Compactness regularization in the analysis of dipolar EPR spectroscopy data

Originally published in:
Journal of Magnetic Resonance 339, https://doi.org/10.1016/j.jmr.2022.107218

Funding acknowledgement:
ETH-35 18-2 - Optimal regularization in computation of one- and two-dimensional distance distributions from pulsed dipolar spectroscopy data (ETHZ)

Journal of Magnetic Resonance 339 (2022) 107218

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Compactness regularization in the analysis of dipolar EPR spectroscopy data

Luis Fábregas-Ibáñez [a,*], Gunnar Jeschke [a], Stefan Stoll [b]

[a] ETH Zurich, Laboratory of Physical Chemistry, Vladimir-Prelog-Weg 2, Zurich 8093, Switzerland
[b] University of Washington, Department of Chemistry, Seattle, WA 98195, USA

ARTICLE INFO

Article history:
Received 18 February 2022
Revised 4 April 2022
Accepted 5 April 2022
Available online 09 April 2022

Keywords:
Electron paramagnetic resonance
Dipolar EPR spectroscopy
Pulse dipolar spectroscopy
Identifiablity
Regularization
Profile likelihood
Compactness
DEER
PELDOR
Distance distribution
Data analysis

ABSTRACT

Dipolar electron paramagnetic resonance (EPR) experiments, such as double electron–electron resonance (DEER), measure distributions of nanometer-scale distances between paramagnetic centers, which are valuable for structural characterization of proteins and other macromolecular systems. One challenge in the least-squares fitting analysis of dipolar EPR data is the separation of the inter-molecular contribution (background) and the intra-molecular contribution. For noisy experimental traces of insufficient length, this separation is not unique, leading to identifiability problems for the background model parameters and the long-distance region of the intra-molecular distance distribution. Here, we introduce a regularization approach that mitigates this by including an additional penalty term in the objective function that is proportional to the variance of the distance distribution and thereby penalizes non-compact distributions. We examine the reliability of this approach statistically on a large set of synthetic data and illustrate it with an experimental example. The results show that the introduction of compactness can improve identifiability.

1. Introduction

Dipolar electron paramagnetic resonance (EPR) spectroscopy enables the measurement of distributions of nanometer-scale distances between pairs of electron spin centers in paramagnetic systems [1]. The most widely applied dipolar EPR technique is four-pulse double electron–electron resonance (DEER) [2,3], but a range of other methods exist [4–15]. All these techniques record a time-domain signal modulated by the dipolar interaction between pairs of electrons.

From this time-domain signal, the distance distribution and other quantities of interest are inferred by least-squares fitting a physical model to the signal [16–20]. The model depends on an intra-molecular distance distribution $P(r)$, a modulation depth parameter, and parameter(s) for the inter-molecular contribution to the signal (the background) such as the total spin concentration. The distance distribution P is modeled either as a linear combination of parameterized basis functions, such as Gaussians, or as a non-parametric histogram. The parameters, including the histogram weights, are determined by least-squares fitting.

When a non-parametric histogram is used to represent P, the model has several specific properties that lead to challenges and failures in the ordinary least-squares approach. In the presence of noise and signal truncation, the fits are not unique regarding P. Hence, one cannot define a quality or amount of time-domain data that ensures a desired uncertainty of the distance distribution P. In other words, a subset of different P can fit an experimental signal roughly equally well. These P differ in their spikiness, and in order to stabilize the least-squares fitting and obtain a unique P, a term is added to the objective function that penalizes spiky distributions [21,22,16]. This is called Tikhonov regularization. Second, the fits are not unique regarding the long-distance region of P and the background. Different combinations of long-distance P and background parameters can fit a given experimental signal roughly equally well. Using a parametric model with a small number of components for P alleviates these problems to some extent, but at the cost of putting restrictions on the shape and complexity of P.

From a broader model fitting perspective, both challenges above concern the (practical) identifiability of model parameters. A

* Corresponding author.
E-mail address: luis.fabregas@phys.chem.ethz.ch (L. Fábregas-Ibáñez).

https://doi.org/10.1016/j.jmr.2022.107218
1090-7807/

model parameter is identifiable when distinct parameter values yield distinct model predictions [23] for a given amount and quality of the data. Otherwise, the parameter, and the entire model, is non-identifiable. Identifiability affects models in many different fields and is a broad topic in statistical literature [24–26,26–29,23,30].

In this work, we introduce an approach to address the second identifiability challenge, i.e. the uniqueness of background vs. long-distance contributions in P. We use the profile likelihood approach to examine identifiability. We then propose to add a second penalty term to the objective function that is proportional to the variance of the distribution and penalizes non-compact distributions. This regularization approach converts many, though not all, non-identifiable models into identifiable ones. We demonstrate its effectiveness on individual synthetic examples, on two large synthetic test datasets, and on an experimental example. The Python-based DeerLab software package [19] implements all relevant methodology presented in this work.

2. Modeling dipolar EPR signals

For samples with highly diluted molecules with two spins each, the dipolar EPR signal contains separate contributions arising from intramolecular and intermolecular interactions [1]. The latter is also referred to as background. In this situation, the signal is modeled as an average over the distribution $P(r)$ of spin–spin distances r as

$$V(t) = \int_0^\infty \mathrm{d}r\, K(t,r)P(r) \tag{1}$$

where $K(t,r)$ is the dipolar kernel function that describes the mapping from the distance-domain to the time-domain signal.

In the simple case of four-pulse DEER with no overlap between pump and probe pulse excitation profiles the dipolar kernel reads

$$K(t,r) = V_0[1-\lambda+\lambda K_0(t,r)]\exp\left(-\kappa c\lambda|t|\right) \tag{2}$$

where λ is the modulation depth, c is the spin concentration, V_0 is the overall amplitude, and $\kappa = 8\pi^2/(9\sqrt{3})D$ is numerical constant with the dipolar constant D given by

$$D = \frac{\mu_0}{4\pi}\frac{g_\mathrm{e}^2\mu_\mathrm{B}^2}{\hbar} \tag{3}$$

which includes the g-value of the free electron g_e, the Bohr magneton μ_B, the magnetic constant μ_0, the reduced Planck constant $\hbar$. The elementary dipolar kernel $K_0(t,r)$ function is given by

$$K_0(t,r) = \int_0^1 \mathrm{d}z\, \cos\left[(1-3z^2)\frac{Dt}{r^3}\right] \tag{4}$$

While in this work we assume a spatially homogeneous 3D distribution of isolated spin pairs for the intermolecular contribution, the method is expected to work similarly for models that include volume exclusion effect or fractal dimensions [31–33]. Also, note that the dipolar kernel in Eq. (2) accounts for the intermolecular contributions and bypasses the need for a modification of the primary signal [34].

Experimentally, the echo amplitude $V(t)$ is recorded at a discrete set of n time points t_i, leading to a discretized dipolar signal vector $\mathbf{V} = (V_1,\ldots,V_n)$ with $V_i = V(t_i)$. In the analysis, $P(r)$ is also represented as a vector $\mathbf{P} = (P_1,\ldots,P_m)$ over a discrete set of m equidistant distances r_j, where $P_j = P(r_j)$. With this, Eq. (1) reads

$$\mathbf{V} = \mathbf{K}\mathbf{P} \tag{5}$$

where $\mathbf{K}$ is the $n\times m$ kernel matrix with elements $K_{ij} = K(t_i,r_j)\Delta r$ and Δr is the distance-domain increment. While $P(r)$ is defined for $r \geqslant 0$, in the analysis the range is limited to $r_\mathrm{min} \leqslant r_j \leqslant r_\mathrm{max}$, implicitly assuming that the distance distribution is zero outside of this range. Typically $r_\mathrm{min} \approx 1.5\,\mathrm{nm}$ is selected as the lower limit, and r_max is usually selected according to some empirically derived heuristic [35,36]. Both r_min and r_max are model parameters.

Besides r_min and r_max, the dipolar kernel $\mathbf{K}[\boldsymbol{\theta}]$ depends on the two independent parameters $\boldsymbol{\theta} = (\lambda, c)$. If the shape of the distance distribution is known, the vector $\mathbf{P}$ can be represented by a parametric function (e.g. Gaussian). We will focus on semiparametric models, which use a non-parametric distance distribution $\mathbf{P}(\boldsymbol{\theta})$, where each element P_j is an independent parameter.

The analysis of an experimental signal

$$\mathbf{V}_\mathrm{exp} = \mathbf{V} + \boldsymbol{\delta} \tag{6}$$

with additional noise vector $\boldsymbol{\delta}$, involves the estimation of the model parameters that best fit the model to the signal. When using non-parametric distance distributions, this can be solved via separable non-linear least-squares [37,38,19]. In this approach, all parameters except $\mathbf{P}$ are determined via

$$\boldsymbol{\theta}_\mathrm{fit} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\left\{F(\boldsymbol{\theta})\right\} \tag{7}$$

with the objective function

$$F(\boldsymbol{\theta}) = \|\mathbf{V}_\mathrm{exp} - \mathbf{K}(\boldsymbol{\theta})\mathbf{P}(\boldsymbol{\theta})\|^2 + \alpha^2\|\mathbf{L}\mathbf{P}(\boldsymbol{\theta})\|^2 \tag{8}$$

In each iteration of Eq. (7), $\mathbf{P}(\boldsymbol{\theta})$ is determined for the given $\boldsymbol{\theta}$ via the linear sub-problem

$$\mathbf{P}(\boldsymbol{\theta}) = \underset{\mathbf{P}\geqslant 0}{\operatorname{argmin}}\left\{\|\mathbf{V}_\mathrm{exp} - \mathbf{K}(\boldsymbol{\theta})\mathbf{P}\|^2 + \alpha^2\|\mathbf{L}\mathbf{P}\|^2\right\} \tag{9}$$

The first term in the objective functions in Eqs. (8) and (9) describes the discrepancy between the model prediction and the experimental signal. The second term is the Tikhonov regularization term that penalizes spiky distributions, where $\mathbf{L}$ is the second-order differential operator and α is the regularization parameter. An adequate value of α can be selected via, e.g., an L-curve criterion, the Akaike information criterion (AIC) or generalized cross-validation (GCV) [39].

The least-squares method is closely related to maximum-likelihood estimation. If the noise in $\mathbf{V}_\mathrm{exp}$ is normally distributed (as is the case in dipolar signals [40]), minimization of a least-squares objective function is equivalent to the maximization of the likelihood function $\mathcal{L}(\boldsymbol{\theta})$ which is related to the objective function via

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{F(\boldsymbol{\theta})}{2\sigma^2}\right) \tag{10}$$

where σ is the standard deviation of the noise $\boldsymbol{\delta}$ in V_exp.

3. Identifiability analysis

In the context of least-squares and maximum-likelihood estimation, a model parameter θ_i is said to be identifiable at $\theta_{\mathrm{fit},i}$ on a given interval if for a given experimental dataset all other parameter values $\theta_i \neq \theta_{\mathrm{fit},i}$ yield different objective function values [24,26], formally

$$F(\theta_i) = F(\theta_{i,\mathrm{fit}}) \Longleftrightarrow \theta_i = \theta_{i,\mathrm{fit}} \tag{11}$$

This is illustrated in Fig. 1. Practical non-identifiability corresponds to the presence of *nearly* equivalent minima of the objective function such as a flat region or multiple equivalent global minima, where multiple parameter combinations yield objective-function values for a given noisy and truncated experimental signal.

There are several methods for analyzing identifiability. A straightforward approach involves mapping the objective function $F(\boldsymbol{\theta})$ over the whole parameter space to assess whether the objective function has extended minima regions. This becomes imprac-

tical in the presence of more than a few parameters. Instead, we will use the profile likelihood method [41–43] for this purpose. This method is well established in other fields [29,44–46,30].

The use of the profile likelihood method [41–43] for identifiability analysis as proposed by Raue et al. [29] provides a simple and graphical representation to qualitatively and quantitatively assess the identifiability of a model. The profile likelihood (or likelihood ratio) for a parameter θ_i is calculated by fixing θ_i at different values and maximizing the likelihood with respect to all other parameters. Equivalently, the profile objective function is obtained via

$$\mathcal{F}(\theta_i) = \min_{\boldsymbol{\theta\prime}}\{F(\theta_i, \boldsymbol{\theta\prime})\} \tag{12}$$

where $\boldsymbol{\theta\prime}$ represents the subset of $\boldsymbol{\theta}$ without θ_i. Graphically, as shown in Fig. 2, this corresponds to traversing the parameter space along θ_i at values of $\boldsymbol{\theta\prime}$ that minimize F.

Under certain regularity conditions [47,48], we can define a $\gamma \times 100\%$ likelihood-based confidence region consisting of the set of values of parameter θ_i that satisfy the inequality

$$\mathcal{F}(\theta_i) < F(\boldsymbol{\theta}_{\text{fit}}) + \sigma^2 \chi^2_{\gamma,1} = \Delta_\gamma \tag{13}$$

where $\chi^2_{\gamma,1}$ denotes the γ-quantile of the chi-squared distribution for one degree of freedom. In the context of dipolar EPR spectroscopy, these confidence intervals have been applied to dipolar models based on multi-Gaussian parametrization of the distance distribution [49,17]. In this work, we will use the $\gamma = 0.95$ confidence regions as threshold, i.e. $\chi^2_{0.95,1} = 3.841$.

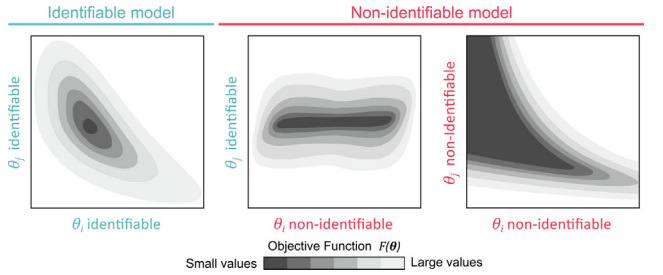

Fig. 1. Schematic representation of (non)-identifiability based on flatness of the objective function. The objective function $F(\boldsymbol{\theta})$ is shown as greyscale contours for three cases: (left) both parameters are identifiable, (center) one of the parameters is non-identifiable, and (right) both parameters are non-identifiable. The presence of a non-identifiable parameter renders the model non-identifiable.

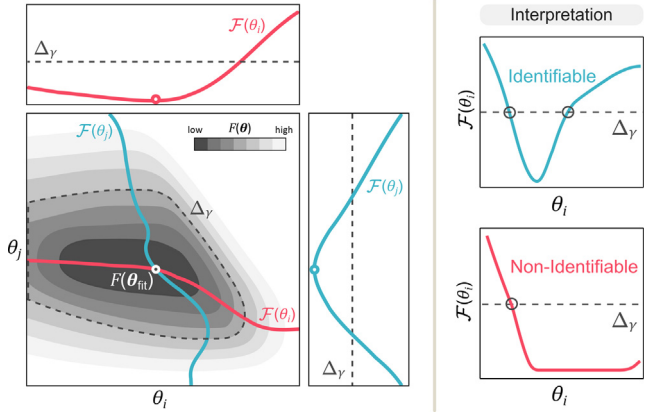

Fig. 2. Profile likelihood method for identifiability analysis. (Left) The objective function for two parameters θ_i and θ_j is shown as greyscale contours and the minimum is indicated as a white circle. The superimposed colored lines represent the isometric likelihood profiles $\mathcal{F}(\theta_i)$ and $\mathcal{F}(\theta_j)$ computed at the model estimate. The insets show the likelihood profiles for both parameters. The dashed grey lines represent the threshold Δ_α. (Right) Identifiability analysis based on the profile likelihood. If the profile $\mathcal{F}(\theta_i)$ crosses the threshold Δ_α twice, shown as black circles, the parameter θ_i is identifiable (green profile). If the profile remains flat and crosses only once, the parameter θ_i is non-identifiable (red profile).

In the framework proposed by Raue et al. we can make a binary decision on the identifiability of individual parameters θ_i by calculating the profile objective functions and comparing them to the threshold Δ_γ [29]. In the case of an identifiable parameter, the profile $\mathcal{F}(\theta_i)$ has a well-defined minimum and crosses Δ_γ twice as depicted in Fig. 2, meaning that the confidence region of θ_i is bounded. For a non-identifiable parameter, the profile has a minimum but does not increase enough to cross the threshold again as depicted in Fig. 2 or will be completely flat, indicating an unbounded confidence region of θ_i. While a binary interpretation might be useful to automated workflows, the most important criterion to apply to these profiles is the presence/absence of significant flat minima regions or multiple equivalent minima.

4. Identifiability in dipolar EPR spectroscopy

Dipolar models feature several parameters, such as the spin concentration, the modulation depth, or the parameters of a parametric distance distribution model. Most of them can suffer from non-identifiability. Some examples of likelihood profiles of multi-Gaussian distance distribution model parameters can be found in the literature [49,17], where some examples of non-identifiable distribution parameters can also be found.

If a non-parametric distance distribution is used, the estimation of $\mathbf{P}(\boldsymbol{\theta})$ constitutes an ill-posed problem. Ill-posedness represents an extreme case of non-identifiability. Still, regularization provides a stable estimate of the distance distribution by imposing a degree of smoothness that renders the histogram coefficients identifiable. However, if the modulation depth or the spin concentration are non-identifiable, the fitted distance distribution, while being stable for the assumed background model, can be an inaccurate estimate of the ground truth and should not be interpreted.

We illustrate this in Fig. 3 for a 4-pulse DEER signal. The objective-function surface $F(\lambda, c)$ from Eq. (8) presents a flat valley showing that both the modulation depth and the spin concentration parameters are non-identifiable. This is supported by the profile objective functions, which indicate the non-identifiability of the spin concentration. While the modulation depth profile is identifiable according to the definition above, it still presents a wide region of flat profile. The set of solutions cover a wide range of values for both the modulation depth (λ=0.3–0.7) and spin concentration (c = 0–150 μ M) that cannot be associated with the typical uncertainty arising from such a noise level. The wide range of solutions for the parameters results in a wide range of distance distributions that contain multiple artifacts or distortions, particularly towards longer distances or at the edge of the distance range. Despite this, all solutions provide a good fit of the data, supported by the reduced-χ^2 values close to 1. These observations exemplify the dangers of neglecting identifiability in dipolar EPR spectroscopy, particularly for automated data analysis workflows. In this example, it would not be obvious to an expert how to distinguish the ground truth from all other solutions.

Fig. 4 illustrates how this non-identifiability arises for truncated signals. For the sake of clarity, we assume that the objective function is solely given by the least-squares goodness-of-fit of the experimental dipolar signal. For the model to be identifiable, according to Eq. (11), the objective function values $F(\boldsymbol{\theta}_1)$ and $F(\boldsymbol{\theta}_2)$ for both of parameter sets must be different, meaning that the model predictions $\mathbf{V}_{\text{fit},1}$ and $\mathbf{V}_{\text{fit},2}$ must also be different. Here is where the length and quality of the dipolar signal come into play. If the dipolar signal length (t_{max}) is long enough, the model predictions $\mathbf{V}_{\text{fit},1}$ and $\mathbf{V}_{\text{fit},2}$ become distinguishable (see Fig. 4), but if t_{max} is too short, the model predictions are indistinguishable, resulting in

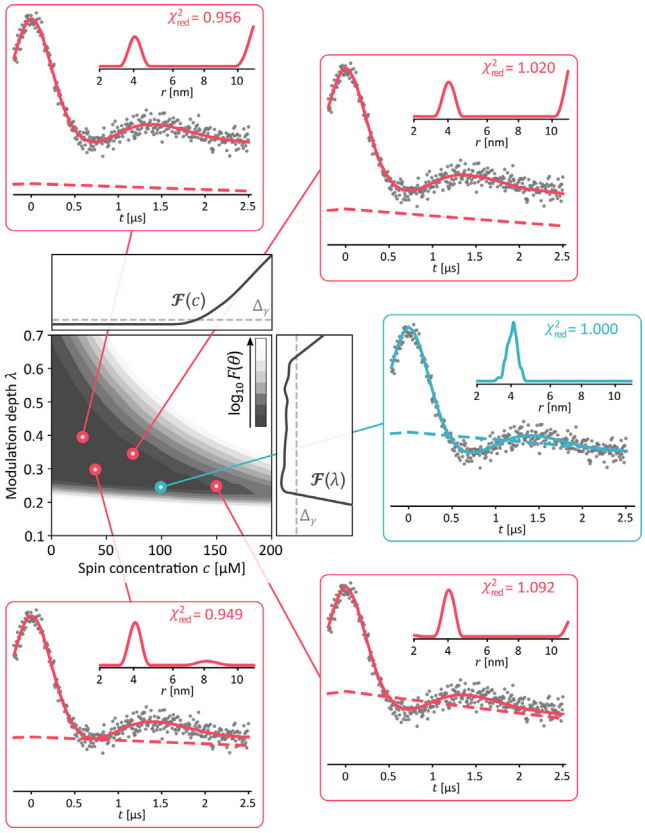

Fig. 3. Example of non-identifiability in a 4-pulse DEER model, using a synthetic dataset with modulation depth $\lambda = 0.25$ and spin concentration $c = 100\,\mu$M. The objective function $F(\boldsymbol{\theta})$ is shown as grey shades on a log scale. The profile objective functions of the modulation depth $\mathcal{F}(\lambda)$ and spin concentration $\mathcal{F}(c)$ are shown as black lines. The threshold $\Delta_{0.95}$ is shown as a grey dashed line for reference. Five parameter sets that minimize the objective function are shown as colored circles. For each, the corresponding model prediction in shown in a separate box. The simulated noisy dipolar signal is shown as grey dots, and the corresponding fitted dipolar signal and background models are shown as solid and dashed colored lines. The associated fitted AIC-Tikhonov distance distributions are shown as colored lines in the insets. For each solution, the corresponding reduced chi-squared χ^2_{red} values are given. The ground truth solution is colored turquoise, while all red solutions are incorrect but still minimize the objective function.

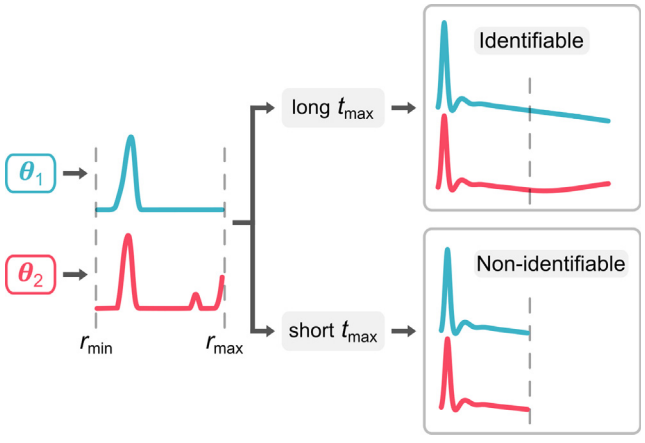

Fig. 4. Equivalence of dipolar model predictions. Two model parameter sets $\boldsymbol{\theta}_1$ (turquoise) and $\boldsymbol{\theta}_2$ (magenta) and their corresponding distance distributions result in two model predictions $\mathbf{V}_{fit,1}$ (turquoise) and $\mathbf{V}_{fit,2}$ (magenta). The equivalency of the model predictions (colored lines) depends on the length of the dipolar signal t_{max}. If t_{max} is not long enough (bottom), $\mathbf{V}_{fit,1}$ and $\mathbf{V}_{fit,2}$ are nearly identical, resulting in similar objective function values $F(\boldsymbol{\theta}_1)$ and $F(\boldsymbol{\theta}_2)$, and in non-identifiability. Inversely, if t_{max} is sufficiently long (top), it will result in identifiability.

equivalent solutions. The importance of the dipolar signal length t_{max} for robust estimation of dipolar models has long been known [50,51]. While not under the scope of identifiability, the discussion of extended trace length to "detect" or "access" longer distances has long been an intensively researched topic and has even motivated the development of new protocols [52,53] and new dipolar EPR experimental techniques (e.g., 5-pulse DEER [6], 7-pulse DEER [7]), and nDEER [8] that aim to surpass the limitations in t_{max}.

The distance vector $\boldsymbol{r}$ is another quantity that critically influences (non)-identifiability of dipolar models. Non-identifiability can occur due to the ability of the distance distribution to allocate distribution mass at long distances to compensate for deviations in spin concentration and modulation depth. Therefore, a larger r range increases the risk of non-identifiability. The distance vector is parametrized by its boundaries r_{min} and r_{max}. While r_{min} usually has minimal effect on model identifiability, r_{max} is essential in terms of identifiability. The larger r_{max}, the more distances will be available, and the more likely the model will be non-identifiable. This was the rationale for empirically deriving a heuristic for determining an appropriate r_{max} given a t_{max} [54].

Fig. 5 illustrates these dependencies on the profile objective functions of a 4-pulse DEER model. As can be seen, unless we choose a sufficiently long t_{max} and short r_{max}, the profile objective functions exhibit extended flat region for one or both model parameters, indicating non-identifiability. Note that the profiles progressively flatten asymmetrically about the ground truth: the modulation depth presents only flat profile regions for values larger than the ground truth and the spin concentration predominantly for values smaller than the ground truth. This can be

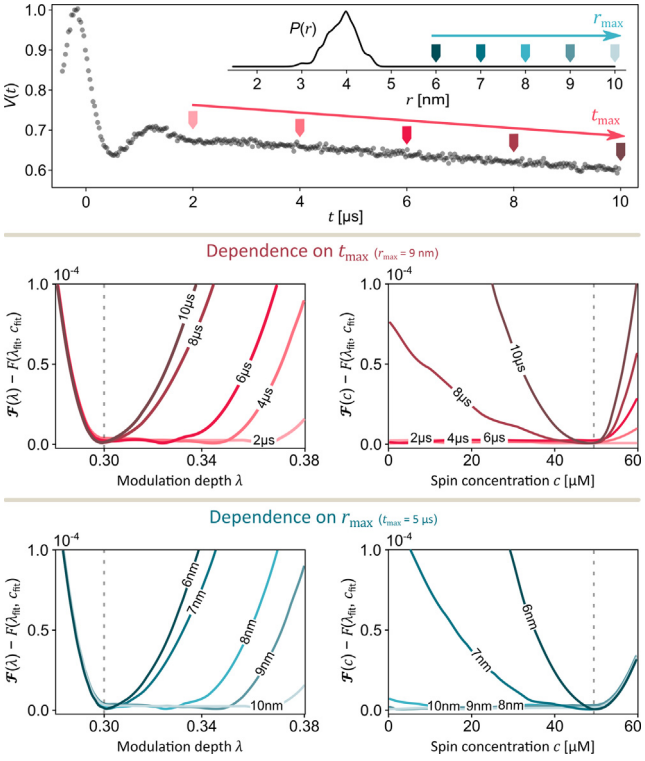

Fig. 5. Dependence of the profile objective functions of a 4-pulse DEER signal on r_{max} and t_{max}. A noisy 4-pulse DEER signal, simulated from the distance distribution shown in the inset with a modulation depth $\lambda = 0.3$ and pump spin concentration $c = 50\,\mu$M, is shown as grey dots for $r_{max} = 11$ nm and $t_{max} = 12\,\mu$s. Different profiles objective functions $\mathcal{F}(\lambda)$ and $\mathcal{F}(c)$ for the modulation depth λ and pump spin concentration c, respectively, as colored lines. The objective profiles dependency on t_{max} at $r_{max} = 9$ nm are shown as red lines and the objective profiles dependency on r_{max} at $t_{max} = 5\,\mu$s are shown as green lines. The parameter ground truth values are given as grey dashed lines for reference. The corresponding t_{max} and r_{max} values are shown on top of each line.

observed as long as the noise level is not excessive. An important consequence of this is that the choice of initial guesses for the optimization of the model parameters becomes relevant. In the DEER case, it is advantageous to start the optimization with an underestimation of the modulation depth and an overestimation of the spin concentration parameters, i.e. where the profile objective functions are less likely to present flat regions.

5. Resolving non-identifiability

Following our discussion thus far, one way to avoid identifiability issues is the acquisition of longer data with less noise. An alternative approach requires the collection of multiple dipolar signals with different configurations or experiments, which can be globally analyzed [19]. However, enhancements to both dipolar trace length and noise level are strongly restricted by experimental and time constraints. Post-acquisition approaches to resolve non-identifiability rely on the introduction of additional (prior) information.

An essential type of prior information that firmly controls identifiability is the aforementioned $r_{\max}$. However, the choice of $r_{\max}$ is non-trivial. As shown, too large values can lead to severe non-identifiability. However, too small values could truncate the distance distribution, rendering a proper analysis unfeasible. Based on structural considerations, we can estimate an upper bound for $r_{\max}$ in specific systems such as proteins or model compounds. However, unless the $r_{\max}$ estimate is accurately close to the edge of the distance distribution or $t_{\max}$ is comfortably long, non-identifiability will remain a problem. It is tempting to rely on heuristic estimators for an optimal $r_{\max}$. The DeerAnalysis and DeerNet software packages include such a heuristic based on the length of the signal [16,36]

$$r_{\max}/\mathrm{nm} = (108 t_{\max}/\mu s)^{1/3} \tag{14}$$

which might provide reasonable estimates in specific instances. Deriving an exact closed-form definition of the optimal $r_{\max}$ is not possible due to its random and recursive nature. Overall, having a reasonable estimate of $r_{\max}$ alleviates non-identifiability but is not necessarily an optimal treatment for data with an unfavorable ratio between the longest distances in the ground-truth distribution and $t_{\max}$.

If reasonable parameter boundaries can be obtained, e.g., for the modulation depth and spin concentrations from independent experiments or simulations, the uncertainty can be significantly reduced. For instance, as shown above, for the 4-pulse DEER case, the uncertainty about the modulation depth and spin concentration parameters can be strongly reduced by a reasonable estimate of an upper bound on the modulation depth or of a lower bound on the spin concentration. From the asymmetry observed in the parameter profiles, it might be tempting to seek the solution with minimal modulation depth and largest concentration as the ground truth. While conceptually reasonable, the ground truth does not always correspond to that solution.

There are many other ways to introduce prior information in the analysis, some of them well established. By modelling a distance distribution with a parametric function (such as a Gaussian distribution), we introduce prior information on the shape and nature of the distribution. Data analysis with neural networks [18] introduces prior information based on the nature of the network's training set. Tikhonov regularization introduces information about the smoothness of the distance distribution, resulting in unique non-parametric distributions. Other parameter constraints such as the non-negativity constraint of the distance distribution introduce information about the nature of the distance distribution. All these forms of prior information contribute towards the identifiability of dipolar models.

It is important to note that prior information is only beneficial as long as it matches the physical reality underlying the data. Otherwise, a dipolar model might become identifiable but no longer an accurate descriptor of the dipolar signal.

6. Compactness regularization

In this context, we aim to find a new type of prior information generally applicable to dipolar EPR data, making a non-identifiable problem identifiable. As shown above, non-identifiability typically manifests as spurious long-distance peaks in the distance distribution. Existing approaches aim to minimize these artifacts by penalizing distribution mass in the long-distance range [55] or optimizing the background to ensure the absence of distribution mass at $r_{\max}$ [16]. Now, if we know that a distance distribution cannot have that spread, we can introduce new prior information in the form of compactness. We define a compact distance distribution as a distribution with a small variance. This is illustrated in Fig. 6. Conceptually, from the set of possible solutions that fit the experimental signal, we penalize those with a larger spread of distances. Compactness can also be justified in terms of protein thermodynamics, since folded and compact proteins are more stable than unfolded ones. A compact distance distribution implies the existence of a single energetic well for an intramolecular distance, which is a common feature of well-structured proteins (e.g. the left panel in Fig. 6) and also very weakly structured proteins (e.g. the middle panel in Fig. 6). A spread-out distribution implies the existence of multiple energetic wells separated by significant energetic barriers (e.g. the right panel in Fig. 6). While specific protein systems can present this scenario, they are uncommon.

Therefore, among the distributions that fit the data equally well, it is reasonable to prefer those distributions that are more compact. To quantify the spread of a distribution, we use the variance (i.e., its second centralized moment) of distances in the distribution defined as

$$\sigma^2[\mathbf{P}(\boldsymbol{\theta})] = \sum_i P_i(\boldsymbol{\theta})\left(r_i - \bar{r}\right)^2 \Delta r \tag{15}$$

where $\bar{r}$ is the mean distance given by

$$\bar{r} = \sum_i P_i(\boldsymbol{\theta}) r_i \Delta r \tag{16}$$

This metric is differentiable, robust with respect to heavy-tailed distributions, and most importantly, location-invariant. Small values of $\sigma^2[\mathbf{P}(\boldsymbol{\theta})]$ indicate a compact distribution, and large values indicate a spread-out distribution (see Fig. 6). We include this term in the non-linear part of the separable non-linear least-squares problem in Eq. (8) as a second weighted penalty

$$F(\boldsymbol{\theta}) = \|\mathbf{V}_{\mathrm{exp}} - \mathbf{K}(\boldsymbol{\theta})\mathbf{P}(\boldsymbol{\theta})\|^2 + \alpha^2\|\mathbf{L}\mathbf{P}(\boldsymbol{\theta})\|^2 + \beta^2\sigma^2[\mathbf{P}(\boldsymbol{\theta})] \tag{17}$$

where β is the parameter that determines the scale of the additional penalty term and $\mathbf{P}(\boldsymbol{\theta})$ remains as defined in Eq. (9). Note that compactness is only imposed upon the search of the non-linear

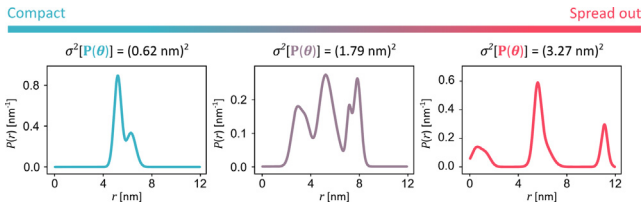

Fig. 6. Illustration of compactness of a distance distribution. Three examples of distance distributions with different variance $\sigma^2[\mathbf{P}(\boldsymbol{\theta})]$ are shown.

parameters, i.e. the background search in a DEER model, whereas no compactness is imposed upon the distribution itself.

The challenge now arises in choosing appropriate values for α and β such that the residual, smoothness, and compactness terms in the objective function are balanced. Due to the separability of the least-squares problems, α and β can be optimized over the linear and non-linear least-squares sub-problems, respectively.

For the selection of β, we require a parsimony criterion that balances the data agreement and identifiability in terms of parameter estimability. This is in analogy to the selection of α in regularization methods [39,56], or model selection in multi-Gaussian models [17] using AIC or related criteria, but with a crucial difference. The AIC and other related criteria balance data agreement with model complexity in terms of the number of model parameters, whereas β must be selected to balance data agreement with parameter estimability. For this, we use the value of β that minimizes the informational complexity criterion (ICC) [57,58], in the following form

$$\mathrm{ICC}(\beta) = \frac{\|\mathbf{V}_{\mathrm{exp}} - \mathbf{V}(\boldsymbol{\theta}_\beta)\|^2}{\sigma^2} + \mathrm{tr}(\ln(\boldsymbol{\Psi}_{\boldsymbol{\theta}_\beta})) + \ln(\det(\boldsymbol{\Psi}_{\boldsymbol{\theta}_\beta})) \tag{18}$$

where $\boldsymbol{\theta}_\beta$ is the parameter set that minimizes the objective function in Eq. (17) for the given value of β and $\boldsymbol{\Psi}_{\boldsymbol{\theta}}$ is the matrix of coefficients of variation of a parameter set $\boldsymbol{\theta}$ given by

$$\boldsymbol{\Psi}_{\boldsymbol{\theta},ij} = \frac{\boldsymbol{\Sigma}_{\boldsymbol{\theta},ij}}{\boldsymbol{\theta}_i \boldsymbol{\theta}_j} \tag{19}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is the covariance matrix of the objective function $F(\boldsymbol{\theta})$ at $\boldsymbol{\theta}$. The first term of the ICC represents data agreement, while the other two terms represent a form of van Emden complexity [59].

Fig. 7 illustrates the effect of compactness on the analysis of four simulated 4-pulse DEER signals. Without the inclusion of compactness, the spin concentration profiles present flat regions falling below the threshold. The modulation depth profiles also present flat regions and the modulation depth is even non-identifiable in the bottom case. The non-identifiability is reflected in the uncertainty of the background fits and the uncertain long-distance artifacts appearing at the edges of the distribution. In all cases, the introduction of the variance penalty with ICC-selected regularization parameter results in well-defined minima of the profiles of both parameters, indicating that the model has a unique solution. Consequently, the uncertainty in the results is reduced, the parameter estimates are improved, and the distance distributions no longer exhibit artifacts.

The ICC criterion adapts the β value to the width of the distribution. For the narrow distribution in Fig. 7 (top), its value β_{ICC} = 0.551 is larger than for the bimodal distribution in Fig. 7 (top middle), where it is β_{ICC} = 0.017. This shows that the addition of the compactness criterion does not automatically artificially compact distributions beyond what is necessary to eliminate long-distance features in P that can be modeled by the background contribution. This is even more evident in the case of the significantly broad distribution in Fig. 7 (bottom middle) which is correctly recovered. However, broad distributions can be prone to artificial compactness in some cases (see later).

Nevertheless, the use of compactness inevitably introduces a certain bias towards a type of distribution. In principle, this criterion can exclude solutions such as the spread-out distribution in Fig. 6, similarly to how the smoothness criterion can exclude distributions with sharp features.

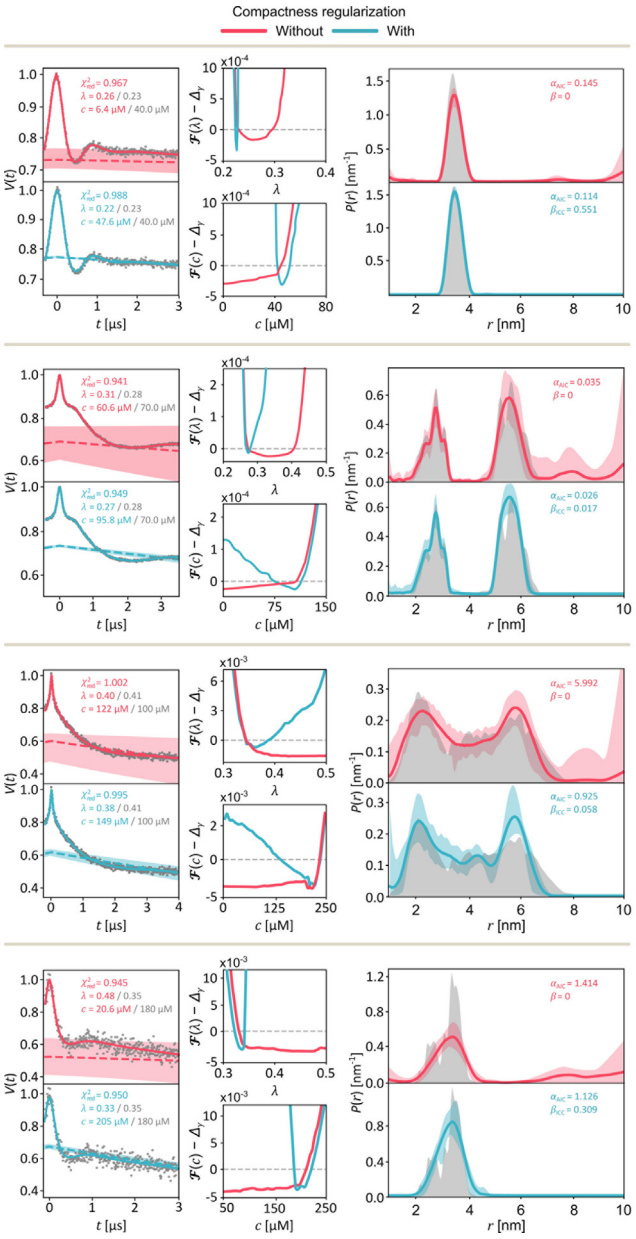

Fig. 7. Compactness regularization in the analysis of 4-pulse DEER data. Four noisy 4-pulse DEER signals simulated from different protein systems (see SI for details) are fitted to a semi-parametric model including the roughness penalty and without (magenta) and with (turquoise) the variance penalty. Each panel contains (left) the data shown as grey dots, the fitted signal shown as colored lines, the background shown as a dashed line, as well as the chi-squared values, the fitted parameter values and the corresponding ground truth values shown as grey text, (middle) objective function profiles of the modulation depth λ and spin concentration c parameters shown as colored lines and the threshold values $\Delta_{\gamma=0.95}$ as dashed grey lines, and (right) the fitted distance distribution as well as the AIC-selected α values and the ICC-selected β values. Each fit is shown as a colored line along their bootstrapped 95%-confidence intervals shown as colored shaded areas. The ground truth distance distribution is shown as a grey shaded area for reference.

7. Performance analysis on a large test set

To assess whether the compactness criterion is robust, we performed a statistical analysis on a large set of dipolar signals. As in previous studies [39,34], we used a library of distance distributions based on MMM [60] simulations of T4 lysozyme (T4L, pdb 2LZM) [39], a small and rigid protein, with relatively short and narrow distance distributions (mean distances of 2–6 nm, inter-quartile ranges of 0.1–0.9 nm) typically encountered in small soluble proteins. We similarly generated a library of distance distributions based on MMM simulations (using the same principles described by Edwards and Stoll [39]) from a preliminary structural ensemble of the protein polypyrimidine-tract binding protein 1 (PTBP1) [61], a large protein with flexible linkers, with relatively long and very

broad distance distributions (mean distances of 2–9 nm, interquartile ranges of 2–6 nm). Both libraries are available in the SI. For each protein, we simulated a set of 5000 dipolar signals from a set of randomly sampled ground truth distance distributions from the corresponding libraries and model parameters randomly sampled from the grids listed in the provided scripts.

For each protein, we fitted 4-pulse DEER models to all 5000 dipolar signals at different values of r_{max} using DeerLab [19]. We repeated the analysis with and without the use of the compactness criterion as described above. We also performed an analysis using the heuristic values of r_{max} given by Eq. (14) as a benchmark against one of the most established approaches. To quantify the quality of the fits, we quantified the error in the fit of the modulation depth and spin concentration parameters by their deviation from the known true values. Similarly, we compared the fitted distance distribution to ground truth through the earth mover's distance (EMD), which quantifies the minimum cost required to transform one distribution into the other by shifting distribution mass. Other metrics such as the overlap index, Kullback–Leibler divergence, and Jensen–Shannon divergence yielded the same conclusions (see the SI).

Fig. 8 shows the results of the statistical analysis. The results illustrate again the dependency of the non-identifiability on r_{max} if compactness is not included. As r_{max} increases, the agreement of the fitted distance distribution with the ground truth quickly deteriorates, as well as the agreement of the fitted modulation depth and spin concentration. The results also show that non-identifiability results in overestimation of the modulation depth and underestimation of the spin concentration, as noted above. Including the compactness criterion leads to fits of higher agreement with the ground truth, generally, as summarized in Fig. 9. The recovery of the ground-truth distance distributions improved considerably for 94.8% (T4L) and 76% (PTBP1) of the fits, whereas for 3.2% (T4L) and 9.1% (PTBP1) no significant changes were observed. Somewhat inferior results are seen in 2.0% (T4L) and 14.9% (PTBP1) of cases. We found no direct correlations between any of the signal simulation parameters (e.g., trace length, noise levels, mean distances, modality, distribution spread, etc.) and the resulting EMD values. However, we observed a bimodality of distributions of the relative EMD values shown in Fig. 9. This bimodality originates from the overlap of a distribution of cases where the compactness criterion works as expected leading to greatly improved results and a narrower distribution of cases where the compactness criterion leads to unchanged or inferior results. The distribution of inferior cases seems to weakly correlate with distance distributions containing the longest distances and short trace lengths. Furthermore, when compactness is included, we observed a strong statistical correlation between inaccurate estimates of the distance distribution (with large EMD values) and significant overestimation of either the spin concentration or modulation depth and underestimation of the other. Therefore, prior information on estimates of the modulation depth and spin concentration can help identify those cases where the compactness criterion is unsuccessful.

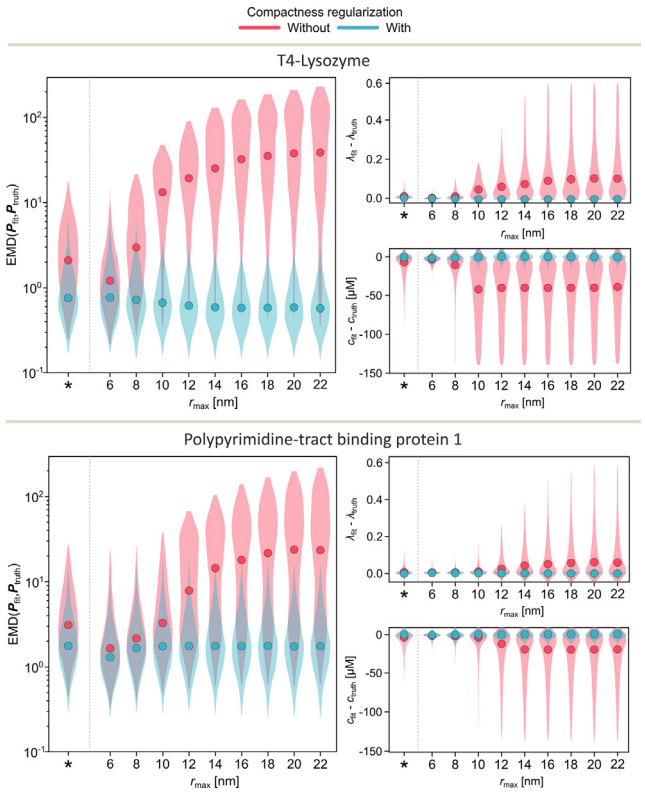

Fig. 8. Statistical analysis of the compactness criterion in the analysis of 4-pulse DEER data. The statistics were obtained from the analysis 5000 different dipolar signals simulated from the structures of T4 lysozyme (top panel) and polypyrimidine-tract binding protein 1 (bottom panel) and analyzed for different values of r_{max}. The analysis with the heuristic r_{max} given by Eq. (14) is denoted by an asterisk. For each signal the agreement between the fitted distance distribution and the ground truth is quantified by the earth-mover distance $\mathrm{EMD}(\mathbf{P}_{\mathrm{fit}}, \mathbf{P}_{\mathrm{truth}})$. The agreement between the fitted modulation depth λ and spin concentration c with their respective ground truths is quantified by their differences $\lambda_{\mathrm{fit}} - \lambda_{\mathrm{truth}}$ and $c_{\mathrm{fit}} - c_{\mathrm{truth}}$. The distributions for each quantity are summarized in violin plots shown as colored areas, with median values indicated wi.th colored dots.

The compactness criterion also shows a better performance in comparison to the established approach of choosing the heuristic r_{max}. Generally, the fits of the distance distribution show a higher agreement with the ground truth for the well structured globular protein T4L than for the partially disordered protein complex PTBP1. The percentage of cases where compactness does not lead to an improvement or leads to an inferior solution strongly correlate with the width of the underlying distance distribution. Such wide distributions are characteristic of the PTBP1 library.

In addition, Fig. 8 shows that the use of the compactness criterion gives results that are overall independent of r_{max}. The analysis is stable even for a vastly overestimated r_{max}, making it safer to use an r_{max} that does not truncate the distance distribution. The overall good performance of this analysis also illustrates the significant advantage of this method towards establishing robust and reliable automated data analysis protocols.

8. Experimental illustration

To illustrate the performance of the compactness regularization on experimental datasets, we use the publicly available 4-pulse DEER datasets measured at Q-band on a modified Yersinia outer protein O (YopO) [63,62] acquired by multiple labs on different constructs. We chose three experimental DEER signals from different constructs measured on different labs based on noise level similarity, long trace length, and the absence of additional pathway contributions. We considerably truncated these signals and analyzed them with and without the compactness criterion. As reference, we analyzed the full-length signals.

Fig. 10 shows the results of this analysis. The compactness criterion on these truncated signals leads to results closer to those obtained when analyzing the entire signals by looking at the background and distribution fits. The introduction of compactness also reduces the uncertainty of the fits. These examples also showcase three possible scenarios. In the middle example in Fig. 10, while the results' quality obtained without the compactness criterion would be acceptable, including the criterion leads to a more certain distribution, particularly towards its longer distances. The left

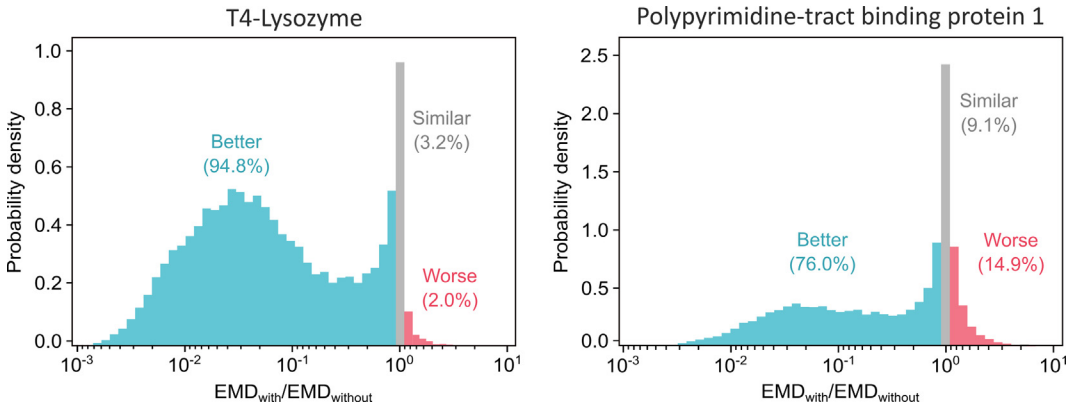

Fig. 9. Statistical analysis of the effects of the compactness criterion on the agreement between the fitted distance distribution and the ground truth. The histograms show the distribution of the relative change in the agreement between the fitted distance distribution and the ground truth quantified by the earth-mover's distance (EMD) over all evaluated r_{max} values. The percentages of cases where the introduction of the compactness criterion leads to a better (EMD is reduced by more than 1%), worse (EMD is increased by more than 1%) or similar (EMD changes by less than 1%) are shown next to the histograms. The colors indicate the type of change: improvement (turquoise), deterioration (magenta) or similar (grey).

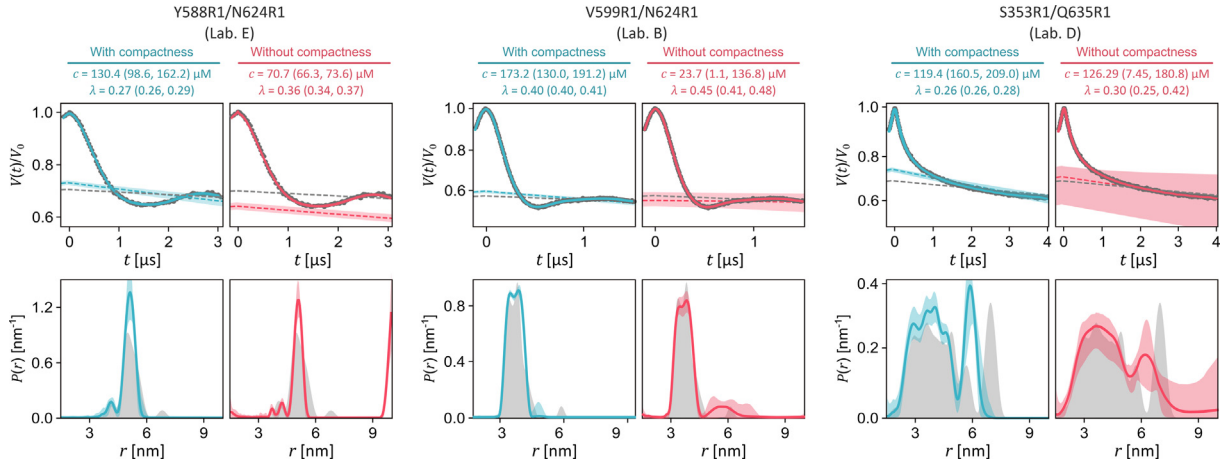

Fig. 10. Analysis of experimental 4-pulse DEER data from different constructs of a modified Yersinia outer protein O (YopO). For each construct, one of the experimental datasets published by Schiemann et al. [62] was truncated and is shown in the top panels as grey dots. The truncated datasets were analyzed with (turquoise) and without (magenta) the compactness criterion. The fits of the dipolar signals and distance distributions are shown as colored lines, and the corresponding 95% bootstrapped confidence intervals are shown as colored areas. The scaled background fits are shown as colored dashed lines. For reference, the fits of the background and distance distribution obtained using the full-length datasets of 7 μs (left), 5 μs (middle), and 7 μs (right) are shown as grey dashed lines and area.s, respectively.

example in Fig. 10 represents a more severe case which is significantly improved by the use of the variance penalty. The right example in Fig. 10 represents another severe case evidenced by the large long-distance artifact. Here, the use of the compactness criterion results in an arguably inferior fitted distance distribution. This case illustrates the effects we observed in the statistical analysis, namely that the use of the compactness criterion on very broad distributions can result in the elimination of the long-distance distribution mass (leading to an identifiable solution) at the cost of a not-so-obvious bias by strongly reducing the uncertainty in the long-distance range. Nonetheless, we have shown that the compactness criterion results in unique and realistic solutions otherwise.

9. Conclusions

In this work we have examined identifiability in the context of data analysis in dipolar EPR spectroscopy. Non-identifiability manifests as flat regions of the objective function and is detected by examining one-dimensional profiles. Non-identifiable dipolar EPR models typically feature artifacts or large uncertainties in the distance distribution at long distances in conjunction with erroneous background parameters and are sensitive to the choice of the upper distance limit.

To mitigate this and to render the model identifiable, we propose using compactness of the distance distribution P as an additional restraint when extracting distance distributions from the DEER data. Compactness is imposed via an additional penalty in the objective function that is proportional to the variance of the distance distribution. The weight of this penalty is selected via the ICC. We have shown how this compactness improves the quality and robustness of the fits in most, but not all, cases. Additionally, the use of compactness lifts the strong dependence of the results on the upper limit of the distance range.

Ensuring identifiability is an important aspect of any reliable data analysis protocol. It ensures the uniqueness and interpretability of the results. If disregarded, it poses an impediment for reliable automated analysis workflows. Therefore, practitioners should keep an open eye to symptoms of non-identifiability. In the case of non-identifiable models, additional prior information must be incorporated to ensure their identifiability. In such cases, it is important that the prior information is reported as well. This will

generally improve both the quality and reliability of dipolar EPR spectroscopy results.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by the ETH Zurich (ETH-35 18-2, G.J. and L.F.I.) and by the National Institutes of Health (GM125753, S.S.).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jmr.2022.107218.

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, Soviet Phys.- Solid State 23 (1981) 565–569.
[2] R.E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich, H.W. Spiess, Determination of End-to-End Distances in a Series of TEMPO Diradicals of up to 2.8 nm Length with a New Four-Pulse Double Electron Electron Resonance Experiment, Angew. Chem. Int. Ed. 37 (1998) 2833–2837.
[3] 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.
[4] S. Saxena, J.H. Freed, Double quantum two-dimensional Fourier transform electron spin resonance: Distance measurements, Chem. Phys. Lett. 251 (1996) 102–110.
[5] S. Saxena, J.H. Freed, Theory of double quantum two-dimensional electron spin resonance with application to distance measurements, J. Chem. Phys. 107 (1997) 1317–1340.
[6] P.P. Borbat, E.R. Georgieva, J.H. Freed, Improved Sensitivity for Long-Distance Measurements in Biomolecules: Five-Pulse Double Electron-Electron Resonance, J. Phys. Chem. Lett. 4 (2013) 170–175.
[7] P.E. Spindler, I. Waclawska, B. Endeward, J. Plackmeyer, C. Ziegler, T.F. Prisner, Carr-Purcell Pulsed Electron Double Resonance with Shaped Inversion Pulses, J. Phys. Chem. Lett. 6 (2015) 4331–4335.
[8] A. Doll, G. Jeschke, Double electron–electron resonance with multiple non-selective chirp refocusing, Phys. Chem. Chem. Phys. 19 (2017) 1039–1053.
[9] L.V. Kulik, S.A. Dzuba, I.A. Grigoryev, Y.D. Tsvetkov, Electron dipole-dipole interaction in ESEEM of nitroxide biradicals, Chem. Phys. Lett. 343 (2001) 315–324.
[10] S. Milikisyants, F. Scarpelli, M.G. Finiguerra, M. Ubbink, M. Huber, A pulsed EPR method to determine distances between paramagnetic centers with strong spectral anisotropy and radicals: The dead-time free RIDME sequence, J. Magn. Reson. 201 (2009) 48–56.
[11] G. Jeschke, M. Pannier, A. Godt, H.W. Spiess, Dipolar spectroscopy and spin alignment in electron paramagnetic resonance, Chem. Phys. Lett. 331 (2000) 243–252.
[12] M. Di Valentin, M. Albertini, E. Zurlo, M. Gobbo, D. Carbonera, Porphyrin Triplet State as a Potential Spin Label for Nanometer Distance Measurements by PELDOR Spectroscopy, J. Am. Chem. Soc. 136 (2014) 6582–6585.
[13] C. Hintze, D. Bücker, S. Domingo Köhler, G. Jeschke, M. Drescher, Laser-Induced Magnetic Dipole Spectroscopy, J. Phys. Chem. Lett. 7 (2016) 2204–2209.
[14] S. Pribitzer, M. Sajid, M. Hülsmann, A. Godt, G. Jeschke, Pulsed triple electron resonance (TRIER) for dipolar correlation spectroscopy, J. Magn. Reson. 282 (2017) 119–128.
[15] S. Milikisiyants, M.A. Voinov, A.I. Smirnov, Refocused Out-Of-Phase (ROOPh) DEER: A pulse scheme for suppressing an unmodulated background in double electron-electron resonance experiments, J. Magn. Reson. 293 (2018) 9–18.
[16] 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.
[17] R.A. Stein, A.H. Beth, E.J. Hustedt, Chapter Twenty - A Straightforward Approach to the Analysis of Double Electron-Electron Resonance Data, in: P. Z. Qin, K. Warncke (Eds.), Methods in Enzymology, Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, vol. 563, Academic Press, 2015, pp. 531–567.
[18] S.G. Worswick, J.A. Spencer, G. Jeschke, I. Kuprov, Deep neural network processing of DEER data, Sci. Adv. 4 (2018) 1–17.
[19] L. Fábregas-Ibáñez, G. Jeschke, S. Stoll, DeerLab: a comprehensive software package for analyzing dipolar electron paramagnetic resonance spectroscopy data, Magnetic Reson. 1 (2020) 209–224.
[20] S.R. Sweger, S. Pribitzer, S. Stoll, Bayesian Probabilistic Analysis of DEER Spectroscopy Data Using Parametric Distance Distribution Models, J. Phys. Chem. A 124 (2020) 6193–6202.
[21] A.N. Tikhonov, Solution of incorrectly formulated problems and the regularization method, Soviet Math. Dokl. 4 (1963) 1035–1038.
[22] 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.
[23] J.H.A. Guillaume, J.D. Jakeman, S. Marsili-Libelli, M. Asher, P. Brunner, B. Croke, M.C. Hill, A.J. Jakeman, K.J. Keesman, S. Razavi, J.D. Stigter, Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose, Environ. Modell. Softw. 119 (2019) 418–432.
[24] R. Bellman, K.J. Åström, On structural identifiability, Math. Biosci. 7 (1970) 329–339.
[25] J. DiStefano, C. Cobelli, On parameter and structural identifiability: Nonunique observability/reconstructibility for identifiable systems, other ambiguities, and new definitions, IEEE Trans. Autom. Control 25 (1980) 830–833, Conference Name: IEEE Transactions on Automatic Control.
[26] C. Cobelli, J.J. DiStefano, Parameter and structural identifiability concepts and ambiguities: a critical review and analysis, Am. J. Physiol.-Regul., Integrative Comparat. Physiol. 239 (1980) R7–R24.
[27] K.R. Godfrey, J.J. DiStefano, Identifiability of Model Parameter, IFAC Proc. Vol. 18 (1985) 89–114.
[28] C. Cobelli, M.P. Saccomani, Unappreciation of a priori identifiability in software packages causes ambiguities in numerical estimates, Am. J. Physiol.-Endocrinol. Metabol. 258 (1990) E1058–E1059.
[29] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, J. Timmer, Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, Bioinformatics 25 (2009) 1923–1929.
[30] F.-G. Wieland, A.L. Hauber, M. Rosenblatt, C. Tönsing, J. Timmer, On structural and practical identifiability, Current Opin. Syst. Biol. 25 (2021) 60–69.
[31] D.R. Kattnig, J. Reichenwallner, D. Hinderberger, Modeling Excluded Volume Effects for the Faithful Description of the Background Signal in Double Electron-Electron Resonance, J. Phys. Chem. B 117 (2013) 16542–16557.
[32] A.D. Milov, Y.D. Tsvetkov, Double electron-electron resonance in electron spin echo: Conformations of spin-labeled poly-4-vinilpyridine in glassy solutions, Appl. Magn. Reson. 12 (1997) 495–504.
[33] Y.E. Kutsovsky, A.G. Mariasov, Y.I. Aristov, V.N. Parmon, Electron spin echo as a tool for investigation of surface structure of finely dispersed fractal solids, React. Kinet. Catal. Lett. 42 (1990) 19–24.
[34] L. Fábregas-Ibáñez, G. Jeschke, Optimal background treatment in dipolar spectroscopy, PCCP 22 (2020) 1855–1868.
[35] G. Jeschke, DeerAnalysis2019 Documentation, 2019.
[36] J. Keeley, T. Choudhury, L. Galazzo, E. Bordignon, A. Feintuch, D. Goldfarb, H. Russell, M.J. Taylor, J.E. Lovett, A. Eggeling, L. Fábregas-Ibáñez, K. Keller, M. Yulikov, G. Jeschke, I. Kuprov, Neural networks in pulsed dipolar spectroscopy: a practical guide, arXiv:2106.07465 [physics] (2021).
[37] G. Golub, V. Pereyra, Separable nonlinear least squares: the variable projection method and its applications, Inverse Prob. 19 (2003) R1–R26, Publisher: IOP Publishing.
[38] D.M. Sima, S. Van Huffel, Separable nonlinear least squares fitting with linear bound constraints and its application in magnetic resonance spectroscopy data quantification, J. Comput. Appl. Math. 203 (2007) 264–278.
[39] T.H. Edwards, S. Stoll, Optimal Tikhonov regularization for DEER spectroscopy, J. Magn. Reson. 288 (2018) 58–68.
[40] T.H. Edwards, S. Stoll, A Bayesian approach to quantifying uncertainty from experimental noise in DEER spectroscopy, J. Magn. Reson. 270 (2016) 87–97.
[41] D.J. Venzon, S.H. Moolgavkar, A Method for Computing Profile-Likelihood-Based Confidence Intervals, J. Roy. Stat. Soc. Series C (Appl. Stat.) 37 (1988) 87–94, Publisher: [Wiley, Royal Statistical Society].
[42] S.A. Murphy, A.W. Van Der Vaart, On Profile Likelihood, J. Am. Stat. Assoc. 95 (2000) 449–465.
[43] D.A. Sprott, Division of Sample Information II: Likelihood Structure, in: Statistical Inference in Science, Springer Series in Statistics, Springer, New York, NY, 2000, pp. 49–72.
[44] C. Kreutz, A. Raue, D. Kaschek, J. Timmer, Profile likelihood in systems biology, FEBS J. 280 (2013) 2564–2571.
[45] S.M. Baker, C.H. Poskar, F. Schreiber, B.H. Junker, A unified framework for estimating parameters of kinetic biological models, BMC Bioinformat. 16 (2015) 1–21.
[46] T. Maiwald, H. Hass, B. Steiert, J. Vanlier, R. Engesser, A. Raue, F. Kipkeew, H.H. Bock, D. Kaschek, C. Kreutz, J. Timmer, Driving the Model to Its Limit: Profile Likelihood Based Model Reduction, PLOS ONE 11 (2016) e0162366, Publisher: Public Library of Science.
[47] W.Q. Meeker, L.A. Escobar, Teaching about Approximate Confidence Regions Based on Maximum Likelihood Estimation, Am. Stat. 49 (1995) 48–53.
[48] E.L. Lehmann, G. Casella, Asymptotic Optimality, in: Theory of Point Estimation, Springer Texts in Statistics, Springer, New York, NY, 1998, pp. 429–519.
[49] S. Brandon, A.H. Beth, E.J. Hustedt, The global analysis of DEER data, J. Magn. Reson. 218 (2012) 93–104.

[50] G. Jeschke, Y. Polyhach, Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance, PCCP 9 (2007) 1895.
[51] M. Azarkh, M. Drescher, Long-Range Distance Constraints in Biomacromolecules by a Combined Approach of Site-Directed Spin Labeling and Double Electron-Electron Resonance (DEER) Spectroscopy, in: G.A. Webb (Ed.), Modern Magnetic Resonance, Springer International Publishing, Cham, 2018, pp. 1013–1030.
[52] T. Schmidt, M.A. Wälti, J.L. Baber, E.J. Hustedt, G.M. Clore, Long Distance Measurements up to 160 Å in the GroEL Tetradecamer Using Q-Band DEER EPR Spectroscopy, Angew. Chem. (International Ed. in English) 55 (2016) 15905–15909.
[53] A. Scherer, S. Tischlik, S. Weickert, V. Wittmann, M. Drescher, Optimising broadband pulses for DEER depends on concentration and distance range of interest, Magnetic Reson. 1 (2020) 59–74, Publisher: Copernicus GmbH.
[54] Y. Polyhach, A. Godt, C. Bauer, G. Jeschke, Spin pair geometry revealed by high-field DEER in the presence of conformational distributions, J. Magn. Reson. 185 (2007) 118–129.
[55] C. Altenbach, LongDistances, 2020. http://www.biochemistry.ucla.edu/Faculty/Hubbell/.
[56] L. Fábregas-Ibáñez, G. Jeschke, General regularization framework for DEER spectroscopy, J. Magn. Reson. 300 (2019) 28–40.
[57] H. Bozdogan, Akaike's Information Criterion and Recent Developments in Information Complexity, J. Math. Psychol. 44 (2000) 62–91.
[58] J.M. Ferguson, M.L. Taper, R. Zenil-Ferguson, M. Jasieniuk, B.D. Maxwell, Incorporating Parameter Estimability Into Model Selection, Front. Ecol. Evol. 7 (2019) 1–15, Publisher: Frontiers.
[59] M.H.V. Emden, An Analysis of Complexity, Mathematisch Centrum, 1971.
[60] G. Jeschke, MMM: A toolbox for integrative structure modeling, Protein Sci. 27 (2018) 76–85.
[61] C. Gmeiner, Integrative Structure Modelling Based on EPR Distance Restraints Uncovers the Role of PTBP1 in the Ires-Mediated Translation Initiation on EMCV, Doctoral Thesis, ETH Zurich, 2018.
[62] O. Schiemann, G. Jeschke, Original data sets of a DEER/PELDOR ring test of four doubly spin-labelled mutants of the protein YopO, 2021.
[63] 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, G. Jeschke, Benchmark Test and Guidelines for DEER/PELDOR Experiments on Nitroxide-Labeled Biomolecules, Journal of the American Chemical Society (2021). Publisher: American Chemical Society.