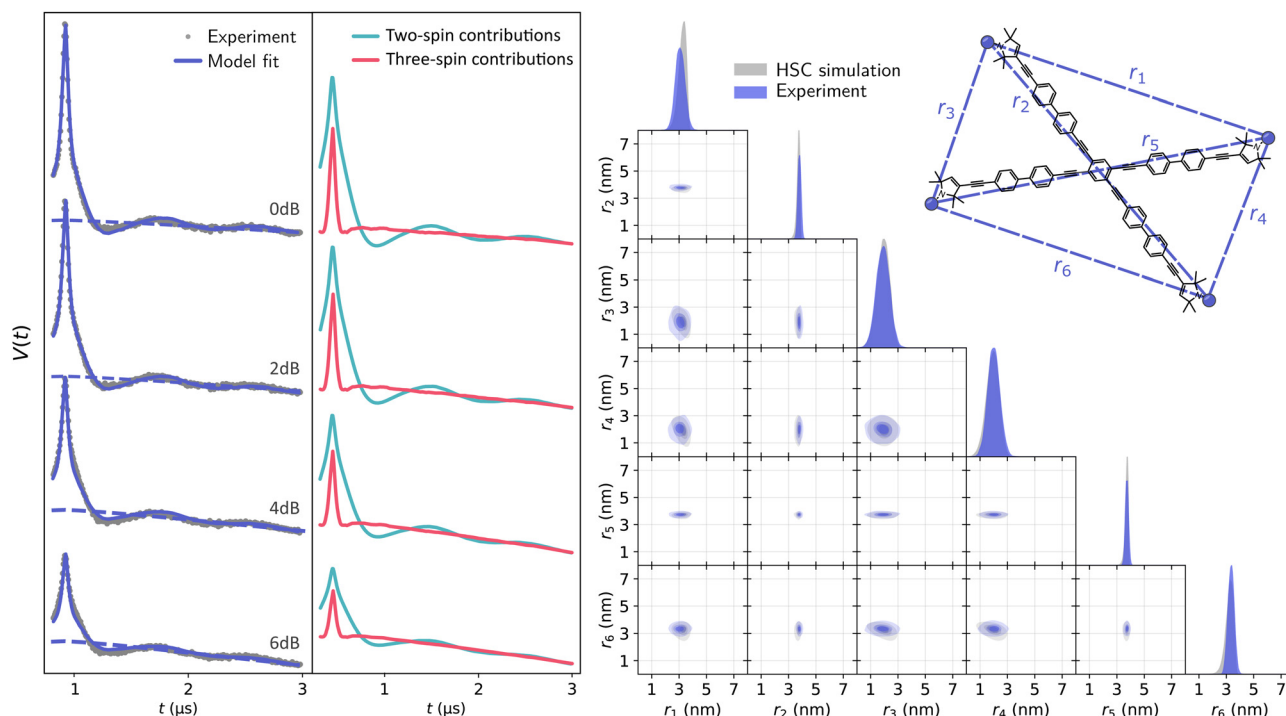


Fig. 10 Global analysis with DeerLab of a series of X-band 4-pulse DEER of an oligoPPE tetradical (top right), acquired with $\tau_1 = 0.4 \mu\text{s}$ and $\tau_2 = 8 \mu\text{s}$, and different levels of pump pulse power. The experimental datasets are shown in the left panel as grey dots and the model fits and unmodulated contributions shown as solid and dashed blue lines, respectively. The contributions from two-spin interactions are shown as turquoise lines and the contributions arising from three-spin interactions are shown as red lines. For clarity, only the first $3 \mu\text{s}$ out of the $8.1 \mu\text{s}$ of the recorded trace are shown (the full traces are shown in Fig. S3 in the ESI†). The right panel shows the globally fitted hexavariate distance distribution. The univariate marginal distributions are shown as filled areas and the bivariate marginal distributions are shown as colored contours. The fitted distribution is shown in blue and the HSC simulation is shown in grey.

experimental data. Some discrepancies can be observed in the signal, particularly in the t_1 -modulated integral at around $t_1 \approx 1\text{--}2 \mu\text{s}$, which could not be resolved by considering additional pathways and are possibly due to moderate orientation selection (which has been observed in the past for such rigid polymers in 4-pulse DEER data). The fit reveals that the main contributions are from pathways ① and ②, as expected, with other non-negligible minor contributions from ④, ⑥, ⑦, ⑩, ⑪, and ⑮. In the case of the T011 TRIER dataset, we also see a high agreement of the fitted model with the experimental data. The fitted main contributions again arise from pathways ① and ②, with additional non-negligible contributions from ③, ④, ⑨, ⑫, and ⑯.

The presence of these pathways indicates that there is a non-negligible overlap between the excitation bandwidth of the probe and pump pulses or a relatively low inversion efficiency by the probe pulses. Therefore, the presence of these additional pathways is plausible. The fitted trivariate distance distributions also show a good agreement both with the HSC simulations and with the results obtained from the analysis of the 4-pulse DEER datasets. There are again apparent differences in the fitted widths of individual distances, which could be caused by the different factors listed and discussed above.

Next, we analyzed the TRIER data from the triply labeled Rpo47 protein complex previously published by Pribitzer *et al.*⁵²

The dataset presents additional noise along the t_1 -dimension due to technical issues with the hardware of the home-built spectrometer at the time when this data was acquired. We employed the same TRIER model as with the oligoPPE radicals, including pathways ① to ⑯ and a three-dimensional homogeneous background model. The results of the analysis are shown in Fig. 13. Despite the increased noise level, the fit shows a good agreement with the experimental data. The fitted distance distribution also agrees reasonably with an MMMx simulation as well as with the results obtained from the 4-pulse DEER experiments.

In general, we see that, despite the high complexity of the TRIER dipolar signal, the multi-pathway model provides accurate descriptions of the experimental data and reasonable estimates of the multivariate distance distributions. We can also see good agreement between the results obtained from a series of 4-pulse DEER experiments with variable pump pulse flip angle and TRIER experiments. In contrast to 4-pulse DEER, the two-dimensional nature of the TRIER signal provides enough information to robustly fit the model without the need for a global-analysis approach. While TRIER data have been analyzed in the frequency domain,¹⁵ the multi-pathway model provides the first reliable approach for analyzing TRIER data directly in the distance domain. Our previous work⁵² relied on an incomplete model of the dipolar signal as well as on a data



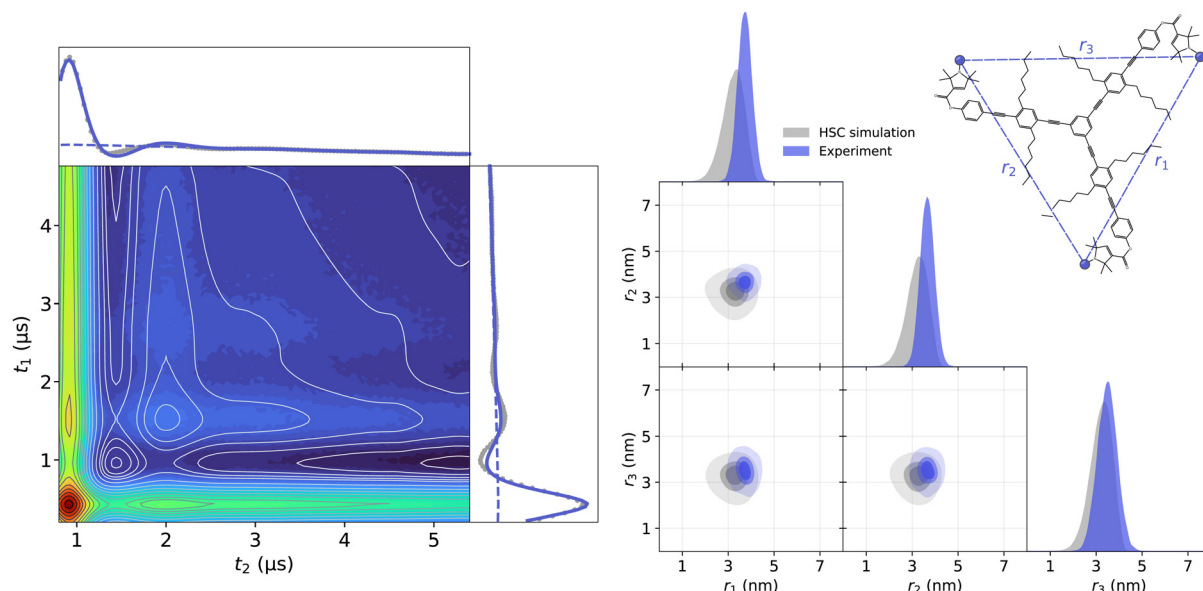


Fig. 11 Analysis with DeerLab of the Q-band TRIER data on oligoPPE triradical T111 (top right), acquired with $\tau_1 = 0.4 \mu\text{s}$, $\tau_2 = 4.6 \mu\text{s}$, and $\tau_3 = 0.9 \mu\text{s}$. The two-dimensional experimental dataset is shown in the left panel as filled colored contours along with the model fit shown as greyscale contour lines. The signal integrals along each dimension are shown in the insets as grey dots for the experimental data and as a solid blue line for the model fit and a dashed blue line for the unmodulated contribution. The right panel shows the fitted trivariate distance distribution. The univariate marginal distributions are shown as filled areas and the bivariate marginal distributions are shown as colored contours. The fitted distribution is shown in blue and the HSC simulation is shown in grey for reference.

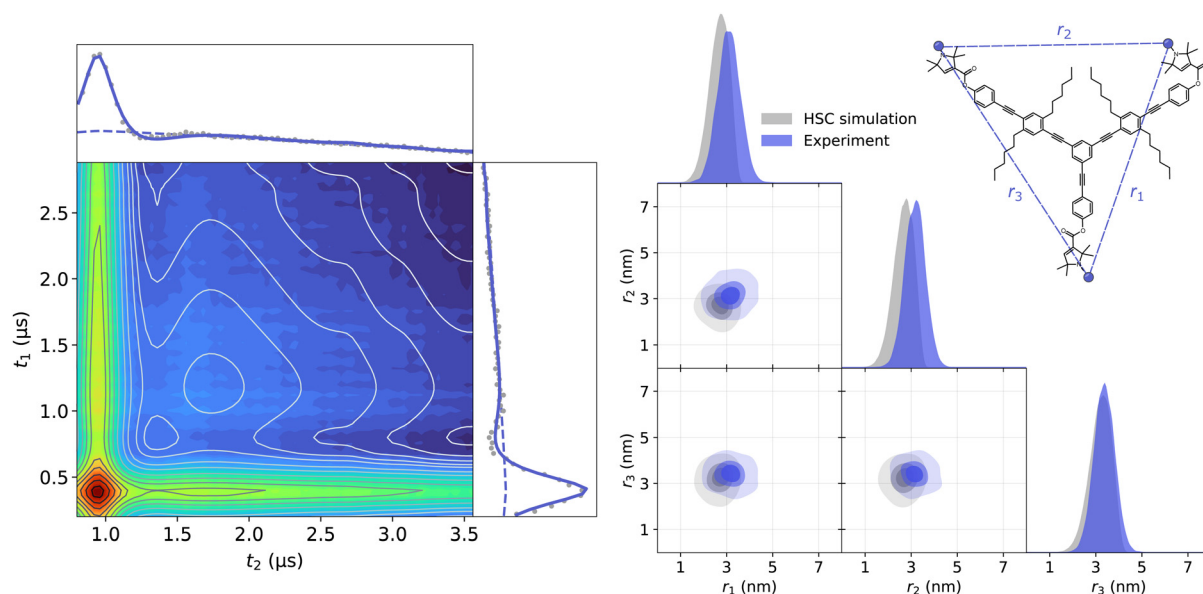


Fig. 12 Analysis with DeerLab of the Q-band TRIER data on oligoPPE triradical T011, (top right) acquired with $\tau_1 = 0.4 \mu\text{s}$, $\tau_2 = 2.8 \mu\text{s}$, and $\tau_3 = 0.9 \mu\text{s}$. The two-dimensional experimental dataset is shown in the left panel as filled colored contours along with the model fit shown as greyscale contour lines. The signal integrals along each dimension are shown in the insets as grey dots for the experimental data and as a solid blue line for the model fit and a dashed blue line for the unmodulated contribution. The right panel shows the fitted trivariate distance distribution. The univariate marginal distributions are shown as filled areas and the bivariate marginal distributions are shown as colored contours. The fitted distribution is shown in blue and the HSC simulation is shown in grey.

analysis method, which we now consider to be inappropriate. This method was based on a simplistic background correction, which fails to account for important background features,⁶⁵ the

application of a smoothing filter that can enhance correlations or possibly even introduce spurious ones, and the use of a two-dimensional approximate Pake transformation based on

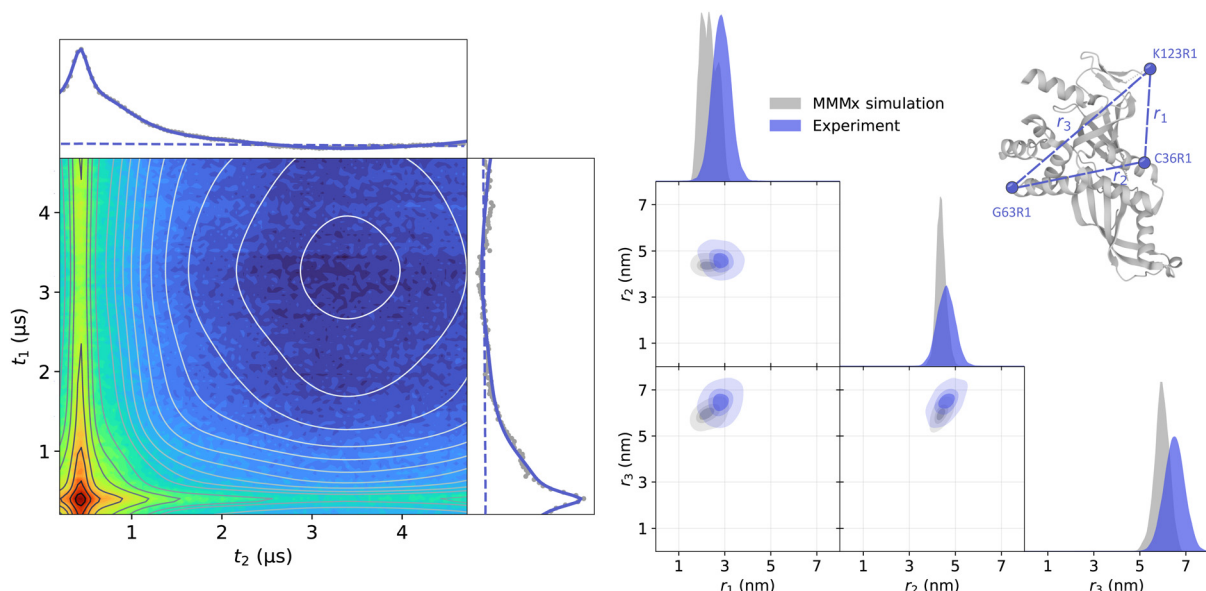


Fig. 13 Analysis with DeerLab of the Q-band TRIER data of triply MTSL-labeled Rpo47 protein (ribbon model top right, PDB: 1GO3), acquired with $\tau_1 = 0.4 \mu\text{s}$, $\tau_2 = 2.8 \mu\text{s}$, and $\tau_3 = 0.4 \mu\text{s}$. The two-dimensional experimental dataset is shown in the left panel as filled colored contours along the model fit shown as greyscale contour lines. The signal integrals along each dimension are shown in the insets as grey dots for the experimental data and as a solid blue line for the model fit and a dashed blue line for the unmodulated contribution. The right panel shows the fitted trivariate distance distribution. The univariate marginal distributions are shown as filled areas and the bivariate marginal distributions are shown as colored contours. The fitted distribution is shown in blue and the MMMx simulation is shown in grey.

the two-spin dipolar kernel, which is inaccurate for the multi-spin case.

5 Conclusions

Multi-spin systems have long been an attractive target of dipolar EPR spectroscopy studies. We have presented a comprehensive model for describing multi-dimensional dipolar EPR experiments on multi-spin systems within the set of assumptions listed in this work. This model expands our previous model for two-spin systems based on dipolar pathways. Under currently achievable experimental conditions, a dipolar signal for any multi-spin system can be described by its contributions arising from two- and three-spin interactions, which in turn can be well approximated in terms of pair dipolar pathways. We have illustrated this by examining the commonly employed one-dimensional 4-pulse DEER experiment and the more recent two-dimensional TRIER experiment. The presence of more than two spins can significantly increase the complexity of the models compared to the two-spin case. However, when applying these models to experimental multi-spin data, a large majority of dipolar pathways can be reasonably neglected under typical experimental conditions. The contributions from many pathways can be minimized experimentally by minimizing the excitation band overlap and maximizing the inversion efficiency of the pulses. Only if the contribution of a dipolar pathway exceeds the noise level must it be included in the analysis of the experimental data. Furthermore, unless the overall modulation depth of a dipolar signal is extremely low,

the three-spin contributions cannot be safely neglected, and the analysis of the dipolar signal in terms of just two-spin interactions is inaccurate. We have tested the multi-pathway model by analyzing different series of 4-pulse DEER and TRIER datasets acquired on different three- and four-spin model compounds as well as on a triple-labeled protein complex. The model, with a proper selection of dipolar pathways, is able to accurately describe the experimental data.

The model-based analysis of experimental multi-spin dipolar data allows the extraction of multivariate distance distributions. We have also shown that the extraction of such multivariate distance distributions and the correlation information between the different distances are not limited to three-frequency experiments such as TRIER but can as well rely on a series of two-frequency four-pulse DEER experiments with varying pump pulse flip angles. The fitted multivariate distance distributions are expected to provide precious information, *e.g.*, for the characterization of conformation ensembles of partially disordered protein systems. Since non-parametric approaches are currently unfeasible, the most challenging aspect is the choice of model for the multivariate distance distribution and the size of the parameter space associated with that model.

The multi-pathway model further provides the means to analyze complex experiments, such as the three-frequency TRIER experiment, for which no complete model had been established before to accurately describe the data and analyze them in terms of a complete multivariate distance distribution. While the TRIER experiment is theoretically and experimentally complex, the two-dimensional information provided by its signals represents an important means of studying multi-spin systems.



Future work is needed to address the experimental and numerical conditions under which the model information can be best extracted from multi-spin dipolar signals. Particularly, additional efforts will be required to improve the numerical fitting protocols and to ensure the uniqueness and robustness of the analysis. The study of other types of multivariate distance distribution models will also be required for the extended applicability of model-based analysis of multi-spin dipolar EPR spectroscopy experiments.

Author contributions

L. F. I., S. S. and G. J. developed the theory. V. M. performed the oligoPPE TRIER measurements, T. v. H. performed the oligoPPE DEER measurements, I. R. performed the Rpo47 DEER and TRIER measurements. L. F. I. implemented and performed the analysis of all experimental datasets. L. F. I., V. M., S. S. and G. J. contributed towards writing the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank Olav Schiemann for providing the tetraradical oligoPPE sample and Daniel Klose for support in generating the Rpo47 sample. This work was funded by the ETH Zurich (ETH-35 18-2, G. J. and L. F. I.), by the Swiss National Science Foundation (SNSF200020_188467, G. J. and V. M.), and by the National Institutes of Health (GM125753, S. S.).

References

- 1 A. D. Milov, K. M. Salikhov and M. D. Shchirov, *Soviet Phys. Solid State*, 1981, **23**, 565–569.
- 2 O. Schiemann, C. A. Heubach, D. Abdullin, K. Ackermann, M. Azarkh, E. G. Bagryanskaya, M. Drescher, B. Endeward, J. H. Freed, L. Galazzo, D. Goldfarb, T. Hett, L. Esteban Hofer, L. Fábregas Ibáñez, E. J. Hustedt, S. Kucher, I. Kuprov, J. E. Lovett, A. Meyer, S. Ruthstein, S. Saxena, S. Stoll, C. R. Timmel, M. Di Valentin, H. S. Mchaourab, T. F. Prisner, B. E. Bode, E. Bordignon, M. Bennati and G. Jeschke, *J. Am. Chem. Soc.*, 2021, **143**, 17875–17890.
- 3 O. Schiemann and T. F. Prisner, *Q. Rev. Biophys.*, 2007, **40**, 1–53.
- 4 G. Jeschke, *Annu. Rev. Phys. Chem.*, 2012, **63**, 419–446.
- 5 M. Drescher, in *EPR Spectroscopy: Applications in Chemistry and Biology*, ed. M. Drescher and G. Jeschke, Springer, Berlin, Heidelberg, 2012, pp. 91–119.
- 6 R. E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich and H. W. Spiess, *Angew. Chem., Int. Ed.*, 1998, **37**, 2833–2837.
- 7 M. Pannier, V. Schädler, M. Schöps, U. Wiesner, G. Jeschke and H. W. Spiess, *Macromolecules*, 2000, **33**, 7812–7818.
- 8 M. Pannier, S. Veit, A. Godt, G. Jeschke and H. W. Spiess, *J. Magn. Reson.*, 2000, **142**, 331–340.
- 9 P. P. Borbat, E. R. Georgieva and J. H. Freed, *J. Phys. Chem. Lett.*, 2013, **4**, 170–175.
- 10 P. E. Spindler, I. Wacławska, B. Endeward, J. Plackmeyer, C. Ziegler and T. F. Prisner, *J. Phys. Chem. Lett.*, 2015, **6**, 4331–4335.
- 11 A. Doll and G. Jeschke, *Phys. Chem. Chem. Phys.*, 2017, **19**, 1039–1053.
- 12 M. Di Valentin, M. Albertini, M. G. Dal Farra, E. Zurlo, L. Orian, A. Polimeno, M. Gobbo and D. Carbonera, *Chem. – Eur. J.*, 2016, **22**, 17204–17214.
- 13 C. Hintze, D. Bucker, S. Domingo Köhler, G. Jeschke and M. Drescher, *J. Phys. Chem. Lett.*, 2016, **7**, 2204–2209.
- 14 T. Bahrenberg, Y. Yang, D. Goldfarb and A. Feintuch, *Magnetochemistry*, 2019, **5**, 20.
- 15 S. Pribitzer, M. Sajid, M. Hülsmann, A. Godt and G. Jeschke, *J. Magn. Reson.*, 2017, **282**, 119–128.
- 16 L. V. Kulik, S. A. Dzuba, I. A. Grigoryev and Y. D. Tsvetkov, *Chem. Phys. Lett.*, 2001, **343**, 315–324.
- 17 S. Milikisyants, F. Scarpelli, M. G. Finiguerra, M. Ubbink and M. Huber, *J. Magn. Reson.*, 2009, **201**, 48–56.
- 18 D. Abdullin, M. Suchatzki and O. Schiemann, *Appl. Magn. Reson.*, 2021, 1–16.
- 19 S. Saxena and J. H. Freed, *Chem. Phys. Lett.*, 1996, **251**, 102–110.
- 20 S. Saxena and J. H. Freed, *J. Chem. Phys.*, 1997, **107**, 1317–1340.
- 21 L. Fábregas-Ibáñez, M. H. Tessmer, G. Jeschke and S. Stoll, *Phys. Chem. Chem. Phys.*, 2022, **24**, 2504–2520.
- 22 P. P. Borbat and J. H. Freed, in *Structural Information from Spin-Labels and Intrinsic Paramagnetic Centres in the Biosciences*, ed. C. R. Timmel and J. R. Harmer, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, vol. 152, pp. 1–82.
- 23 K. M. Salikhov and I. T. Khairuzhdinov, *Appl. Magn. Reson.*, 2015, **46**, 67–83.
- 24 K. M. Salikhov, I. T. Khairuzhdinov and R. B. Zaripov, *Appl. Magn. Reson.*, 2014, **45**, 573–619.
- 25 B. E. Bode, D. Margraf, J. Plackmeyer, G. Dürner, T. F. Prisner and O. Schiemann, *J. Am. Chem. Soc.*, 2007, **129**, 6736–6745.
- 26 G. Jeschke, M. Sajid, M. Schulte and A. Godt, *Phys. Chem. Chem. Phys.*, 2009, **11**, 6580–6591.
- 27 T. von Hagens, Y. Polyhach, M. Sajid, A. Godt and G. Jeschke, *Phys. Chem. Chem. Phys.*, 2013, **15**, 5854–5866.
- 28 G. Hagelueken, W. J. Ingledew, H. Huang, B. Petrovic-Stojanovska, C. Whitfield, H. ElMkami, O. Schiemann and J. H. Naismith, *Angew. Chem., Int. Ed.*, 2009, **48**, 2904–2906.
- 29 B. Endeward, J. A. Butterwick, R. MacKinnon and T. F. Prisner, *J. Am. Chem. Soc.*, 2009, **131**, 15246–15250.
- 30 M. J. N. Junk, H. W. Spiess and D. Hinderberger, *J. Magn. Reson.*, 2011, **210**, 210–217.
- 31 P. Braun, B. Nägele, V. Wittmann and M. Drescher, *Angew. Chem., Int. Ed.*, 2011, **50**, 8428–8431.
- 32 O. Dalmás, H. C. Hyde, R. E. Hulse and E. Perozo, *J. Am. Chem. Soc.*, 2012, **134**, 16360–16369.
- 33 A. Giannoulis, R. Ward, E. Branigan, J. H. Naismith and B. E. Bode, *Mol. Phys.*, 2013, **111**, 2845–2854.



- 34 S. Valera, K. Ackermann, C. Pliotas, H. Huang, J. H. Naismith and B. E. Bode, *Chem. – Eur. J.*, 2016, **22**, 4700–4703.
- 35 K. Ackermann and B. E. Bode, *Mol. Phys.*, 2018, **116**, 1513–1521.
- 36 S. Stoll and B. Kasumaj, *Appl. Magn. Reson.*, 2008, **35**, 15–32.
- 37 S. R. Sweger, V. P. Denysenkov, L. Maibaum, T. F. Prisner and S. Stoll, *Magn. Reson.*, 2022, **3**, 101–110.
- 38 A. F. Bedilo and A. G. Maryasov, *J. Magn. Reson., Ser. A*, 1995, **116**, 87–96.
- 39 D. Abdullin, P. Brehm, N. Fleck, S. Spicher, S. Grimme and O. Schiemann, *Chem. – Eur. J.*, 2019, **25**, 14388–14398.
- 40 A. Weber, O. Schiemann, B. Bode and T. F. Prisner, *J. Magn. Reson.*, 2002, **157**, 277–285.
- 41 C. Riplinger, J. P. Y. Kao, G. M. Rosen, V. Kathirvelu, G. R. Eaton, S. S. Eaton, A. Kutateladze and F. Neese, *J. Am. Chem. Soc.*, 2009, **131**, 10092–10106.
- 42 Y. Liu, F. A. Villamena, A. Rockenbauer, Y. Song and J. L. Zweier, *J. Am. Chem. Soc.*, 2013, **135**, 2350–2356.
- 43 K. Keller, M. Qi, C. Gmeiner, I. Ritsch, A. Godt, G. Jeschke, A. Savitsky and M. Yulikov, *Phys. Chem. Chem. Phys.*, 2019, **21**, 8228–8245.
- 44 A. A. Nevzorov and J. H. Freed, *J. Chem. Phys.*, 2001, **115**, 2401–2415.
- 45 J. Soetbeer, M. Hülsmann, A. Godt, Y. Polyhach and G. Jeschke, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1615–1628.
- 46 J. L. Baber, J. M. Louis and G. M. Clore, *Angew. Chem., Int. Ed.*, 2015, **54**, 5336–5339.
- 47 F. D. Breitgoff, J. Soetbeer, A. Doll, G. Jeschke and Y. O. Polyhach, *Phys. Chem. Chem. Phys.*, 2017, **19**, 15766–15779.
- 48 M. Teucher and E. Bordignon, *J. Magn. Reson.*, 2018, **296**, 103–111.
- 49 G. Jeschke and R. Bittl, *Chem. Phys. Lett.*, 1998, **294**, 323–331.
- 50 A. D. Milov, A. B. Ponomarev and Y. D. Tsvetkov, *Chem. Phys. Lett.*, 1984, **110**, 67–72.
- 51 D. R. Kattinig, J. Reichenwallner and D. Hinderberger, *J. Phys. Chem. B*, 2013, **117**, 16542–16557.
- 52 S. Pribitzer, L. F. Ibáñez, C. Gmeiner, I. Ritsch, D. Klose, M. Sajid, M. Hülsmann, A. Godt and G. Jeschke, *Appl. Magn. Reson.*, 2018, **49**, 1–27.
- 53 L. Fábregas-Ibáñez, G. Jeschke and S. Stoll, *Magn. Reson.*, 2020, **1**, 209–224.
- 54 Z. Ugray, L. Lasdon, J. Plummer, F. Glover, J. Kelly and R. Martí, *INFORMS J. Comput.*, 2007, **19**, 328–340.
- 55 Y. Kurihara, *Mon. Weather Rev.*, 1965, **93**, 399–415.
- 56 D. Wang and G. R. Hanson, *J. Magn. Reson., Ser. A*, 1995, **117**, 1–8.
- 57 S. Stoll and A. Schweiger, *J. Magn. Reson.*, 2006, **178**, 42–55.
- 58 A. Godt, M. Schulte, H. Zimmermann and G. Jeschke, *Angew. Chem., Int. Ed.*, 2006, **45**, 7560–7564.
- 59 G. Jeschke, M. Sajid, M. Schulte, N. Ramezani, A. Volkov, H. Zimmermann and A. Godt, *J. Am. Chem. Soc.*, 2010, **132**, 10107–10117.
- 60 M. F. Czar, F. D. Breitgoff, D. Sahoo, M. Sajid, N. Ramezani, Y. Polyhach, G. Jeschke, A. Godt and R. Zenobi, *J. Phys. Chem. Lett.*, 2019, **10**, 6942–6947.
- 61 G. Jeschke and L. Esteban-Hofer, in *Methods in Enzymology*, ed. R. D. Britt, Academic Press, 2022, vol. 666 of Advances in Biomolecular EPR, pp. 145–169.
- 62 T. von Hagens, PhD thesis, ETH Zurich, Zurich, 2014.
- 63 I. Ritsch, PhD thesis, ETH Zurich, Zurich, 2019.
- 64 A. Doll and G. Jeschke, *J. Magn. Reson.*, 2017, **280**, 46–62.
- 65 L. Fábregas-Ibáñez and G. Jeschke, *Phys. Chem. Chem. Phys.*, 2020, **22**, 1855–1868.

