Applied Magnetic Resonance
https://doi.org/10.1007/s00723-023-01616-w

Applied
Magnetic Resonance

ORIGINAL PAPER

Differentiating Unimodal and Multimodal Distributions in Pulsed Dipolar Spectroscopy Using Wavelet Transforms

Aritro Sinha Roy[1,2] · Jack H. Freed[1,2] · Madhur Srivastava[1,2,3]

Received: 30 July 2023 / Revised: 8 September 2023 / Accepted: 9 September 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2023

Abstract

Site-directed spin labeling has enabled protein structure determination using electron spin resonance pulsed dipolar spectroscopy (PDS). Small details in a distance distribution can be key to understanding important protein structure–function relationships. A major challenge has been to differentiate unimodal and overlapped multimodal distance distributions. They often yield similar distributions and dipolar signals. Current model-free distance reconstruction techniques, such as Srivastava-Freed singular value decomposition and Tikhonov regularization, can suppress these small features in uncertainty and/or error bounds, despite being present. In this work, we demonstrate that continuous wavelet transform (CWT) can distinguish PDS signals from unimodal and multimodal distance distributions. We show that periodicity in CWT representation reflects unimodal distributions, which is masked for multimodal cases. This work is meant as a precursor to a cross-validation technique, which could indicate the modality of the distance distribution.

✉ Madhur Srivastava
ms2736@cornell.edu

Aritro Sinha Roy
as836@cornell.edu

Jack H. Freed
jhf3@cornell.edu

[1] Department of Chemistry and Chemical Biology, Cornell University, Baker Laboratory, Ithaca, NY 14853, USA

[2] National Biomedical Resource for Advanced ESR Spectroscopy, Cornell University, Baker Laboratory, Ithaca, NY 14853, USA

[3] Cornell Atkinson Center for Sustainability, Cornell University, 340 Tower Road, Ithaca, NY 14853, USA

Published online: 22 September 2023

1 Introduction

Protein structure determination remains one of the most challenging and open research subjects. Site-directed spin labeling (SDSL) [1–4] in combination with pulsed dipolar spectroscopy (PDS) has enabled protein structure determination using electron spin resonance (ESR) spectroscopy [5–10]. Typically, a pair of spin probes are attached to the domain of interest in a protein utilizing SDSL. The dipolar coupling between the spin probes at distance r apart is inversely proportional to r^3. Thus, measuring the dipolar coupling by PDS yields inter-spin distance information. Such distances can resolve aspects of protein structures directly and serve as crucial constraints in structural studies.

Proteins are highly dynamic entities and their conformational ensembles give rise to distance distributions between the spin pairs, $P(r)$, rather than a single value of r, as shown in Fig. 1. The process of deriving $P(r)$ from a PDS time domain signal, $S(t)$, is an ill-posed problem [11, 12]. In general, a PDS signal can be expressed as

$$S(t) = \int \mathrm{d}r\, \kappa(t, r)\, P(r), \tag{1}$$

where $\kappa(t, r)$ is the kernel that depends on t and r, averaged over the angle θ between the inter-spin dipolar vector and the direction of the external magnetic field. For all PDS techniques, $\kappa(t, r)$ is singular, and therefore, Eq. (1) cannot be solved for $P(r)$ by a simple inversion of $\kappa(t, r)$. Various techniques have been proposed over the years to derive $P(r)$ from PDS signals, including model-free [13–18], model-based [19–21], and training-based methods [22, 23]. In model-free approaches, such as Tikhonov regularization (TIKR) [13] and SF-SVD [16, 17], distance distributions are heavily reliant on the PDS time-domain signals, as they operate independently of a priori information. Because of the nature of the problem of determining $P(r)$, the solutions raise uncertainties, especially when the $P(r)$ contains weak and/or shoulder peaks. In such cases, often there is no way to confirm whether the solutions truly represent the $P(r)$ or they are artifact driven.

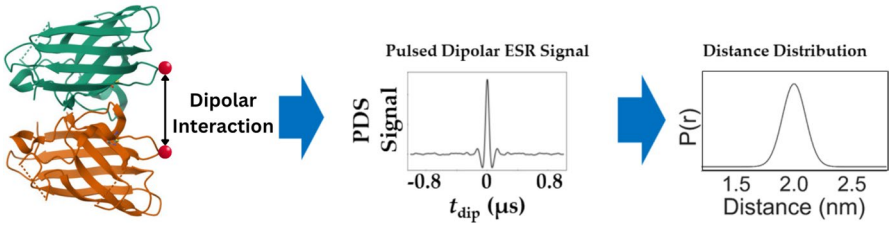

Fig. 1 PDS signals capture the dipolar interaction between a pair of spin labels attached to a protein molecule (PDB code: 3ECV). Post-processing of the signal yields the distance distribution, $P(r)$, between the spin pair

1.1 Major Challenge: Reconstruction of Small Details in *P*(*r*)

The kernel for DEER is given by

$$\kappa(t, r) = \int_0^1 \cos\left[(1 - 3\,\cos^2\theta)\, a\, t/r^3\right] \mathrm{d}\cos\theta, \tag{2}$$

where the dipolar constant, $a = \mu_0\, \mu_\mathrm{B}^2\, g_e^2/2\,h$, μ_0 is the magnetic constant, μ_B the Bohr magneton, g_e the free-electron g value, and h the Planck constant. The DEER signal in its discrete form can be written as

$$S = K\,P, \tag{3}$$

where K and P are the kernel matrix of size $M \times N$ and distance distribution vector with N rows. Note that the expression given in Eq. (3) corresponds to the DEER signal originating from the interaction of an isolated pair of spin-1/2 particles. In a standard DEER experiment, the inter-molecular signal (or background) must be removed first.

The $P(r)$ used in the simulations are shown in Fig. 2. All the distance distributions in model-I (top row) were produced by mixing different Gaussian distributions (shaded area). For model-II (bottom row), both Gaussian and Cauchy distributions were mixed in producing the distance distributions. The probability distribution functions can be defined as: Gaussian distribution, $f_G(r, \mu_r, \sigma_r) = \left(1/\sigma_r\,\sqrt{2\,\pi}\right) e^{-(r-\mu_r)^2/2\,\sigma_r^2}$ and Cauchy distribution, $f_C(r, \mu_r, \Gamma_r) = \Gamma_r/\left[(r-\mu_r)^2 + \Gamma_r^2\right]$. The $P(r)$ for each model is so designed that they are very similar with minor differences. Such small differences in the $P(r)$ can be key in understanding protein structure–function relationships and structural changes. The parameters used in creating the $P(r)$ are shown in Table 1.

It is often challenging to reconstruct such small details in the $P(r)$ with great confidence. The DEER signals utilized for model-I and -II distance distributions are shown in Fig. 3. Visual inspection of the DEER time traces hardly shows any differences, while the differences in their distance distributions are visible in the overlapped plots of the scaled $P(r)$ (the left panel of Fig. 3).

Reconstructions by the SF-SVD [16, 17] and the DEERLab TIKR method [24] were compared with the model distance distributions in Figs. 4 and 5. While both methods captured major parts of the distance distribution patterns, the solutions varied significantly in some cases, e.g., (I.C-D) and (II.G-H) in Fig. 4 and (II.B-D) and (II.F-H) in Fig. 5. This raises considerable doubt over the true nature of the $P(r)$ in such cases. The shaded (gray) region in those figures represents the uncertainty (SF-SVD), the 50% and 95% confidence intervals (DEERLab TIKR). For the SF-SVD solutions, the uncertainty is much less than those of the TIKR, especially for the multimodal distributions. It is visible in both cases, but mainly for TIKR, that the uncertainty is greater in regions near the minor peak positions in the model $P(r)$. More importantly, the 95% confidence interval for TIKR solutions, especially for model-II, shows large uncertainty associated with the solutions. At present, no cross-validation method exists to confirm the existence of multimodal distance distributions with one or more minor (or shoulder) peaks, which is necessary to improve the robustness of PDS analysis.

Table 1 Parameters used in defining the model probability distributions

$P(r)$	Number of peaks	Distribution type	Parameters
I.A	1	Gaussian	μ_r: 3.4 nm, σ_r: 0.25 nm
I.B	2	Gaussian	$\mu_r(1)$: 3.4 nm, $\sigma_r(1)$: 0.25 nm, $\mu_r(2)$: 3.9 nm, $\sigma_r(2)$: 0.1 nm, Gaussian(1):Gaussian(2) = 1.0:0.1
I.C	3	Gaussian	$\mu_r(1)$: 3.4 nm, $\sigma_r(1)$: 0.25 nm, $\mu_r(2)$: 3.9 nm, $\sigma_r(2)$: 0.1 nm, $\mu_r(3)$: 3.15 nm, $\sigma_r(3)$: 0.1 nm, Gaussian(1):Gaussian(2):Gaussian(3) = 1.0:0.1:0.26
I.D	4	Gaussian	$\mu_r(1)$: 3.4 nm, $\sigma_r(1)$: 0.25 nm, $\mu_r(2)$: 3.9 nm, $\sigma_r(2)$: 0.1 nm, $\mu_r(3)$: 3.15 nm, $\sigma_r(3)$: 0.1 nm, $\mu_r(4)$: 2.95 nm, $\sigma_r(4)$: 0.15 nm, Gaussian(1):Gaussian(2):Gaussian(3):Gaussian(4) = 1.0:0.1:0.26:0.15
II.A	1	Gaussian	μ_r: 3.4 nm, σ_r: 0.25 nm
II.B	2	Gaussian + Cauchy	$\mu_r(1)$: 3.4 nm, $\sigma_r(1)$: 0.25 nm, $\mu_r(2)$: 3.7 nm, $\Gamma_r(2)$: 0.3 nm, Gaussian(1):Cauchy(2) = 1.0:0.01
II.C	3	Gaussian + Cauchy	$\mu_r(1)$: 3.4 nm, $\sigma_r(1)$: 0.25 nm, $\mu_r(2)$: 3.7 nm, $\Gamma_r(2)$: 0.3 nm, $\mu_r(3)$: 3.25 nm, $\Gamma_r(3)$: 0.3 nm, Gaussian(1):Cauchy(2):Cauchy(3) = 1.0:0.01:0.015
II.D	4	Gaussian + Cauchy	$\mu_r(1)$: 3.4 nm, $\sigma_r(1)$: 0.25 nm, $\mu_r(2)$: 3.7 nm, $\Gamma_r(2)$: 0.3 nm, $\mu_r(3)$: 3.25 nm, $\Gamma_r(3)$: 0.3 nm, $\mu_r(4)$: 3.1 nm, $\Gamma_r(4)$: 0.3 nm, Gaussian(1):Cauchy(2):Cauchy(3):Cauchy(4) = 1.0:0.01:0.015:0.01

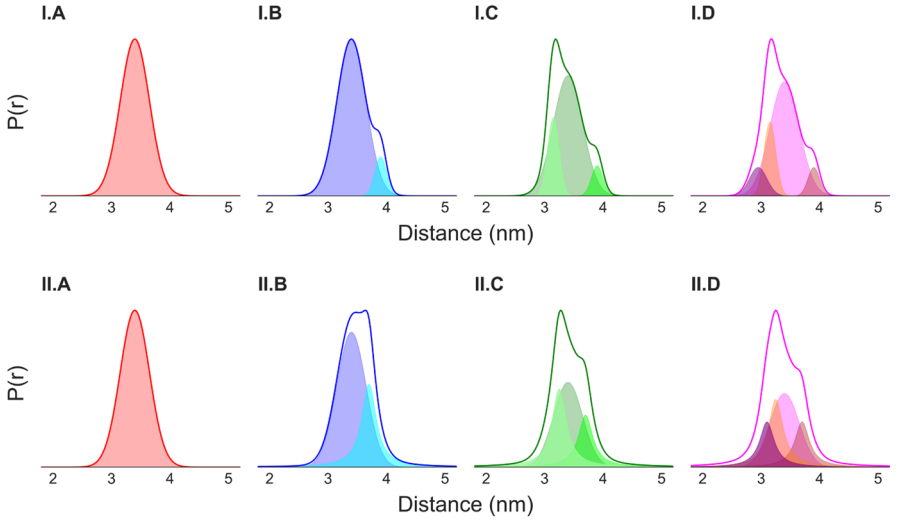

Fig. 2 The different distance distributions, $P(r)$, used in the analysis. The traces show the $P(r)$ and the shaded regions show the different components of a distance distribution. Set-I $P(r)$ were produced by mixing Gaussian distributions and set-II were produced by mixing Gaussian and Cauchy distributions with different means (μ_r), standard deviations (σ_r) or width factor (Γ_r) in different proportions. The parameters are given in Table 1

1.2 Proposed Method

Time-frequency analysis [25–30] is a reliable method to decouple a signal into its distinct constituent components by projecting it on the time-frequency plane. Short Time Fourier Transform (STFT) [25, 26, 29] is another strategy for such analysis, but a fixed window associated with STFT makes it unsuitable for separation of overlapping signal components. Wavelet transform is a powerful method with great flexibility in time–frequency analysis, and hence, it is extremely useful in extracting localized information from various types of signals [31, 32]. We propose the application of continuous wavelet transforms (CWT) in time–frequency analysis to (1) identify differences in $P(r)$ for practically identical PDS signals and (2) confirm the existence of multimodal $P(r)$.

2 Method

2.1 Wavelet Transform

A wavelet transform (WT) can simultaneously represent the time–frequency information for analysis through signal localization and is defined as [33–35]

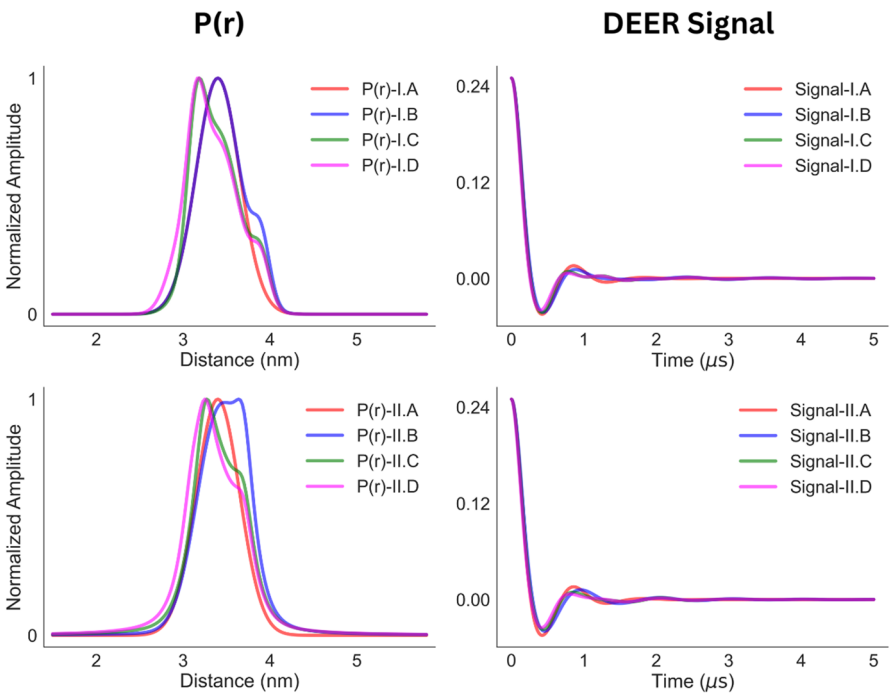

Fig. 3 Two different sets of $P(r)$ (left panel) and the corresponding DEER time traces (right panel) are shown. The DEER signals were calculated with dipolar evolution time of 5 µs and a time increment of 6 ns

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} f(t)\, \psi^* \left(\frac{t-\tau}{s} \right) \mathrm{d}t, \tag{4}$$

where s is the inverse frequency (or frequency range) observing parameter (also called scale parameter), τ is the signal localization parameter (also called translation parameter), t represents the signal location, $f(t)$ is the signal, $F(\tau, s)$ is the wavelet-transformed signal at a given signal localization and frequency, and $\psi^* \left(\frac{t-\tau}{s} \right)$ is the signal probing function obtained from a function called the "wavelet", $\psi(t)$. The functions $\psi(t)$ and $\psi^* \left(\frac{t-\tau}{s} \right)$ are commonly referred as "mother" and "daughter" wavelet, respectively, because $\psi^* \left(\frac{t-\tau}{s} \right)$ is derived from $\psi(t)$. $\psi^*(t)$ is the complex conjugate of $\psi(t)$, which for a real function is the same, $(\psi^*(t) = \psi(t))$.

The Fourier Transform (FT) can be considered as a special limiting case of the WT wherein $s \to (-i\omega)^{-1}$, $\tau \to 0$, and $\psi \to e^{-i\omega t}$. Whereas an FT integrates out the time dependence, the WT is a function of both frequency, s^{-1} and time, τ and thus can display correlations in the signal between them.

Unlike STFT, the WT employs a variable window width and a frequency parameter incorporated in the wavelet, that allows variation in both signal (e.g., time) and frequency. This informs about locations of a particular frequency in the signal domain as

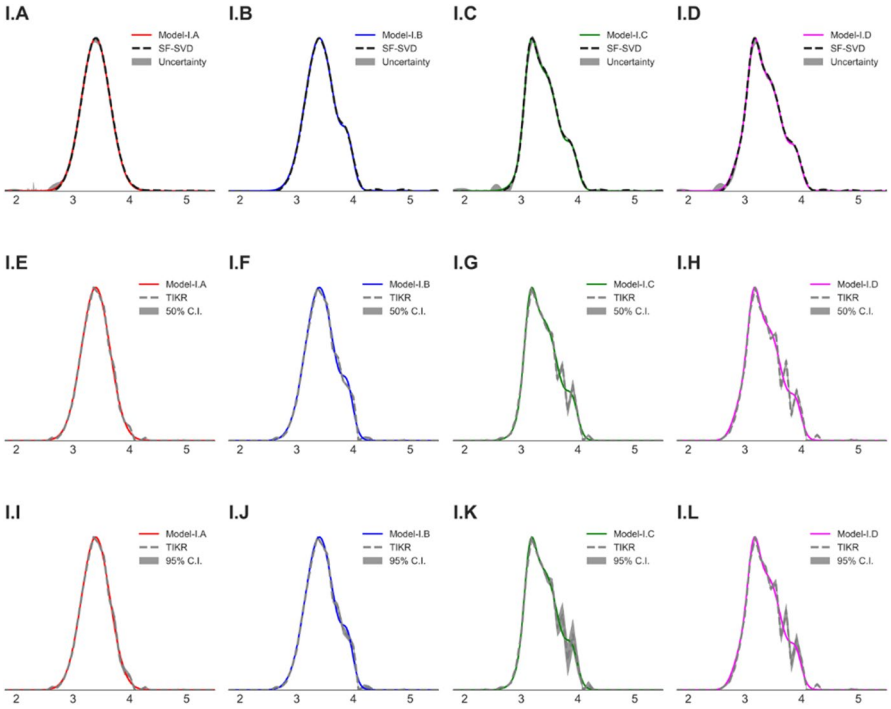

Fig. 4 Model-I $P(r)$ along with the reconstructed distance distributions by the **A**–**D** SF-SVD method and **E**–**L** the DEERLab Tikhonov regularization method (TIKR). The gray shaded regions represent the uncertainty for SF-SVD (**A**–**D**), the 50% (**E**–**H**) and the 95% (**I**–**L**) confidence intervals for TIKR

well as identifies all frequencies that are present at a particular signal location or interval. It results in analyzing a signal into different frequencies at different resolutions, allowing what is known as "multiresolution analysis".

The wavelet-transformed signal $F(\tau, s)$ is represented in the signal domain at a frequency or frequency range, unlike the FT and STFT that represents signal just in the frequency domain. The location of data points in the wavelet domain is spatially correlated with the location of the signal domain. This reveals how a signal looks when observed from a specific frequency or frequency range.

The signal is reconstructed by inverse WT which is given as

$$f(t) = \frac{1}{C_\psi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{s^2} F(\tau, s)\, \psi\left(\frac{t-\tau}{s}\right) \mathrm{d}\tau\, \mathrm{d}s, \tag{5}$$

where C_ψ is admissibility constant which is written as

$$C_\psi = \sqrt{2\pi \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} \mathrm{d}\omega} < \infty, \tag{6}$$

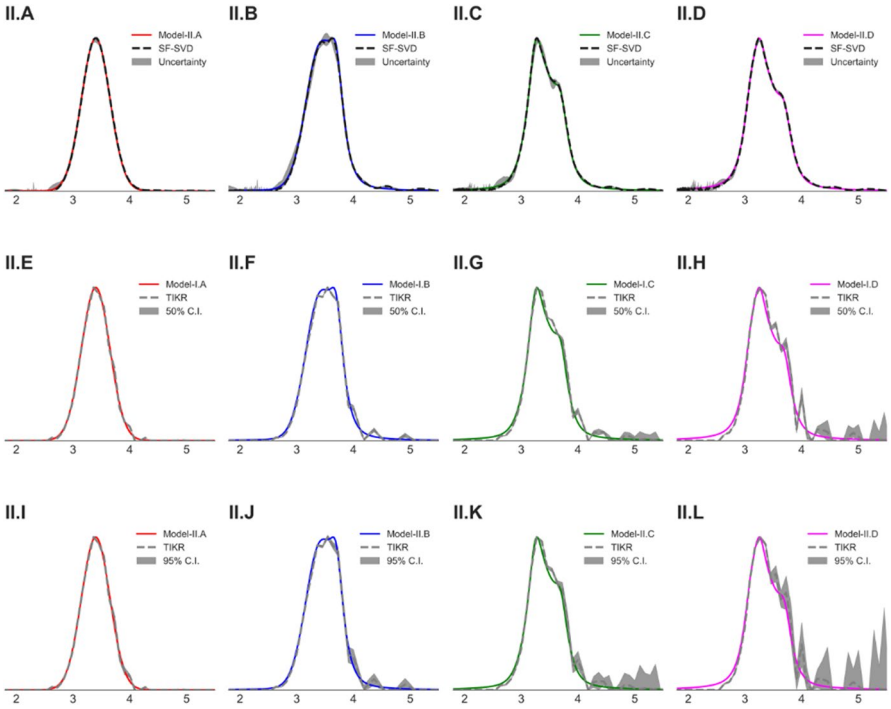

Fig. 5 Model-II $P(r)$ along with the reconstructed distance distributions by the **A**–**D** SF-SVD method and **E**–**H** the DEERLab Tikhonov regularization method (TIKR). The gray shaded regions represent the uncertainty for SF-SVD (**A**–**D**), the 50% (**E**–**H**) and 95% (**I**–**L**) confidence intervals for TIKR

where $\Psi(\omega)$ is the FT of the wavelet function $\psi(t)$. The constraint in Eq. (6) implies that the wavelet function $\psi(t)$ must also be oscillatory with zero mean, i.e., $\int_{-\infty}^{+\infty} \psi(t)\,\mathrm{d}t = 0$.

2.2 Discretized Continuous Wavelet Transform (CWT)

Similar to the Fourier Transform, the WT in Eq. (4) is impractical for discrete data and a discretized version of CWT is used. For practical purposes, the translation parameter and the scale parameter are discretized as $\tau = a$ and $s = b$, a and b both being integers. The CWT of a discrete input signal is defined as

$$C[a,b] = \frac{1}{\sqrt{b}} \sum_{m=0}^{p-1} f[t_m]\, \psi\left[\frac{t_m - a}{b}\right], \tag{7}$$

where $C[a, b]$ is the wavelet-transformed signal and $f[t_m]$ is the discrete input signal.

It should be noted that the discrete wavelet transform (DWT) is computationally more efficient than the CWT and applied more frequently. However, it is appropriate for extracting specific information from a signal. The CWT, on the other hand,

is better suited for scanning all the time–frequency components in a signal for finer details, and hence, it is better suited for this work.

2.3 CWT Time–Frequency Analysis in Python

Time–frequency analysis decouples a signal into its distinct constituent components by projecting it on the time–frequency plane. In this work, we used CWT time–frequency analysis of PDS signals and the Python script for that is as follows.

```
## Python version 3.9,12
# set time domain
tm = 5.0; # in micro second
dt = 0.006; # in micro second
tx = numpy.arange(dt, tm + dt, dt);

# wavelet scale to use
import pywt
scales = np.array([i**2 for i in range(1, 19)]);

# set wavelet
waveletname = 'gaus2';

# calculate the cwt of the signal at given scales
[coef1, freq1] = pywt.cwt(signal1, scales, waveletname, dt)

# contour plots
import matplotlib.pyplot as plt
plt.contourf(tx, np.log(freq1), abs(coef1)**0.5)
```

3 Results and Discussion

3.1 Time–Frequency Analysis of PDS Signals

We calculated the CWT for the simulated DEER traces and plotted those in Figs. 6 and 7. The first and second rows in those figures show the component $P(r)$ and the component DEER traces. Starting with a unimodal distribution (the far left column), minor components comprising Gaussian and/or Cauchy distributions were added to create the other $P(r)$ and DEER signals. The resultant DEER signals were produced by adding the component DEER time-domain signals (third row). In Figs. 6 and 7, the resultant DEER signals are practically identical to those of the dominant component signals in the second rows of the figures because of the dominance of $P(r)$- *I.A* and $P(r)$- *II.A* in all the distance distributions in the two sets of model systems. In

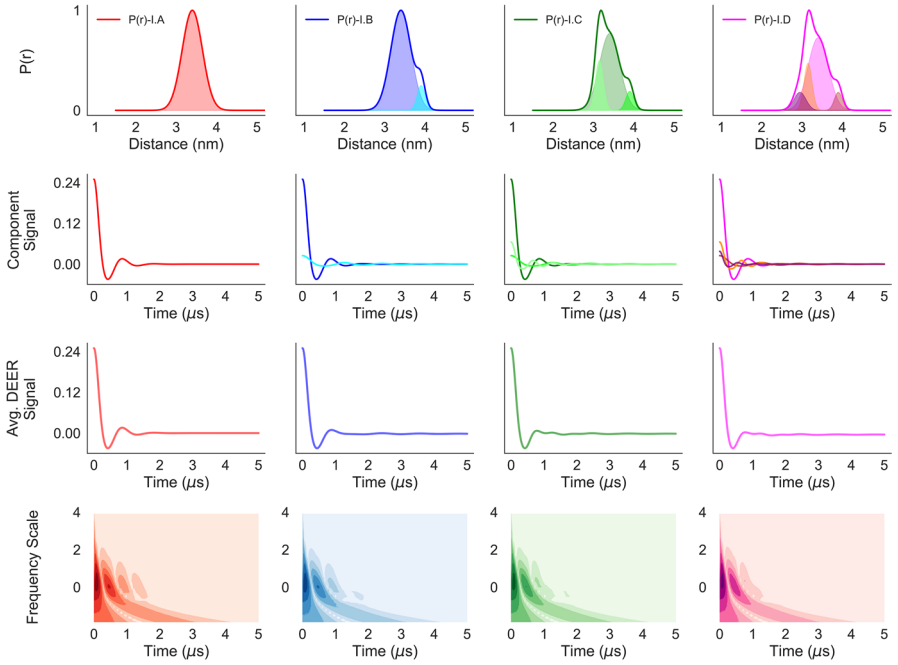

Fig. 6 The similarity of the four DEER traces for model-I systems is reduced via the time–frequency analysis (bottom row) using CWT. The average $P(r)$ (traces) along with the components (shaded distributions) are plotted in the top row. The resulting components of DEER signals and the average DEER traces are shown in the next two rows. The bottom row illustrates the time–frequency contour plots resulting from CWT of the signals using 'Gaus2' wavelet of the pywt Python library

addition, summing of such signals with slightly different time-periods of oscillations causes destructive interference or smoothing of the features associated with the individual components. The bottom rows of those two figures show contour plots of the CWT of the DEER signals for different frequency scales vs. time. The Python script for calculation is given in Sect. 2. It can be noticed immediately that the CWT time–frequency contour plots in both Figs. 6 and 7 illustrate minor, but clearly visible differences, suggesting strong similarity among all the distance distributions, with minor, but detectable differences among all of them. Thus, the time–frequency analysis reveals significant information about different samples prior to the $P(r)$ reconstruction process. On the other hand, such results for identical samples could indicate artifacts, reproducibility issues, and inconsistency in sample preparation.

3.2 Time–Frequency Analysis and the Modality of the Distance Distributions

For a qualitative analysis of the correlation between the differences in $P(r)$ and the corresponding time–frequency contour plots, we have plotted the $P(r)$ and DEER trace components along with their time–frequency plots for $P(r)$- $I.D$ in Fig. 8 (top four rows). The CWT time–frequency contour plots show that both the frequency and

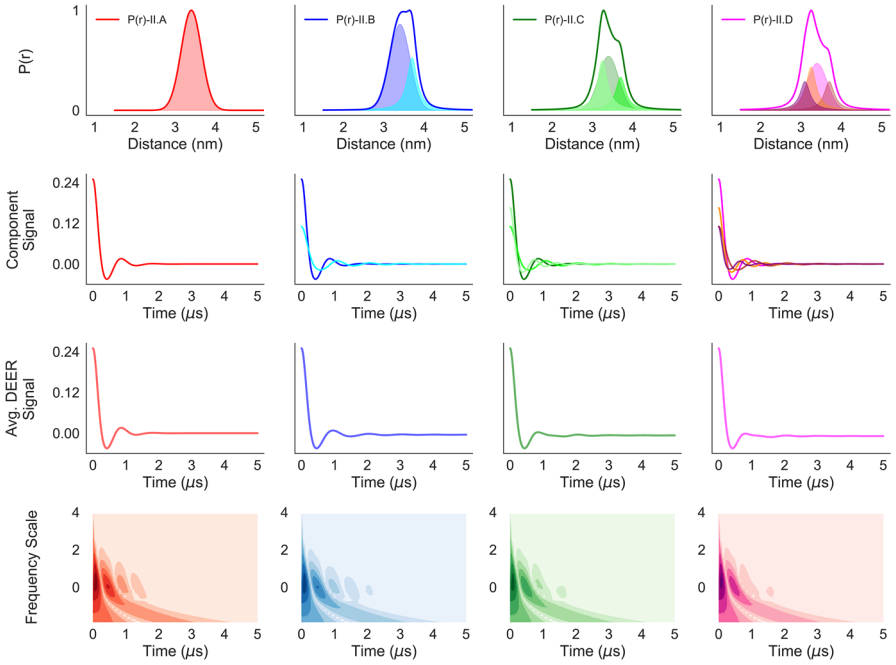

Fig. 7 The similarity of the four DEER traces for model-II systems is reduced via the time–frequency analysis (bottom row) using CWT. The average $P(r)$ (traces) along with the components (shaded distributions) are plotted in the top row. The resulting components of DEER signals and the average DEER traces are shown in the next two rows. The bottom row illustrates the time–frequency contour plots resulting from CWT of the signals using 'Gaus2' wavelet of the pywt Python library

pattern along the time-domain varies with the modal distance of the distribution as well as its width. While the sum of the time–frequency plot shows close resemblance to the top row plot, indicating the dominance of the corresponding $P(r)$ in the mixture, it also demonstrates clear differences. It can be seen that the time–frequency plots of the unimodal distributions have more prominent and periodic features compared to that of the summed signal. In multimodal and overlapping distance distributions, such periodic patterns tend to cancel out. Thus, a time–frequency plot with truncated features suggests the presence of such multimodal and overlapped distance distributions. The same observations are the case for model-II in Fig. 9. The level of loss of the features is proportional to the number of closely spaced modal distances present in a distribution. Hence, it may be possible to train machine learning clustering algorithms against a large dataset of model distance distributions and their time–frequency patterns, which could then indicate the number of modal distances in a distribution.

3.3 Time–Frequency Analysis Using Different Wavelets

In Fig. 10, we repeated the time–frequency analysis for model-I DEER time-domain signals using Gaussian-4 ('Gaus4') and Mexican Hat ('Mexh') wavelets and plotted

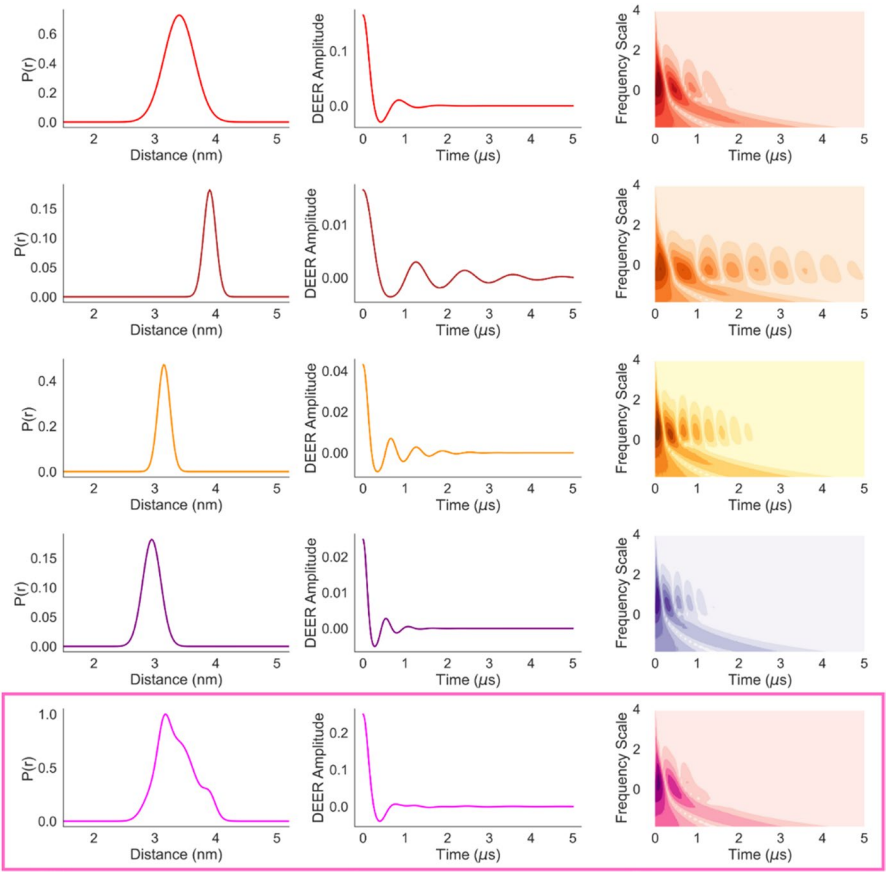

Fig. 8 The four individual components of $P(r)$- $I.D$ (left column), the corresponding DEER signals (middle column), and their CWT time–frequency plots (right column) are presented in the top four rows. The bottom row shows the summed $P(r)$, DEER signal, and its time–frequency plot (magenta rectangular box)

the results along with that of the 'Gaus2' analysis, shown previously. Despite capturing different time–frequency sensitivity, core features of the CWT spectral pattern remain the same. Small variations between the CWT spectra reveal the sensitivity patterns for the different wavelets, but the core features remain the same. Hence, it can been seen that CWT spectra obtain for the dipolar signal is largely independent of the type of wavelet used, demonstrating the robustness of the analysis. We found that among the wavelet families available for CWT time–frequency analysis at present, the Gaussian family is the most suited for the analysis performed in this work. It is worth mentioning that other standard wavelets, such as Coiflet and Daubechies, are not available for CWT analysis in Python or MATLAB at present. Therefore, we plan to develop new wavelets for deeper time–frequency analysis of PDS signals in the near future.

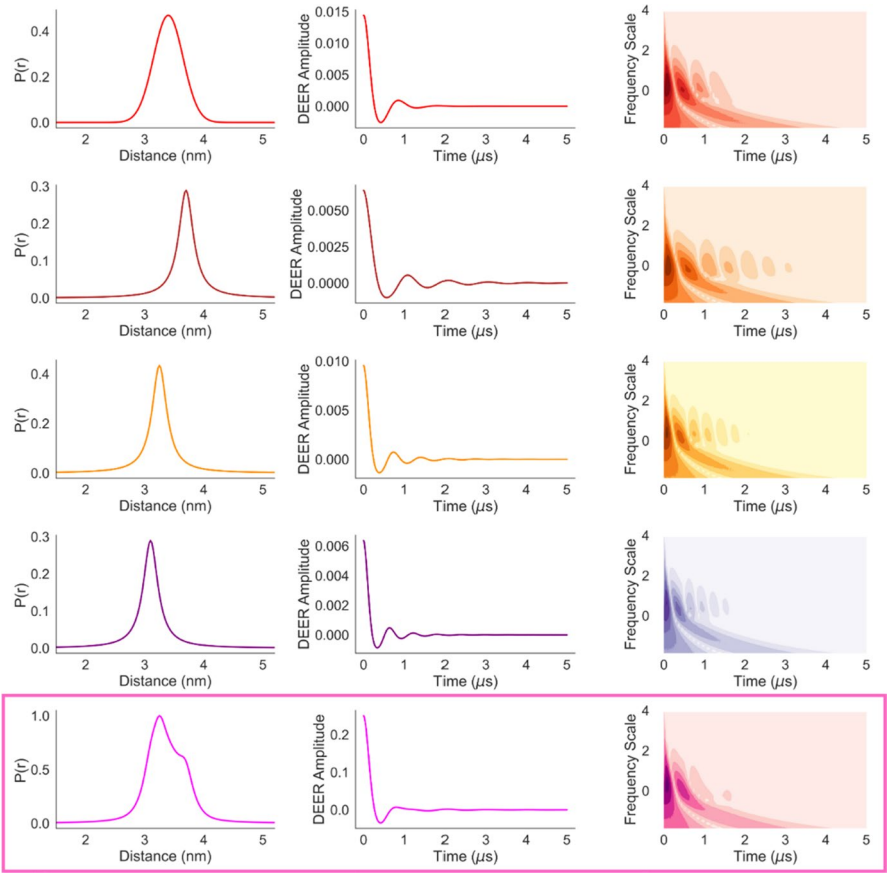

Fig. 9 The individual components of $P(r)$- *II.D* (left column), the corresponding DEER signals (middle column), and their CWT time-frequency plots (right column) are presented in the top four rows. The bottom row shows the summed $P(r)$, DEER signal, and its time–frequency plot (magenta rectangular box)

3.4 Effect of Background Correction Error and Noise

So far, in our discussion, we have used simulated DEER spectra for an isolated pair of spin labels. Background signal originating from inter-molecular interactions is significant in real DEER experiments, while noise is an inherent part of all spectroscopic signals. An error in background correction and the presence of signal noise, both of which can rarely be bypassed, could affect the reconstruction of distance distributions significantly. In the next set of analysis, we simulated DEER signals at various concentrations with background signals and Gaussian noise to achieve a signal-to-noise ratio (SNR) of 63. The DEER signal with background and noise were simulated using the following expression:

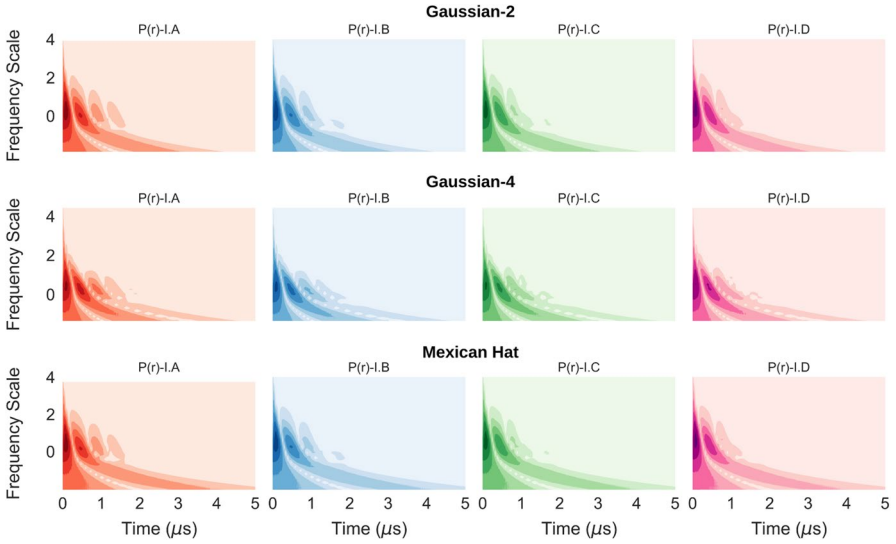

Fig. 10 Time–frequency plots for model-I DEER time-domain signals using three symmetric but different wavelets, namely: 'Gaus2' (top row), 'Gaus4' (middle row), and 'Mexh' (bottom row) wavelets of the pywt Python library

$$V(t) = \left\langle (1-\Delta) + \Delta\, \kappa(t,r)\, P(r) \right\rangle_r \times e^{-\nu\, \Delta\, C\, |t|^{d/3}} + \text{Gaussian noise}, \tag{8}$$

where $S(t)$ is the isolated pair signal expression given in Eq. (1), $\nu \approx 10^{-3}\, \mu\text{M}^{-1}\, \mu\text{s}^{-1}$, the modulation depth Δ (set to 0.5), a parameter associated with the distribution of spin-labeled molecules in the sample d, and the doubly labeled protein concentration C given in μM [36]. We introduced error in background correction using different values of C and d while simulating model-I DEER signals and predicting the background and the parameters are given in Table 2.

We summarized the results in Fig. 11 for model-I distance distributions. We compared the time–frequency plots for the simulated isolated pair DEER signal (I.A3 to I.D3) and the noise-added DEER signals with background error (I.A6 to I.D6). The differentiating time–frequency pattern for the unimodal and multimodal distance distributions show up in both the analysis in the region of frequency scale ≥ 0, while the differences in the latter analysis emerges for frequency scale < 0. We

Table 2 Parameters used in simulating DEER signal with background and noise

$P(r)$	Actual background	SNR	Predicted background
I.A	$C = 100$ μM, $d = 3.0$	63	$C = 85$ μM, $d = 3.0$
I.B	$C = 130$ μM, $d = 2.9$	63	$C = 150$ μM, $d = 3.0$
I.C	$C = 150$ μM, $d = 2.8$	63	$C = 120$ μM, $d = 3.0$
I.D	$C = 180$ μM, $d = 2.85$	63	$C = 207$ μM, $d = 3.0$

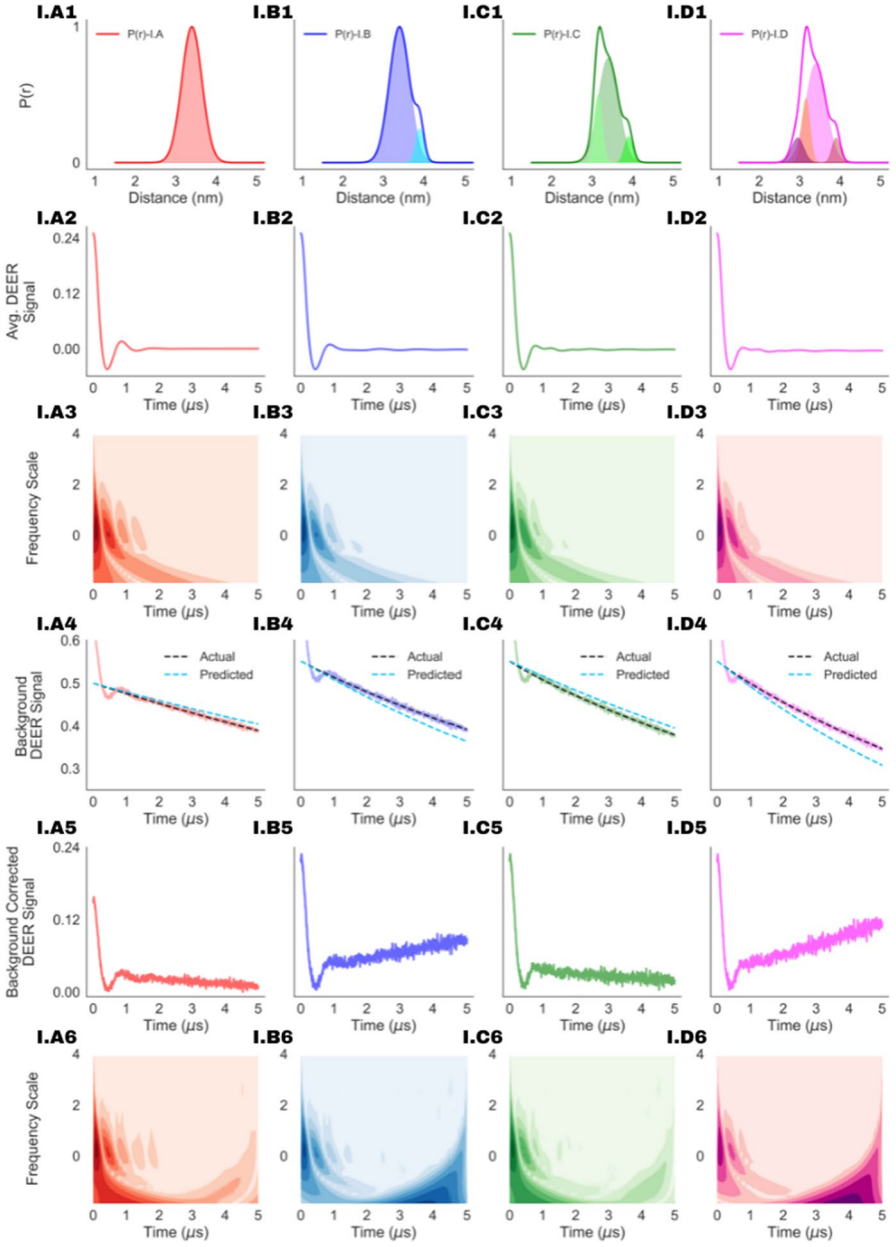

Fig. 11 Simulated noise-free DEER time-domain signals for model-I distance distributions (I.A1 to I.D1). The isolated spin pair signals (I.A2 to I.D2) and the corresponding time–frequency plots (I.A3 to I.D3) are compared with the background corrected signals (I.A5 to I.D5) and their time–frequency analysis (I.A6 to I.D6). The error in the background correction process is demonstrated in (I.A4 to I.D4), which shows the actual (dotted black line) and predicted backgrounds (dotted cyan line). Both sets of time–frequency analysis were done using 'Gaus2' wavelet of the pywt Python library

introduced significant error in concentrations (15–20%) during the background correction to emphasize its effect in the time–frequency analysis. The test confirmed that the time–frequency analysis in differentiating different distance distributions is effectively unperturbed to error in background signal removal and the presence of signal noise.

4 Conclusion

Through SDSL, spin labels are attached to specific domains of a protein, and then, application of PDS yields targeted structural information, i.e., distance distributions between the spin probes. The derivation of distance distributions is an ill-posed problem and in many cases, and the results vary with the methods of reconstruction used in the analysis. In such cases, it is important to cross-validate the results and propose a solution with the least uncertainty. We illustrated in this work that continuous wavelet transform-based time–frequency analysis could be used for distinguishing unimodal and multimodal distance distributions. We used eight model distance distributions and compared the solutions obtained from SF-SVD and the DEERLab Tikhonov regularization methods to illustrate the issue. The CWT time–frequency analysis reliably distinguishes between such very similar PDS signals, indicating the presence of unimodal vs. multimodal distance distributions. This method could be further developed for analysis of PDS signals to cross-validate derived distance distributions and reduce the uncertainty associated with such analysis. In addition to model-free methods that generate distance distributions from dipolar signals, this method can also be potentially used with training-based $P(r)$ reconstruction methods. We demonstrated its robustness against effects such as error in the DEER background correction process and signal noise.

The future work will include development of Coiflet and Daubechies wavelets in CWT. Additionally, we plan to analyze a large dataset of PDS signals and employ appropriate machine learning tools to quantify the correlation between CWT time–frequency patterns and the number of peaks in the distance distributions.

Appendix A: DEERLab Script

```
## P(r) reconstruction by Tikhonov regularization using DEERLab

import deerlab as dl
import numpy as np
import pandas as pd
import os

# filename without file format extension
tag = 'signal2C'

# read experimental data
_mdir = "local repository"
_fname = tag+'.txt'
_data = np.loadtxt(os.path.join(_mdir,_fname), delimiter = '\t')
t, Vexp = _data[:,0], _data[:,1]

# define model
tmin, tmax = 0.004, 2.
r = np.linspace(1,10,100)
expinfo = dl.ex_4pdeer(tau1=tmin, tau2=tmax, pathways=[1])
Vmodel = dl.dipolarmodel(t, r, Pmodel=None, Bmodel=None,
                         experiment=expinfo)

# fit the model
compactness_penalty = dl.dipolarpenalty(None, r, 'compactness', 'icc')
result = dl.fit(Vmodel, Vexp, penalties=compactness_penalty)

# read the P(r)
pr = result.P
_norm = pr.max()
pr /= _norm
pci50 = result.PUncert.ci(50)
pci95 = result.PUncert.ci(95)
```

Author Contributions Conceptualization, ASR, JHF, and MS; methodology, ASR and MS; formal analysis, ASR; writing—original draft preparation, ASR; writing—review and editing, JHF and MS; visualization, ASR; supervision, MS; project administration, MS; funding acquisition, JHF and MS. All authors have read and agreed to the published version of the manuscript.

Funding This research was funded by the NSF Grant No. 2044599 and the National Institute of General Medical Sciences/National Institutes of Health under Grant No. R24GM146107 and R35GM148272.

Availability of Data and Materials Data and the code used in this work have been stored in the public github repository (https://github.com/Signal-Science-Lab).

Declarations

Conflict of Interest The authors declare no competing financial interest.

Ethical Approval Not applicable.

References

1. W.L. Hubbell, H.S. Mchaourab, C. Altenbach, M.A. Lietzow, Watching proteins move using site-directed spin labeling. Structure **4**(7), 779–783 (1996)
2. W.L. Hubbell, A. Gross, R. Langen, M.A. Lietzow, Recent advances in site-directed spin labeling of proteins. Curr. Opin. Struct. Biol. **8**(5), 649–656 (1998)
3. W.L. Hubbell, D.S. Cafiso, C. Altenbach, Identifying conformational changes with site-directed spin labeling. Nat. Struct. Biol. **7**(9), 735–739 (2000)
4. C.S. Klug, J.B. Feix, Methods and applications of site-directed spin labeling EPR spectroscopy. Methods Cell Biol. **84**, 617–658 (2008)
5. P. Borbat, A. Costa-Filho, K. Earle, J. Moscicki, J. Freed, Electron spin resonance in studies of membranes and proteins. Science **291**(5502), 266–269 (2001)
6. P.P. Borbat, H.S. Mchaourab, J.H. Freed, Protein structure determination using long-distance constraints from double-quantum coherence ESR: study of T4 lysozyme. J. Am. Chem. Soc. **124**(19), 5304–5314 (2002)
7. P.P. Borbat, J.H. Freed, *Measuring Distances by Pulsed Dipolar ESR Spectroscopy: Spin-labeled Histidine Kinases* (Elsevier, New York, 2007)
8. G. Jeschke, Distance measurements in the nanometer range by pulse EPR. ChemPhysChem **3**(11), 927–932 (2002)
9. G. Jeschke, DEER distance measurements on proteins. Ann. Rev. Phys. Chem. **63**, 419–446 (2012)
10. A. Schweiger, G. Jeschke, *Principles of Pulse Electron Paramagnetic Resonance* (Oxford University Press, Oxford, 2001)
11. M. Bertero, T.A. Poggio, V. Torre, Ill-posed problems in early vision. Proc. IEEE **76**(8), 869–889 (1988)
12. M. Bertero, *Linear Inverse and iII-posed Problems* (Elsevier, New York, 1989)
13. 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**(2), 279–295 (2005)
14. Y.-W. Chiang, P.P. Borbat, J.H. Freed, Maximum entropy: a complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR. J. Magn. Reson. **177**(2), 184–196 (2005)
15. S. Brandon, A.H. Beth, E.J. Hustedt, The global analysis of DEER data. J. Magn. Reson. **218**, 93–104 (2012)
16. M. Srivastava, J.H. Freed, Singular value decomposition method to determine distance distributions in pulsed dipolar electron spin resonance. J. Phys. Chem. Lett. **8**(22), 5648–5655 (2017)
17. M. Srivastava, J.H. Freed, Singular value decomposition method to determine distance distributions in pulsed dipolar electron spin resonance: II. Estimating uncertainty. J. Phys. Chem. A **123**(1), 359–370 (2018)
18. T.H. Edwards, S. Stoll, Optimal Tikhonov regularization for DEER spectroscopy. J. Magn. Reson. **288**, 58–68 (2018)
19. A.G. Matveeva, Y.V. Yushkova, S.V. Morozov, I.A. Grygor'ev, S.A. Dzuba, Multi-Gaussian Monte Carlo analysis of PELDOR data in the frequency domain. Z. Phys. Chem. **231**(3), 671–688 (2017)
20. I.O. Timofeev, O.A. Krumkacheva, M.V. Fedin, G.G. Karpova, E.G. Bagryanskaya, Refining spin–spin distance distributions in complex biological systems using multi-Gaussian Monte Carlo analysis. Appl. Magn. Reson. **49**, 265–276 (2018)

21. S.R. Sweger, S. Pribitzer, S. Stoll, Bayesian probabilistic analysis of DEER spectroscopy data using parametric distance distribution models. J. Phys. Chem. A **124**(30), 6193–6202 (2020)
22. S.G. Worswick, J.A. Spencer, G. Jeschke, I. Kuprov, Deep neural network processing of DEER data. Sci. Adv. **4**(8), 5218 (2018)
23. J. Keeley, T. Choudhury, L. Galazzo, E. Bordignon, A. Feintuch, D. Goldfarb, H. Russell, M.J. Taylor, J.E. Lovett, A. Eggeling et al., Neural networks in pulsed dipolar spectroscopy: a practical guide. J. Magn. Reson. **338**, 107186 (2022)
24. L.F. Ibáñez, G. Jeschke, S. Stoll, DeerLab: a comprehensive software package for analyzing dipolar electron paramagnetic resonance spectroscopy data. Magn. Reson. (Gottingen, Germany) **1**(2), 209 (2020)
25. C.J. Langmead, B.R. Donald, Extracting structural information using time-frequency analysis of protein NMR data, in *Proceedings of the Fifth Annual International Conference on Computational Biology*, RECOMB01, Montreal, Quebec, Canada, Association for Computing Machinery, New York, NY, United States, (2001), p. 164–175
26. I. Daubechies, The wavelet transform, time-frequency localization and signal analysis. IEEE Trans. Inf. Theory **36**(5), 961–1005 (1990)
27. W. Popiński, Wavelet transform for time-frequency representation and filtration of discrete signals. Appl. Math. **23**(4), 433–448 (1996)
28. R. Constable, R. Thornhill, Using the discrete wavelet transform for time-frequency analysis of the surface EMG signal. Biomed. Sci. Instrum. **29**, 121–127 (1993)
29. T. Brotherton, T. Pollard, R. Barton, A. Krieger, L. Marple, Application of time-frequency and time-scale analysis to underwater acoustic transients, in *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis* (IEEE, 1992), p. 513–516
30. L. Qin, B. He, A wavelet-based time-frequency analysis approach for classification of motor imagery for brain–computer interface applications. J. Neural Eng. **2**(4), 65 (2005)
31. A.S. Roy, M. Srivastava, Hyperfine decoupling of ESR spectra using wavelet transform. Magnetochemistry **8**(3), 32 (2022)
32. A. Sinha Roy, M. Srivastava, Analysis of small-molecule mixtures by super-resolved 1H NMR spectroscopy. J. Phys. Chem. A **126**(48), 9108–9113 (2022)
33. I. Daubechies, *Ten Lectures on Wavelets* (SIAM, New Delhi, 1992)
34. S. Mallat, *A Wavelet Tour of Signal Processing* (Elsevier, New York, 1999)
35. P.S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance* (CRC Press, Boca Raton, 2017)
36. L. Fábregas Ibáñez, G. Jeschke, S. Stoll, DeerLab: a comprehensive software package for analyzing dipolar electron paramagnetic resonance spectroscopy data. Magn. Reson. **1**(2), 209–224 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.